SimAUD 2010

Orlando FL USA

2010 Proceedings of the
Symposium on Simulation for Architecture and Urban Design

Edited by
Azam Khan

2010 Proceedings of the

Symposium on Simulation for Architecture and Urban Design

Edited by

Azam Khan

Cover by **Justin Matejka**
Layout by **Michael Glueck**

Contents

Preface

Azam Khan
Autodesk Research

Welcome to SimAUD, the first annual Symposium on Simulation for Architecture and Urban Design at the 2010 Spring Simulation Conference (SpringSim'10).

Buildings are the largest consumers of energy responsible for 48% of all Green House Gas (GHG) emissions. Due to the complexity and multidisciplinary aspects of architecture design and construction, and urban design and society, modeling and simulation become valuable techniques to understand and optimize this enormous challenge. Although research on building simulation has a long history, simulation research has also progressed significantly over the last few decades. We believe that these areas can contribute greatly to each other, enabling the construction of an advanced systems-based building simulation framework. This ambitious research agenda will require multi-disciplinary cooperation and a more explicit research outcome than these printed proceedings. To this end, in the first year of SimAUD, we have introduced the Data Set track where participants can submit their data set files to be archived at www.simaud.org. As large high-quality data sets are always difficult for researchers to find, we hope this mechanism encourages the community to share complex building models, GIS data, or other helpful data sets. Going forward, we hope that SimAUD will innovate within the research process itself and include additional non-paper tracks to further advance and accelerate progress on holistic systems-based simulation for minimizing resource consumption and GHG emissions, while maximizing occupant comfort.

This new Symposium on Simulation for Architecture and Urban Design offers a venue for architecture researchers and simulation researchers to come together to focus on this important area. By seating SimAUD in the context of simulation research at SpringSim, we hope to start a cross-pollination process that will inform both communities. I would like to thank Gabriel Wainer, Michael Jemtrud, and Ramtin Attar for early discussions about creating a unique concept like SimAUD and for their support in making it happen. Also, I thank Michael Glueck for all of his hard work in designing and building both the website and the proceedings, Justin Matejka for his excellent logo design, and Ruslana Steininger for organizational assistance. The support of Gord Kurtenbach and Jeff Kowalski of Autodesk was invaluable. Finally, a world-class Committee shows that SimAUD is off to a great start.

Azam Khan
SimAUD Founder and 2010 General Chair

All accepted papers will be published in the ACM Digital Library at the SpringSim Archive.
Sponsored by The Society for Modeling and Simulation International.

Beyond Simulation

Multi-Objective Optimization in Architectural Design

Ian Keough

Senior Technical Designer
Buro Happold Consulting Engineers,
ian.keough@burohappold.com

David Benjamin

Adjunct Assistant Professor
Columbia University Graduate School of Architecture, Planning, and Preservation
d@thelivingnewyork.com

Abstract

The challenge of the architect is to create a high-performing building design that is the result of often competing objectives. There are programmatic requirements, aesthetic objectives, and structural criteria which must all be carefully balanced. This paper describes the creation of an automated workflow using parametric modeling, links to structural analysis and a multi-objective optimization engine to act as a tool for the exploration of a wide design space, and as an aid in the decision making process. The design of our custom software CatBot for the linking of Catia and Robot is described, and the further challenge of generalizing the structural inputs, as a set of Catia parameters, to be accessible by students while still providing rigorous structural results is also described. The defining characteristic of this workflow, the ability to trigger topological variation of the model as part of the optimization, is exemplified by *Living Light* in Seoul, South Korea, by the Living Architecture Lab at Columbia University Graduate School of Architecture, Planning and Preservation. This project is presented as a case study.

1. INTRODUCTION

The context for our research includes several recent trends in digital design in the fields of architecture and engineering. First, in architecture, building designs are increasingly subjected to performance evaluation using simulation software such as finite element analysis (FEA) for structure, computational fluid dynamics (CFD) for flow of air and water, environmental analysis for solar gain and energy use, and crowd behavior simulation for safety and evacuation. This is especially true as consultant engineers apply this software and analysis for large and multi-faceted projects. But as simulation software has become more available and more useable, architects themselves are beginning to use it, and they are applying it to smaller projects as well as larger ones.

Second, in architecture, CAD/CAM software is now commonly used to design and fabricate buildings with complex geometry. While an initial era of using *modeling software* to design new forms was followed by an era of adding *manufacturing software* to the process and determining how to fabricate and construct those new forms, the current era is increasingly including *simulation software* and evaluating the performance of complex forms. The integrated use of software for modeling, manufacturing, and simulation allows complex and non-standard designs to be evaluated almost as easily as simple and traditional designs.

Third, in engineering (and in architecture to a smaller extent), evolutionary computation and multi-objective optimization are being used to allow automated algorithmic processes to generate, evaluate, and improve the performance of possible design options. Some projects involve custom-written search algorithms and a single precise fitness criterion. For example, a team of electrical engineers and computer scientists from academia and private industry has designed analog electrical circuits to filter high frequencies for stereo speakers [1], computer scientists at Brandeis University have designed truss-like walking robots to move rapidly in a given direction [2], and civil engineers and a computer scientist at George Mason University have designed steel structures in tall buildings to perform well structurally [3]. Other projects involve existing optimization software and multiple fitness criteria, which are often competing objectives. For example, Bombardier has designed the nose cone of a high speed train to optimize both for drag and for cross-wind stability [4].

Within this context of precedents, we believe our research is unique in several ways. Unlike the majority of current architecture projects that apply simulation software to a small number of designs late in the design process, we have been experimenting with applying simulation software to a very large number of designs early in the design process. We take advantage of two computational features of the simulation software: the ability to simulate the performance of complex design permutations as well as simple permutations, and the ability to automate the simulation process. This potentially allows for the results of the simulation to have a greater impact on the final design result than in traditional design processes.

In addition, unlike previous work in the realm of optimization and evolutionary computation for architectural design, our workflow allows for the use of typical parametric modeling software and typical simulation software without advanced scripting or programming knowledge. In other words, our workflow allows architects themselves to control the processes of simulation and optimization without relegating this work exclusively to consulting engineers. Rather than writing our own search algorithm, we use an existing software application for multi-objective optimization (ModeFrontier) and a custom piece of software for exchanging geometry and performance results (CatBot). By emphasizing flexible and intuitive tools, we are able to engage more non-specialist architects and architecture students, to conduct a wider range of experiments, and to explore more types and combinations of objectives.

Finally, unlike much of the current work with multi-objective optimization in engineering fields that aims to zero in on the single best-performing design, we aim to discover designs that are both high-performing and unexpected. We are interested in using the workflow to achieve performance above a minimum threshold, but also to find designs that are novel and that would not have otherwise occurred to us. While most engineering precedents use an automated search process for the sake of *exploitation*, our experiments use an automated search process for the sake of *exploration*. While most precedents in engineering tend to establish a design space that is limited, continuous, and straightforward to solve, we aim to create a design space that is wide, discontinuous, and difficult to solve. While prior research aims to use multi-objective optimization to create an improved version of iterative design in search of a single solution, we aim to use multi-objective optimization to create a new method of design in search of multiple results that would not have been possible without the computational process—ideally, results that are different from known rules of thumb and different from an initial sketch of the solution by an expert. Moreover, we believe that our workflow is well-suited to the general task of the architect to integrate a wide variety of objectives (goals related to site, program, aesthetics, structure, environment, and many other factors) into a single building design.

In the past two years, we have used our workflow in a series of architecture design studios and technology workshop classes at Columbia University Graduate School of Architecture, Planning and Preservation (GSAPP). This research has shown us that architecture students with no prior experience in computer programming, parametric modeling, or structural simulation are able to learn the process quickly and apply it to a large range of architectural designs with a wide variety of design goals. We have also used the workflow for an advanced research project of the Living Architecture Lab at GSAPP. In this project—called *Living Light* and described below as a case study—we used our automated process to evolve an unexpected, high-performing design. We then verified the performance of the evolved design with more extensive manual simulation. Finally, we constructed the evolved design as a permanent pavilion in a public park in Seoul, Korea.

2. CATBOT

Our workflow includes the use of three existing software applications: Catia [5] for parametric modeling, modeFrontier [6] for multi-objective optimization, and Robot [7] for structural analysis. The unique functionality of our workflow is performed by a piece of custom software, called CatBot, written from scratch in C#.

To facilitate the modeling and optimization process, we designed this piece of software to transfer model geometry from Catia to Robot, conduct a structural analysis, and output results. The challenge was to design this link to be intuitive for students, allow topological variation in the model, and conduct a rigorous structural analysis. Previous workflows combining architectural form optimization and structural analysis have either constrained the user to creation of specific structural types, such as tensile structures [8], or have limited optimization to member sizing or simple member addition or deletion in a topologically fixed system [9]. Our software, CatBot, is available through Google Code [10].

Catia's parameter structure gave us the ability to store values which were not directly related to model elements. A template of structural parameters for loading, slab thicknesses, and section dimensions is included in the Catia model simply to transfer structural information. During every loop triggered by modeFrontier, the geometry of the Catia model is updated according to modeFrontier's design table, then the CatBot node is fired using a batch script. Using the Catia API, CatBot queries the geometry and the parameters values. We then construct a structural analysis model in Robot, apply loads and boundary conditions as described in the parameters, run an analysis, and output the results to a text file.

The Catia workflow was designed to mimic a typical analytical center-line model, comprised solely of linear elements and plates. For this, we reduced the number of elements the students could use to points, lines, and polylines. In order to map analytical information to the model elements we created a simple naming structure where a line might be named "myLine_TUBE01," "TUBE01" designating a section size defined in our parameters. This line might have been created between two points, "point01_FIXED," and "point02_FREE" where "FIXED," designates a point as being fixed in translation, while "FREE" designates a point as free to translate. As Catia is associative, when querying the geometry using the API, a line returns two points, "PtOrigine," and "PtExtremite," corresponding to the start and end point of the line. Parsing the names of these points returns two string values, which should be unique, which can be associated programmatically to integer values corresponding to nodes in the Robot model.

For Catia's line elements, the three available section types are first created in Robot as tube sections with diameters and thicknesses as defined in the Catia parameters. We chose tube sections, as opposed to a section with a primary bending axis (i.e., a wide flange) so that the students would not need to define an angle of rotation of the primary axis of the element. All members are defined as 50ksi steel. By default end releases on all members are defined as fully fixed in translation and rotation. Allowing the students to define the end releases individually per bar would have required an additional portion of the bar naming scheme in Catia to describe the twelve possible degrees of freedom (three rotation and three translation for both the start and the end of the bar). For a given design of the structure you could also not be guaranteed that your choice of end releases, unless varied with each design, would not result in an unstable structure.

Plate elements are defined in Catia as closed polylines, which are then converted to single finite element panels in Robot. By using single finite elements for slabs and panels in Robot we can assure that loads applied to these elements will be conveyed directly to the structure, with no displacement of the FE mesh internal to the panel. The downside to this is that there is no way to carry curved surface information from Catia to Robot. A student defines a polyline in Catia with the name "myPanel_XX," where "XX" corresponds to one of three values: "WL" for wind loading, "LL" for live loading, and "COMB" for a combination of wind and live loading. Wind panels are defined in Robot as negligibly light, in order to reduce the affect of their weight on the structure. Wind forces are applied as a pressure acting along a "wind" vector that is defined by the student in their Catia parameters. Live loaded and combination plates are defined as homogenous reinforced concrete of a thickness defined in the Catia parameters. All live loads are defined as acting in the negative global "z" direction as area loads. An additional point load of 300 pounds is applied to a randomly selected node on the structure during each iteration.

Two additional checks occur during each CatBot iteration that allow for quick rejection of failed designs. Using Robot's "Bar Structure Design," tool, CatBot checks slenderness ratios according to LRFD [14]. Members considered "unstable" are marked as "failed bars" in CatBot. Additionally, Robot nodes whose displacement is greater than the "Nodal Displacement" parameter in Catia are marked as "failed nodes" by CatBot. In this way, students can control the allowable nodal displacement based on the type of structure that they are designing. For example, in long span structures, large nodal displacements would be allowed.

3. CASE STUDY : LIVING LIGHT

In early 2008, the Living Architecture Lab at GSAPP submitted an entry for a design competition sponsored by City Gallery of the Municipal Government of Seoul. The proposal, *Living Light*, involved a pavilion in a public park that glows and blinks according to real-time air quality data as well as public interest in the environment [11]. The pavilion was designed to be a floating dome-like canopy in the shape of the map of Seoul. Each of the city's neighborhoods was a flat polygon of transparent acrylic with etched roads that could be illuminated independently. The size and shape of the canopy was designed according to material properties, lighting performance, and conditions of the site. The steel structural frame for the canopy was designed with a modified space-frame system [12] to allow for laser cutting and non-expert assembly of unique members and joints. The members of the system were deepened to include two bolts at each member connection with top and bottom plates at each node creating a moment connection. The entire canopy surface was then a rigid frame. With the panels and frame of the canopy fixed, the supporting structure was created with a series of branching columns. These columns were constructed from welded square tube sections, and each upper support point was then welded to the nodes of the canopy frame. The column bases were welded to steel plates which were then bolted to the foundation. In order to design the arrangement of branching columns, we used our process of multi-objective optimization and our custom workflow.

For this process, we created a design space with many possible configurations of branching columns as supports for the canopy. For each of the 27 polygons associated with a neighborhood in Seoul, we allowed a column to exist or to be absent. If the column existed, we created a trunk and at least two branches. The trunk was located at the center of the polygon and it extended vertically from the ground up to a "spring point" of a variable height. Each branch extended from the spring point to one of the vertices of the polygon. For example, in a neighborhood at the center of the city represented by a four-sided polygon, the possible configurations were: a tree with one trunk and four branches, four variations of a tree with one trunk and three branches, six variations of a tree with one trunk and two branches, and a configuration with no tree at all. [Figure 1] For feasibility, we did not allow a trunk with no branches, a trunk with one branch, or branches with no trunk. In addition, the spring point could be located at six discrete heights. So for each configuration of trunk and branches, there were actually six different permutations corresponding to the six possible heights of the spring point. In total, there were 67 permutations for this four-sided polygon. Polygons with more sides had more permutations. And each polygon was allowed to vary independently, creating a design space with many millions of possible arrangements of support members.

One of the central challenges of this workflow was keeping the number of inputs to be updated by modeFrontier to a minimum. The workflow would be cumbersome if each potential column or branch location had its own binary variable representing off or on. In order to encode the creation of columns and branches in an integer representation that could be easily updated by modeFrontier, we developed a scheme for mapping all possible configurations of trunks and branches for a panel, or "site," to a binary number. For example, a tree with one column and 5 possible branches can be represented by a binary number with 6 digits, "111111," the first "1" representing the trunk and each successive "1" representing a branch. This maps to integers between 32 and 64. By allowing the genetic algorithm in modeFrontier to choose an integer in this range, the corresponding binary zeroes and ones can turn on or off trunks and branches in each site.

With this setup for the parametric model, we followed our general workflow. The multi-objective optimization application created a random population of 50 designs. At the beginning of each iteration, the Catia model was updated with parameter values in the modeFrontier design table. CatBot was then triggered by modeFrontier to create and evaluate the analysis model. ModeFrontier then mined the results and ranked each design. This process was executed for each design in the generation. After each generation was completed, the multi-objective optimization application used a genetic algorithm to create a new generation with the aim of producing higher-performing results.

Of course, one key element in any process of optimization is defining the objectives. For *Living Light*, we used five outputs and two objectives. Three structural outputs mined from the FEA application—bar failure, node failure, and global extreme forces—were aggregated to create a "combined structural objective." Bar failure was here defined as any bar whose bending stress exceeded the strength of the material (50ksi). Node failure was here defined as the displacement of any node beyond a tolerance set by the user in his structural parameters. For *Living Light*, the maximum nodal displacement was set to 0.5 inches in order to avoid excessive movement of the structure which could damage the acrylic panels. Two geometric outputs read directly from the parametric modeling application—number of trees and sum of the lengths of all branches and trunks—were aggregated to create a "combined material objective." The fitness function included the competing demands of minimizing the "combined structural objective" and minimizing the "combined material objective."

The load combination used to analyze the structure included a snow load of 10psf, a wind load of 30psf, our randomly placed 300lb point load, and the dead load of the structure. All wind and snow loads were applied normal to the panel surfaces.

With this setup, we ran an automated test that produced 25,000 individual designs (500 generations with 50 designs per generation). [Figures 2 and 3] Each design was processed in about 70 seconds and the entire experiment lasted for about twenty days of continuous run-time on a single CPU with a 3.00 GHz processor. At the conclusion of the test, the multi-objective optimization software was used to identify a Pareto frontier—the set of non-dominated designs, or designs that cannot move closer to one objective without moving farther from another objective. It is important to note that the process of multi-objective optimization does not produce a single best-performing design, but rather a set of designs that are all equally high-performing based on the fitness function. Then, further selection criteria—such as aesthetics or further performance analysis—are necessary to select a single design.

Given our set of Pareto designs from the automated test, we manually filtered them to allow only permutations with zero bar failures, zero node failures, and fewer than ten columns. In this narrowed set of designs, there were a variety of formal configurations, but we observed one relatively unexpected trend. In a significant number of designs, the columns clustered in two distinct areas of the canopy. One cluster of 4-5 columns was located at the western perimeter, under the cells lowest to the ground. The other cluster of 4-5 columns was located near (but not at the extent of) the eastern perimeter, approximately under the cells highest from the ground.

Given the observation of this unexpected trend, it was important for us to analyze it in several ways. For our final evolved design, we selected one of the best designs that exhibited this trend and studied its performance in Robot. [13] [Figures 4 and 5] In this design, the structure exhibited a maximum nodal displacement of 0.32 inches, well within our 0.5 inch tolerance. The maximum bending stress within any member of the structure is 6.60ksi, suggesting that member sizes could have been reduced further to take advantage of material strength.

Then, several "fit" versions of the structure that exhibited our observed trend— including our final evolved design—were sent to a local Korean structural engineer who independently validated the results of the analysis.

In addition, in order to understand whether the observed trend of clustering of columns was due to chance, we re-ran the entire experiment two times: once with a different random population of designs in the initial generation, and once with more designs in the initial generation and fewer generations. In both experiments, the clustering trend was replicated.

Finally, we attempted to understand the significance of the clustering trend. We believed that the dome-like shape of the surface grid enhances the spanning potential of the system. As such, one would expect that the final design would have the supports clustered at the farthest edges of the system, allowing the surface to span between them. In fact, this is what happened, but in many cases columns were placed in locations adjacent to the outer-most sites. We believe this occurred because the some of the outer-most sites contain the largest cells in the structure with comparatively long branches which would have increased our combined material objective.

For comparison we created a version of the design in which we created columns at every other exterior cell with two branches each. [Figure 6] The total amount of material is comparable to the evolved design. The maximum nodal displacement of 0.34 inches is similar to the evolved design while the maximum bending stress is 11.54ksi. [Figure 7] almost twice the maximum bending stress in any element of the evolved design. This quick check of the fitness of the evolved design suggests that, at a minimum, our workflow was capable of producing an arrangement of branches which performed comparably to the best "intuitive" design but which was significantly different in form. In other words, we derived an unexpected, high-performing result.

4. CONCLUSION AND FUTURE RESEARCH

In conclusion, we have created and applied an automated workflow that combines parametric modeling, structural analysis, and multi-objective optimization. This process involves existing software applications as well as our custom-written code called CatBot. Though our workflow is still in an early stage of development, we believe that it shows promise for becoming a productive tool for integrating performance evaluation and simulation software into architectural design.

One interesting characteristic of the workflow is its potential ability to generate novel designs. In the case study of *Living Light*, we have seen evidence that our process generated a unique design trend. This trend was different than predicted, but it was understandable—after we observed the generation of high-performing designs, we were able to understand *why* they performed well. The performance of the evolved designs was verified through additional simulation. And the trend was reproduced in additional experiments.

In addition, based on our experience with students at Columbia GSAPP, we believe our process can be used directly by architecture students (and therefore by architects) without extensive knowledge or training. Though we have described only one case study, we have worked with students to apply the workflow to a variety of projects, with a variety of design goals. We have seen that the workflow can aid in the creation of designs that meet specific quantifiable objectives, and that this is especially helpful when the objectives are distinct and competing.

It is important to note that our workflow requires a good *design of experiment* in order to be most successful. We believe that a good experiment involves establishing clear design metrics, defining clear quantifiable objectives (combined into a clear fitness function), and limiting the number of input parameters. Also, good experiments involve a discontinuous design space (one that does not have a simple, obvious solution) and a large design space (one that would be impossible to explore manually).

Furthermore, it is important to be clear that the process does not automatically produce a single answer to a given design problem. Even with a precise fitness function and accurate measurement of the performance of individual designs, interpretation and judgment by the designer are necessary. The process is not meant to remove authorship or agency from design. Instead, it is intended to be used to explore new design possibilities and to directly engage specific aspects of building performance.

Finally, we believe there are several interesting directions to explore in further developing this research. First, we would first like to apply the workflow to more case studies. Also, we would like to conduct experiments that use multiple simultaneous solvers. For example, we would like run experiments in which we subject each design permutation to multiple simulations (CFD, environmental performance, and even crowd behavior, in addition to FEA) during the automated process. We believe this would increase the complexity of competing objectives and allow for more unpredicted solutions. In addition, we would like to incorporate greater use of "generative geometry" into the parametric models in order to create more discontinuous design spaces and produce a wider variety of high-performing designs in a single experiment. Combined with increasingly complex objectives, this could eventually lead to a new aesthetic for designs based on measurable performance. And perhaps most far-reaching of all, we would like to explore the potential for this process to expose values and design goals that are typically hidden, to allow new kinds of debate and discussion about the trade-offs between design goals, and to make the process of architectural design more open and inclusive.

References

1. John R. Koza, Martin A. Keane, and Matthew J. Streeter, "Evolving Inventions," *Scientific American Magazine* (February 2003).
2. H. Lipson and J. B. Pollack, "Automatic design and Manufacture of Robotic Lifeforms", *Nature* (Issue 406, 2000).

3. Rafal Kicinger, Tomasz Arciszewski, and Kenneth DeJong, "Evolutionary Design of Steel Structures in Tall Buildings," *Journal of Computing in Civil Engineering* (Vol. 19, No. 3, July 2005).
4. Unpublished presentation at ModeFrontier User Meeting 2008, Trieste, Italy, October 14-15, 2008.
5. Catia is Copyright Dassault Systemes.
6. modeFrontier is Copyright Esteco.
7. Robot is Copyright Autodesk Inc.
8. Killian, A. and Ochsendorf, J. 2005, "Particle Spring Systems for Structural Form-Finding," *Journal of the International Association for Shell and Spatial Structures*, Vol. 46 (2005) n.147.
9. Von Buelow, P. 2007. "An intelligent Genetic Design Tool (IGDT) Applied to the Exploration of Architectural Trussed Structural Systems." Institute for Lightweight Structures, University of Stuttgart. Ph.D Thesis.
10. CatBot source code available at Google Code: http://code.google.com/p/ikeo-catbot/.
11. *Living Light* web site: http://www.livinglightseoul.net.
12. Trusset system, developed by Phillip Anzalone and the Avery Digital Fabrication Lab at Columbia University Graduate School of Architecture, Planning, and Preservation.
13. The selected design was generated about two-thirds of the way through the experiment. Although this may suggest that the experiment could have been run for fewer generations, in complex problems it is impossible to know with certainty whether an improved design will be generated if the experiment is allowed to continue.
14. American Institute of Steel Construction, *Load & Resistance Factor Design*, Second Edition (American Institute of Steel Construction, 1995), 2-13.

FIGURES

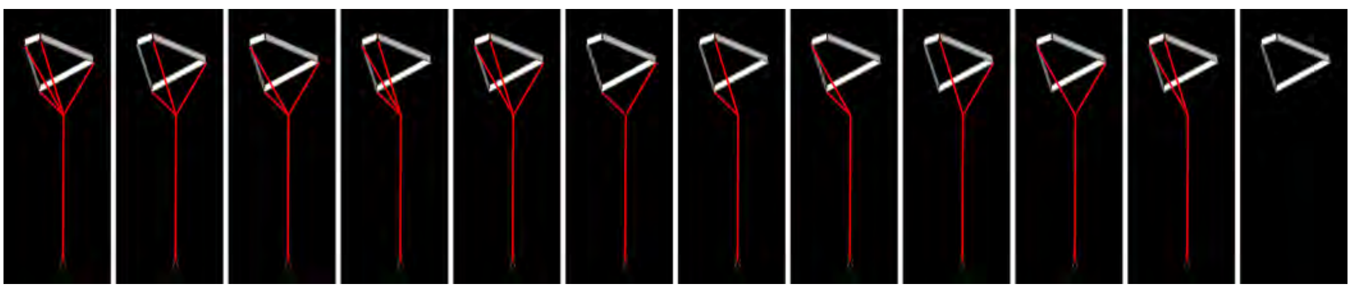


Figure 1. Permutations of a single tree column.

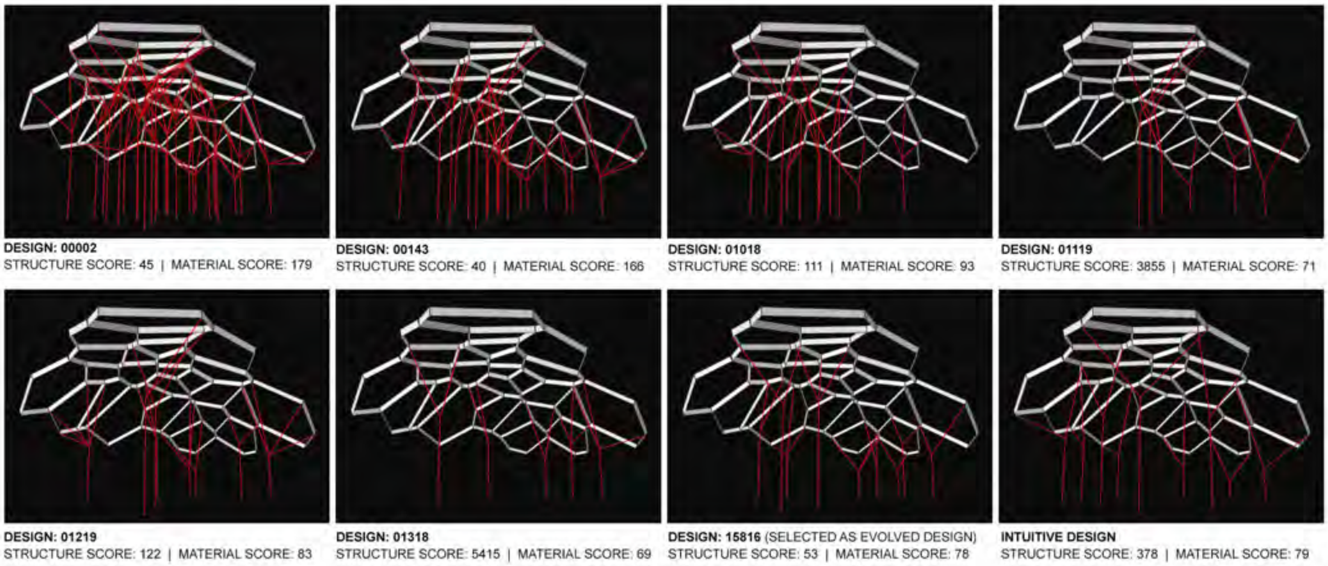


Figure 2. Sample of designs from the automated test.

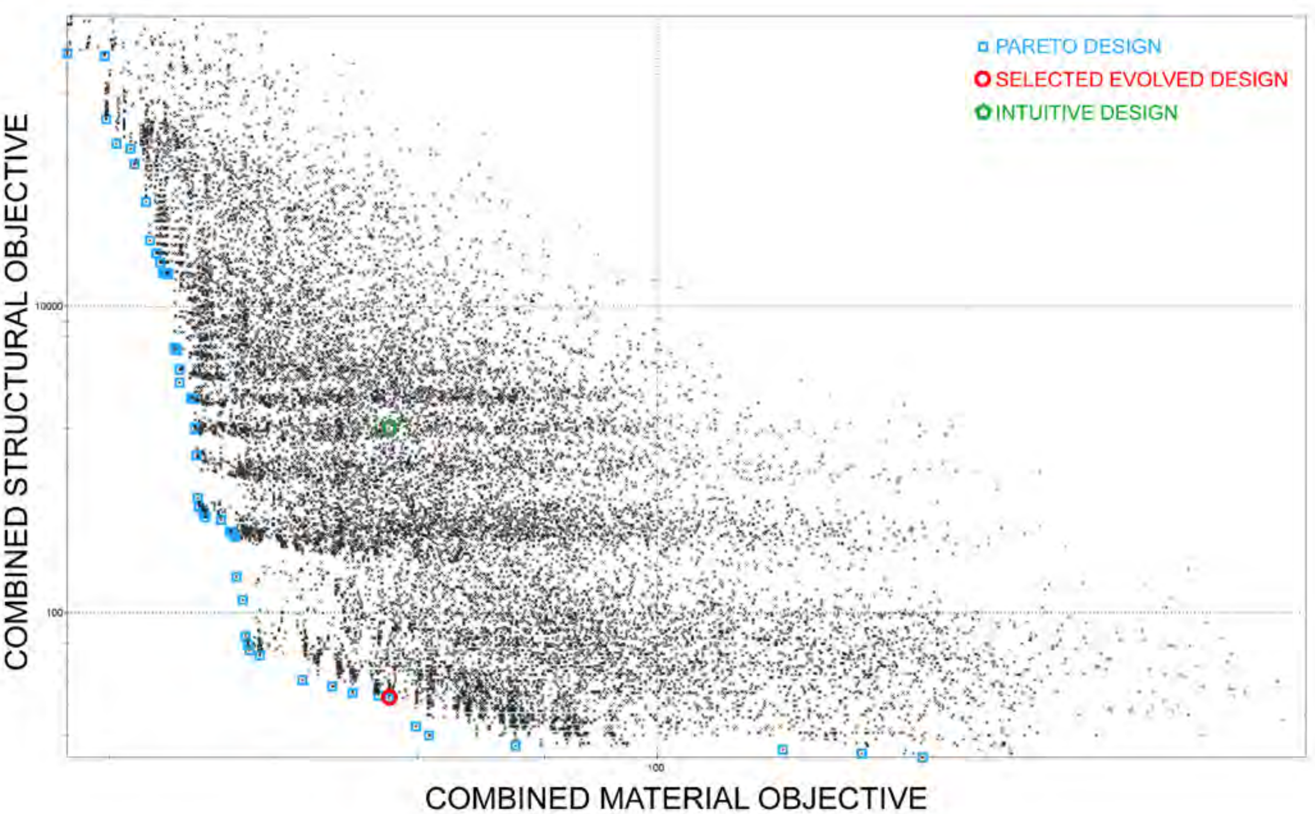


Figure 3. Graph of 25,000 designs.

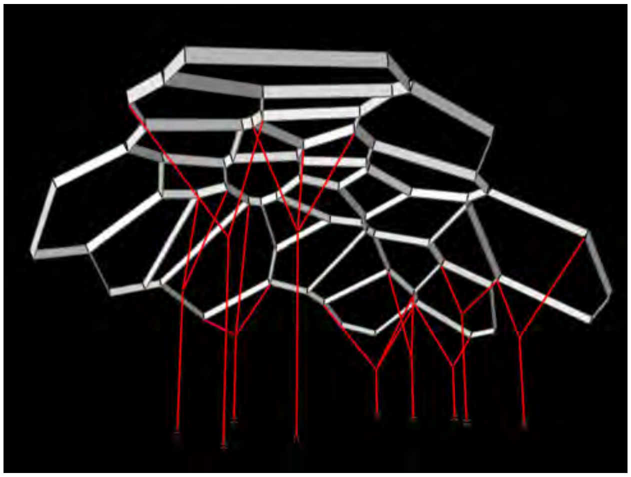


Figure 4. Evolved design selected for construction.

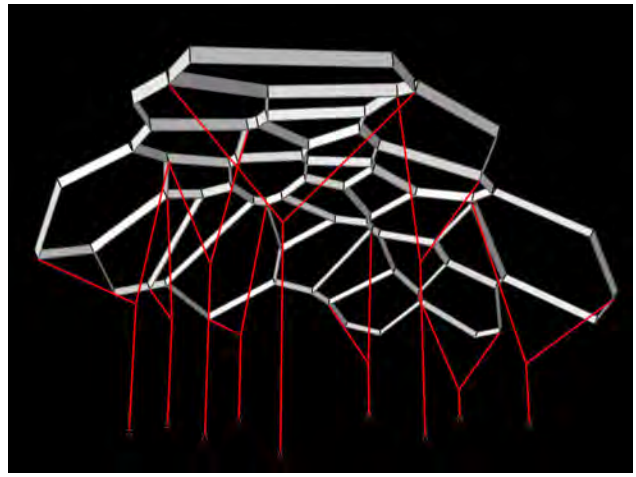


Figure 6. Intuitive design.

Figure 5. Constructed design.

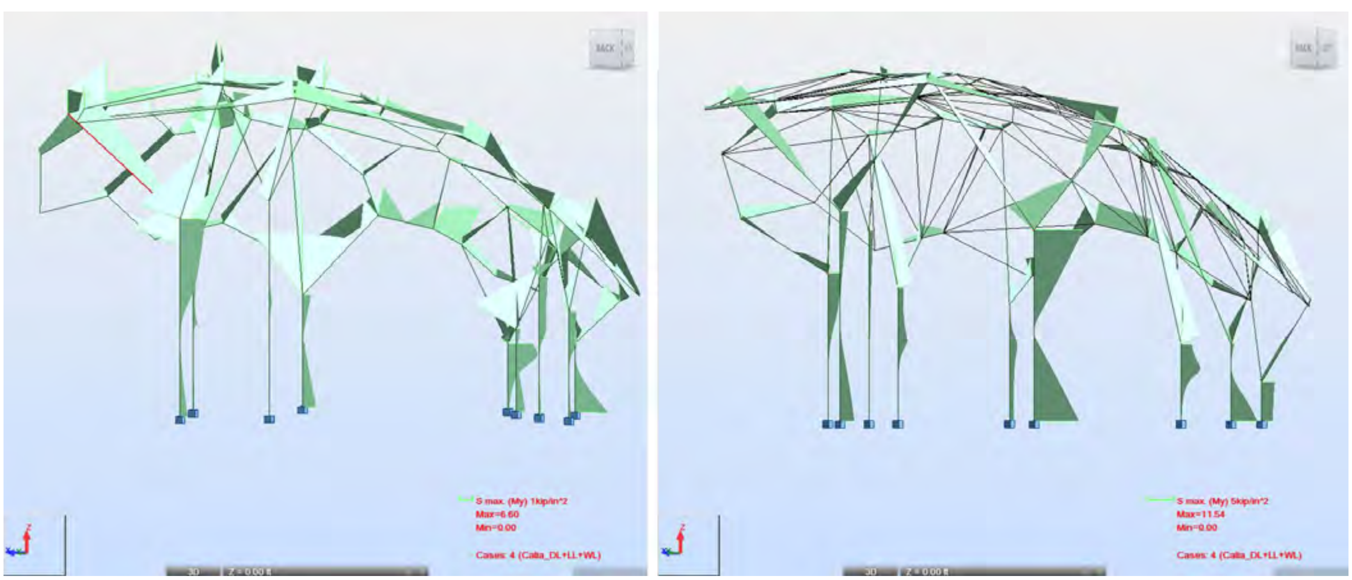


Figure 7. Evolved design bending stress diagram (left) and intuitive design bending stress diagram (right).

Beyond Simulation:
Designing for Uncertainty and Robust Solutions

Sean Hanna
Bartlett School of Grad. Studies
University College London
s.hanna@ucl.ac.uk

Lars Hesselgren
PLP Architecture
LHesselgren@plparchitecture.com

Victor Gonzalez, Ignacio Vargas
Next Limit Technologies
victor@nextlimit.com

Keywords: Design, decision support, interaction techniques, simulation, visualization.

Abstract

Simulation is an increasingly essential tool in the design of our environment, but any model is only as good as the initial assumptions on which it is built. This paper aims to outline some of the limits and potential dangers of reliance on simulation, and suggests how to make our models, and our buildings, more robust with respect to the uncertainty we face in design. It argues that the single analyses provided by most simulations display too precise and too narrow a result to be maximally useful in design, and instead a broader description is required, as might be provided by many differing simulations. Increased computing power now allows this in many areas. Suggestions are made for the further development of simulation tools for design, in that these increased resources should be dedicated not simply to the accuracy of single solutions, but to a bigger picture that takes account of a design's robustness to change, multiple phenomena that cannot be predicted, and the wider range of possible solutions. Methods for doing so, including statistical methods, adaptive modelling, machine learning and pattern recognition algorithms for identifying persistent structures in models, will be identified. We propose a number of avenues for future research and how these fit into design process, particularly in the case of the design of very large buildings.

1. INTRODUCTION

The development of contemporary technologies of simulation has yielded many techniques for deriving results of high quality and precision. As these technologies are predominantly computational, Moore's [1965] law has ensured a steadily increasing speed and precision at an exponential rate that should continue to improve these results into the future, allowing the same tools to apply in greater detail to ever larger projects. But where do we go from here? What are the future directions for research and development in simulation?

Design projects undertaken by architects and engineers are of a scale that is unprecedented in history. Not only do highly programmed buildings such as the international airports of Beijing and Dubai encompass a square kilometre or more of floor area, but the planning of entire new cities, in all their functional complexity, is becoming commonplace, especially in Asia. There are few models for such projects, and little that intuition and experience can hope to contribute, so the need for modelling and simulation in virtually every aspect of the design and planning process has become ever more clear. Moreover, the collaborative teams required to realise these projects are of similarly unprecedented scale, and require effective communication. Their members may be distributed geographically, as well as temporally, and may even change over the duration of the task, making detailed virtual models ever more relevant as a requirement of collaboration.

These factors indicate a greater need for simulation, but also make that simulation far more difficult.
The size and complexity of projects ensures this, as does the obviously disastrous cost of mistakes on such a scale. This position paper outlines current limits or difficulties in the state of the art, then suggests possible solutions and where research efforts should be made.

2. DIFFICULTIES

In practice, there are a number of important limits to what simulation is capable of and how it can be used. Examples are given in this section of a number of current difficulties: resource dependent limits, unknowable design parameters, 'wickedness' of design problems, the process of design in practice, and miscommunication inherent in the use of models.

2.1. Difficult to simulate

For many tasks, the complexity of the situation alone makes simulation exceedingly difficult, either because of the time or resolution required to generate a usable solution.

The Pinnacle, at nearly 300m tall, is designed by Kohn Pedersen Fox to be London's tallest building. In addition to the structural and wind load problems typical of its height, its double-skin of partially overlapping glazing panels introduces additional complexity as it forms a scaled 'snake-skin' of a singly and doubly curved façade. In this case, the effect of the new building on air flow in the area was of concern, particularly the possible impact on pedestrian areas at ground level, which might be adversely effected by winds redirected and amplified by the building's extreme size. In this, a series of computational fluid dynamics simulations were instrumental in guiding the design. Performed in X-Flow by Next Limit, these contributed to a 'skirt' near ground level in which the vertical surfaces of the glazed tower flare outward to form a canopy and redirect air flow at the ground (Figure 1).

But the building skin has other requirements at a finer level of detail that present more difficulties. The double-skin design is intended to perform passive cooling and ventilation via the cavity between the two layers (Figure 2). Outside air is allowed to enter each glazing unit through an opening at the bottom, rises as it draws heat from the building, and is drawn out laterally through the vertical slot between overlapping panels. This flow also is relatively straightforward to simulate, however the building consists of 8,500 units, each with a unique shape angle and position with respect to prevailing winds, and as the local flow of air is altered by any change made to overlapping, neighbouring panels, the evaluation of how any particular design behaves requires the modelling and simulation of air flow with respect to 12,000 independent angles of glazing. Optimisation of the position and angle of each panel required many iterations of this, and thus a cost in computation time of approximately two weeks, each time a significant change was made in the building shape.

Such a task is not uncommon, and a time frame of two weeks is acceptable for occasional testing, but hardly "real-time". The case is particularly noticeable as the overall shape of the building was modelled parametrically using Bentley's Generative Components, and so could be easily modified in many other respects. The optimisation process thus sits somewhat outside the normal process of negotiating the interdependent systems that make up a complex building. Phenomena of much greater precision were ruled out entirely. What about the noise due to the acoustic effect of air flow on each panel? What about rain? These could not have been simulated accurately enough to be of any real use. Although possible in principle, they are at present beyond feasibility due to another level of magnitude in their complexity.

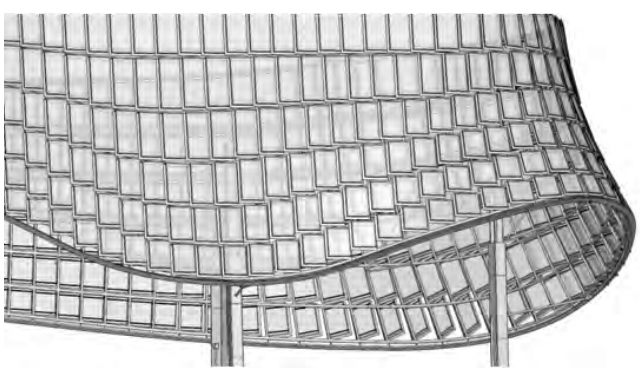


Figure 1. Detail of partially overlapping gazing in the Pinnacle 'skirt'. Image: Kohn Pedersen Fox.

Figure 2. Exterior air enters the cavity through the opening at the bottom (left). Air rises as it heats up and is drawn to the left, exiting through the vertical slots between overlapping planes (right). Image: Kohn Pedersen Fox.

Figure 3. A single simulation gives precise values for wind velocity, but some regions can be particularly sensitive to initial conditions. Image: Next Limit Technologies.

2.2. Impossible to simulate

Compounding the difficulties above, for some design problems we do not even know the values of all the variables involved. In the case of the Pinnacle it is certain that the state of surrounding buildings will change in future, and possibly affecting the air flow drastically enough to make analyses of the current environment obsolete. For many complex problems, the precise states of relevant variables cannot be measured, or (as is the case with many kinds of human behaviour) there is insufficient knowledge on how to even model the system.

Even in the relatively stable state of an unchanged urban environment, most phenomena to be simulated are continuous and can take on any of an infinite range of real values. In such cases, the probability of simulating the exact values for wind speed, direction or other factors approaches zero (Figure 3). This can often pass without causing problems, but if conditions lie within an instability regime, in which a minor change in the wind causes big differences in performance, then the simulation becomes useless.

In practice, one makes a series of best guesses, and then plans for multiple scenarios. A number of other towers are currently planned for the City of London, and one can use the current state of their designs in a model for the vicinity of the building. But these are only coarse guesses, subject to change, and if the result of the simulation is highly sensitive to initial conditions they may not always suffice. Complex phenomena are dependent on many factors, and for many design problems it is impossible to collect all the relevant data at the outset. Unfortunately, this is often just the type of problem designers face.

2.3. 'Wicked' problems

Such difficulties in simulation are made explicit in Rittel and Webber's [1984] definition of the "wicked problem", which by its nature resists any kind of clear definition. Unfortunately, design in disciplines such as architecture and planning is described as dealing almost exclusively with such problems. The brief is relatively ill-defined relative to the real range of problem considerations, the perception of the problem itself may change radically as design progresses, and the solution is typically arrived at by a unique process that cannot be predicted in advance.

Rittel and Webber list ten points that describe the nature of this wickedness. For such problems, the problem domain itself cannot be defined. It is perhaps misleading even to consider the design task as dealing with a *problem* at all, in the traditional sense, as this has no "definitive formulation" [ibid., point 1] in terms of boundaries or objectives. Moreover, wicked problems "do not have an enumerable (or an exhaustively describable) set of potential solutions, nor is there a well-described set of permissible operations that may be incorporated into the plan" [ibid., point 6], so no finite set of rules can be considered that might guarantee they are solved. Even with the fastest computer available, the notion of a solution space can simply expand endlessly.

Even given a wealth of potential solutions, the act of testing them is problematic in itself. Solutions to wicked problems have no "immediate [or] ultimate test" [ibid., point 4]. Aside from the inability to define an objective, the unpredictability of the system in question and emergent nature of its behaviour mean that any proposal may generate repercussions, or "waves of consequences" into a future beyond the point at which the test is made. Moreover, "every solution to a wicked problem is a "one-shot" operation; because there is no opportunity to learn by trial-and-error, every attempt counts significantly" [ibid., point 5]. One therefore cannot experiment with the various possible options, trying out, for example, various versions of a motorway or an urban development, because the cost of these is so high, and each is "essentially unique" [ibid. point 7].

The architect may frequently be of the attitude: "once we define it as problem and solution any competent engineer can deal with it". This is not a comment on the skill level of any member of a particular profession, but merely an observation on the way problems in disciplines are often framed—the classic formulation of engineering problem, or an optimization scenario, is clear. Optimisation might be resource expensive, but it consists simply of defined solution spaces, constraints and performance measures. Unfortunately, these are the wicked part of wicked problems. The bigger and more complex the system, the more we are forced to rely on models to aid in understanding and designing for them, but our certainty about the results of these models also decreases as complexity grows.

2.4. Fit with the design process

The design process is characterised by rapid change, requiring frequent remodelling, re-simulation or re-optimisation. The working relationships of the team both in design and construction also change from project to project, as requirements themselves change, and the structure of communication must be reconstructed to some extent to reflect this.

The features of wicked problems noted above are in stark contrast to the comparatively constrained design domains of, for example, the automotive and aerospace industries—the industries in which optimisation and simulation are used most successfully. These are the industries primarily responsible for the tools (e.g. Abaqus, for finite element analysis and CATIA, for parametric CAD modelling, are both by Dassault, the latter created directly within its aviation division) and currently remain their greatest influence and market. Within these industries, much more is clearly defined and constrained in advance about the design objective, methods of manufacturing and channels of communication. The processes of design and fabrication are usually known, and therefore highly streamlined. A greater proportion of variables are known throughout the process, thereby justifying a greater investment of time and resources to set up a model that is known to be useful at the outset of a new production line. A spoiler on a car, for example, presents a complex aerodynamic situation, but many cars have them and so they are well understood. Architecture has few such spoilers. Where a good deal of systematic, refined and explicit knowledge can be reliably used and traded in these specific domains, the practice is necessarily messy with respect to the built environment.

Attempts have been made to systematise design more generally. In Simon's [1996] proposal for a "science of the artificial", he attempts to rectify the seemingly intuitive and "cookbooky" nature of how it has traditionally been taught and practiced. But such attempts have been opposed, for example by Schön [1983], who criticises this picture of engineering in which problems are well defined and ends are agreed a priori as "Technical Rationality", an essentially Positivist view that is somewhat limited. Schön argues that real design occurs only by an extended practice of re-evaluation and reflection, in which the working definitions of the problems are refined in parallel with their solutions. This cycle of reinterpretation is often observed as an essential feature of design [Snodgrass and Coyne 1999; Lawson 2006] and creativity [Czikszentmihalyi 1988] in practice, and it would seem necessary to support it with different types of tools.

2.5. A false sense of accuracy

The precision of engineering simulations belies the fact that they are ultimately built on statistical measurements of a significant variation of real cases, and this basis may be unknown to the end-user. The simulated behaviour of a single beam of given dimensions is known only because of real tests on many samples of a similar material, with the potential range of structural properties taken into account as a factor of ignorance. But while the models themselves are founded on statistical approximation, this unfortunately doesn't show in the result of simulation. Precise values are given for minimum material tolerances, and these are used in a structural model, along with similarly precise geometry for idealised members, connection details, etc. The result, naturally, is just as precise. For a trained engineer familiar with the factor of ignorance implicit in the original assumptions, the level of real accuracy may be estimated with a little thought, but this can easily be forgotten in practice.

Lawson [2006] gives examples of two types of dangers in how this affects design. The first is that the precision of calculation itself conveys an image of knowledge that does not actually exist. The ease with which computers perform hidden calculation to many decimal places obscures the scientific notion of significant digits, and the polished graphical display can give the uncanny impression of authority. This is typical with students, who "sometimes submit thermal analyses of their buildings […] calculated down to the last watt. Ask them how many kilowatts are lost when a door is left open for a few minutes and they are incapable of answering." [Lawson 2006 p. 70] The unknown quantity of heat loss due to use far outweighs any minor benefits of a few watts here and there, but in this case the designer's ignorance of the major factor in the margin of error is too easily concealed by the clarity of the simulated output. The precision of this output is too often taken to indicate accuracy.

Even if factors contributing to the performance of the design are well known, the second danger is that easily accessible statistics about one factor may influence the designer to emphasise it over more important ones. The kind of precise measurement in Boje's [1971] account of the seconds lost in every opening and closing of an office door, for example, exemplifies a kind of "numerical measuring disease" [Lawson 2006] that might sway a designer toward open plans as they ignore the far more important social and interpersonal factors dependent on spatial separation. This influence may be the result of any of a number of properties inherent to our perception of statistics: they make the factor in question more frequently visible, more explicit and more easily explained to others.

3. POSSIBLE SOLUTIONS & RESEARCH AIMS

In the example of the Pinnacle given above, 8,500 glazing units with 12,000 angles of placement are all unique and require different solutions for optimisation. As a set, however, there are many common features: overall structural hierarchy, materials, but also how they respond to varying wind conditions. There may well be regularities to be found in these that can help in determining the results of the simulation or design optimisation, in the same manner as the statistical regularities assumed in everything from structural capacity of a beam to the variation of annual climate. These regularities are not likely to be statistical, however. In a statistical approach, the variables are defined a priori. Here, and for interesting design tasks, they are unknown at the outset. This section describes directions for research in several areas that may help to deal with the limits and problems of the previous section. Several possible solutions are described: the mapping of broader state spaces, a change in design goals, and 'smarter' modelling techniques.

3.1. Multiple runs of the simulation

The single run of a simulation results in a prediction of behaviour for a precise set of conditions—a specific temperature, humidity, wind direction and velocity, and sunlight, for example, constitute the weather. Designers, however, are usually interested less in this than in climate. Planning for a building that is intended to last for many years, they need a range of varying conditions that must be accommodated, not an instantaneous snapshot of a specific one. In the case of wind, a ‘wind rose’ captures a range of possible input parameters of wind velocity and direction, as they might be distributed probabilistically for a given location. Designing for this range then means running a series of multiple simulations under differing conditions within this given range, and often weighting any conflicting recommendations according to their likelihood or importance.

Multiple runs are often done in an ad-hoc fashion due to time constraints, or more methodically in the case of sensitivity analysis to test perturbations around a single solution under investigation. With enough such runs of a simulation, however, one might begin to build a more systematic overall picture of the effect of a particular design parameter on the behaviour of the system as a whole: The width of building element A has a non-linear but reliable effect on the wind velocity in zone B, at least when the wind direction is between 150 and 230 degrees; If the building remains constant, the wind direction has a quantifiable non-linear relationship with the load on element C, at least below a threshold velocity D. These relationships, as complex as they may be, constitute the state space of the system—an abstract, high-dimensional space in which each point represents a different version of the design, its environment or boundary conditions. Even if the relevant relationships between variables are difficult to know in advance, they can emerge when sufficiently frequent samples are taken. By mapping this in detail, one can discern a great deal more about the kinds of effects that ranges of design choices will have, and the ranges of conditions within which one may operate. Moreover, in the language of complexity science, this state space will likely contain certain regions of divergence and instability, and others that form basins of attraction. Identifying these, and their limits, would allow one to design for stability over time by mapping a (intuitive or systematic) description of the stability and instability of a given configuration.

Many approaches to multi-objective optimisation, including Pareto optimisation [Deb 2001], take a variation on this approach. In these, many solutions are evaluated to determine a range of possible optima that trade one parameter off against another. The final decision as to which solution is used may be deferred to a later time.

It appears that the knowledge gained by many runs of a simulation can have a direct effect on the designer. While the simulation of how people move through spaces is far more complex than the physical behaviour of inanimate systems, Space Syntax methods of analysis [Spiliopoulou and Penn 1999; Hillier and Shu 2001] have proven reliable in doing so. Part of the reason is the acknowledgement that the prediction is ultimately founded on the cumulative results of a vast number of people—in the simulation of visual agents [Turner 2006], a single agent moving through a building will trace a path that appears unlike that of a normal person, however the total effect of a large number of agents in a virtual model will correlate highly with the movement of real people in the actual space. Designers working with such agents in real time have been observed to change their interaction with the developing plan from one of first person manipulation of elements, to one of engagement with or accommodation of the agents themselves. It appears likely that instead of imagining walking a single path through a building as an aid to design, the view of many simultaneous simulations allows the designer to think more abstractly in terms of the overall behaviour relevant to the building.

To fully exploit this exploration of design spaces by multiple simulation, a fuller understanding is required of complex systems in general, and any specific design domain in particular. From its inception over half a century ago, complexity science has explicitly acknowledged the difference between systems that can be reliably predicted statistically, and those complex systems which cannot [Weaver 1948]. Given that the identification of a regularity in a previously unconsidered set of variables may allow the latter to become predictable, this distinction may not be absolute. The relevant questions for any given domain are just what sort of regularities is it possible to find? Structural systems are generally more stable then fluid dynamics, for example. An understanding of what causes phase changes, and what tools, variables and resolutions are most appropriate to model them, is still relatively unexplored in domains relevant to design.

3.2. Change in goals: robustness

If one is to design a built environment that is robust and sustainable as conditions change, the attempt to predict, or futurology, is less tenable than providing an adaptable infrastructure. Designing for a sustainable future is largely about identifying persistent structures across scales—everything from road networks to floor to ceiling heights—that have been viable and robust in the past, and ensuring they continue. The result should maintain adaptability even when more precise predictions inevitably turn out to be wrong.

A change in how we conceptualise our goals for design to explicitly acknowledge robustness in spite of variation may be required. In specifying an invariant objective, optimisation normally targets single optima which may be unstable to perturbation when apparent project goals change rapidly during design or real-world conditions turn out to be somewhat different from those predicted. Somewhat less optimal plateaus of stability are preferable. Technical research required here overlaps with that of the other suggestions in this section: increased computational power allows exploration of search spaces and multi-objective optimisation (§3.1) and structured approximations (e.g. low resolution models) may be derived by running a truncated optimisation process during early stages of design, to be completed in detail later (§§3.3 & 3.4).

3.3. Increased speed and smarter models

Performing optimisation (as in §2.1) or reliably mapping a state space (§3.1) require numbers of simulations of progressively higher orders of magnitude, and making these multiple runs feasible requires faster simulations. These are guaranteed by current trends in the increasing availability of computing power: its cost will continue to decrease (Moore's law [Moore 1965]), and the adoption of grid and cloud (internet based) computing will make better use of it by sharing otherwise dormant resources. But these are only incremental improvements, and a step change from single simulations to an overview of a complex state space requires a vastly larger number of simulations. Moreover, it is likely that constantly growing projects, increased pressure on project timelines and new demands for detail will negate much of this benefit. This is even without the possible counter effects of increased demands from software known as "Wirth's law" [Wirth 1995]. In addition to better hardware, the step change may be affected by the development of smarter models for use in simulation.

In the simplest case these might be based on statistical approximations—low resolution working models, for example. Mesh sizing for finite element analyses takes such an approach in attempting to use the largest element dimensions possible to reliably capture relevant details, thereby increasing resolution in some zones and decreasing it elsewhere [Langham and Grant 1999]. In mapping a large state space of multiple simulations, the basins of attraction or regions of greatest sensitivity to initial conditions may be the same across a broad range of resolutions, and a far lower resolution may be used in some areas.

In more complex cases, patterns or otherwise hidden correlations in data might be found via more advanced statistical techniques or machine learning algorithms. Hanna [2007] uses such an approach in the optimisation of cellular structures consisting of many thousands to millions of unique cells. The modular nature of the individual units provides enough regularity to allow a function to be derived that can replace the optimisation and simulation entirely. A support vector machine is used to map local stress to an optimal cellular structure by training on data taken from between 100 and 600 previously optimised samples. The result can actually improve performance (verified in simulation) over traditionally optimised versions, and increase speed in the order of tens of thousands of times faster. The patterns derived from large data sets may extend to much less clearly defined properties as well. Similar machine learning methods have been used to extract and manipulate arbitrary patterns from spatial arrangements of desks in the workplace [Hanna 2007a] buildings [Laskari et al. 2008] or entire cities. In the latter case [Hanna 2009] the geographical location of a city has been shown to correlate with a number of properties of its form that are non-discursive in the sense that they are not easily spotted or described explicitly by a human observer, but the computer can derive them from plan data and thereby classify cities as to their location with a significant degree of accuracy.

Research into machine learning in general is necessary here, in addition to more domain specific investigation. In many cases, reliable predictions are possible because of the underlying stability of a different part of the system. The consistent patterns predicted by Space Syntax of social interaction [Spiliopoulou and Penn 1999] and crime [Hillier and Shu 2001] in addition to human movement, for example, are due to the relative stability of building layout or street networks over time. Identifying how various subsystems are related for any particular domain will help to identify strategies for design.

3.4. Adaptable, flexible methods of modelling

A design process in which change is frequent (§2.4) and problems are ill-defined (§2.3) would benefit from modelling and simulation methods that can easily and quickly adapt to new problem definitions.

To increase the speed of simulation when ideal levels of resolution are not known in advance, modelling techniques that readily allow changes in resolution may be used. Particle systems may be preferred to meshes of fixed topology, for example, because although the latter allow variation in detail at crucial edges, they are not easy to change over time. In a simulation of air flow over an automobile, Next Limit's X-Flow dynamically updates the number of particles in the system depending on the volume of turbulent air (Figure 4). The process uses a maximum of 50 million particles, but begins with only 5 million for a substantial saving of computation time. This flexibility is potentially more valuable if varying the resolution of particles also becomes possible, with particle (and therefore computation) density increasing over time in turbulent zones where detail is greater and decreasing where it is not needed. This is now being considered for development in the future. The principle can be extended in n-dimensions to the resolution of sampling of state spaces as described above (§3.1).

To fit with the design process and ill-defined problems, there are several ways in which it is possible to use and re-use intermediate results throughout the design process. A typical example of how this is frequently done already is the overall surface shape of a large roof, which might be optimised at a low resolution, while finer details such as the space frame modules that form the actual structure, the dimensions of the structural grid and even the local interruption of the structure by cuts for services might change frequently thereafter. Because their effects on the partial solution are local or non-existent, there is no need to revisit the initial optimisation. The development of a repertoire of such partial solutions is frequently employed in practice.

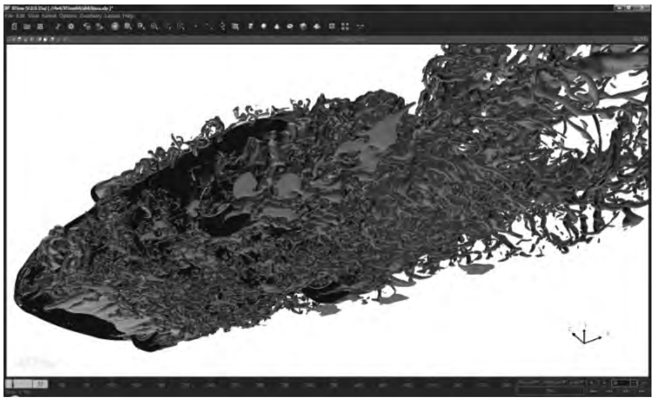


Figure 4. X-Flow dynamically updates the number of particles in the system depending on the volume of turbulent air. Image: Next Limit Technologies.

Figure 5. Multilight renders each light source separately so that they can be adjusted independently after rendering is complete. Different versions of the same scene can be produced without re-rendering. Image: Next Limit Technologies.

Simulation tools can accommodate these. In Maxwell render, Next Limit have developed the 'Multilight', in which each light source is rendered and separately as a partial solution, then to be mixed afterward in the final image. The user is then able to adjust the exact mix of light sources after the fact, in real time, to produce different versions of the scene (Figure 5). This results in increased ease of use for the user, as it does not require decisions about final light to be made early on, but allows a reflective process by which the effect of the light can be seen directly and immediately as the decisions are made later. It also allows far better communication with a client or among design teams. While renderings are often seen as somewhat final, they are crucial to the collective creative process, and such tools encourage engagement.

Basic research is still required. Multilight is possible because light is easily separable, but many design variables are not. The roof example above is strictly hierarchical in that decisions of detail design may be dependent on overall shape, but not vice-versa. Finding the points at which partial solutions to more complex models may be stable enough for re-use is another task for the investigation of state spaces and machine learning research mentioned above (§§3.1 & 3.3). Technical development will then be required in the development of tools that allow the use of partial solutions. This use of simulation and models at intermediate stages, without clearly defined start conditions or a fixed end solution, is intimately related to the way designers work, and ultimately, some re-education of designers themselves may also be required.

4. CONCLUSION

As designers take on larger and more complex tasks, this paper has suggested that the single analyses currently provided by most simulations display too precise and too narrow a result to be maximally useful in design, and instead a broader description of what simulation can do, and how it can be used, is required. It has attempted to outline ways to make our simulations, and our buildings, more robust with respect to the uncertainty we face in design: a better exploration of the range of solutions, changes in how we perceive optimisation goals, and the use of statistical and machine learning algorithms to do so.

In planning efforts for future research, there is certainly the tendency to refine the simulation methods we have to ever finer degrees of precision. This is helpful, but only in context of understanding the real needs of the design tasks. None of the methods in question are in any danger of becoming "technologies looking for an application", but this paper has aimed to present the ways in which they need to be used, in the hope that newly developing technologies will improve designers' understanding of uncertainty and robustness and better equip them to deal with the complex, ill-defined problems they face.

References

Boje A (1971) *Open-plan Offices*. London, Business Books.

Czikszentmihalyi, M (1988) Society, culture, and person: a systems view of creativity, *in* Sternberg, RJ (ed), *The nature of creativity: Contemporary psychological perspectives*, Cambridge University Press, Cambridge, pp. 325-339.

Deb K (2001) *Multi-objective optimization using evolutionary algorithms*, John Wiley and Sons.

Hanna S. (2007) Inductive machine learning of optimal modular structures: Estimating solutions using support vector machines, *Artificial Intelligence for Engineering Design, Analysis and Manufacturing*, 21: 1-16.

Hanna S (2007a) Automated Representation of Style by Feature Space Archetypes: Distinguishing Spatial Styles from Generative Rules, *International Journal of Architectural Computing.*

Hanna S (2009) Spectral Comparison of Large Urban Graphs, In Koch D, Markus L and Steen J, *SSS7: Proceedings of the 7th International Space Syntax Symposium*, Stockholm.

Hillier B and Shu S (2001) Crime and urban layout: The need for evidence, in *Ballintyne S, Pease, K and McLaren V: Secure Foundations: Key Issues in Crime Prevention and Community Safety.* IPPR, London.

Langham AE and Grant PW (1999) Evolving rules for a self-organizing finite element mesh generation algorithm. In *Proc. 1999 Congress on Evolutionary Computation* , pp. 161-168, Washington D.C., USA, July 1999. IEEE Computer Science.

Laskari S, Hanna S and Derix C (2008) Urban identity through quantifiable spatial attributes: Coherence and dispersion of local identity through the automated comparative analysis of building block plans, in Gero JS and Goel A (eds) *Design Computing and Cognition '08,* Springer.

Lawson B (2006) *How Designers Think: The design process demystified, Fourth Edition*. Elsevier, Oxford.

Moore GE (1965) Cramming more components onto integrated circuits, *Electronics* 38(8).

Rittel HWJ and Webber MM (1984) Planning problems are wicked problems. In Cross N (ed.) *Developments in Design Methodology*, John Wiley and Sons.

Schön D A (1983) *The Reflective Practitioner: How Professionals Think in Action*. Basic Books, New York.

Simon HA (1996) *The Sciences of the Artificial.* 3rd ed. MIT Press, Cambridge, MA.

Snodgrass AB and Coyne RD (1997) Is designing hermeneutical? *Architectural Theory Review,* 2(1): 65–97.

Spiliopoulou G and Penn A (1999) Organisations as Multi-Layered Networks: face to face, e-mail and telephone interaction in the workplace, *2nd Intl. Space Syntax Symposium*: 1-24.

Turner A (2006), First steps in evolution of animats for pedestrian simulation. Presented at *Design Computing Cognition '06, Workshop on Models of Machine Learning in Creative Design*, Eindhoven.

Weaver W (1948) Science and Complexity, *American Scientist*, 36: 536.

Wirth N (1995) A Plea for Lean Software, *Computer*, 28(2): 64-68.

Generative Approaches

Real-Time Design Feedback:
Coupling Performance-Knowledge with Design Iteration for Decision-making

Paola Sanguinetti, Marcelo Bernal, Maher El-Khaldi, Matthew Erwin
Georgia Institute of Technology
paolas@gatech.edu, marcelo.bernal@gatech.edu, maher@gatech.edu, matthewerwin1@hotmail.com

Keywords: Generative design, Performance-based analysis, normative calculations, Decision support systems

Abstract

Many commercial environmental analysis tools support the evaluation of a building model based on parameters assigned in the design process. However interoperability issues between different data exchange formats hinder the iteration between design and analysis. Because the engineering calculations involved in analysis and evaluation are not integrated with architectural design parameters, evaluation takes place after the design is already defined, and analysis in real-time is not possible. In this project we integrate architectural design and engineering constraints to support design evolution and decision making by using a set of performance objectives. We propose a framework for coupling performance knowledge with generative synthesis to address multidisciplinary design challenges in the Architecture, Engineering, and Construction (AEC) industry. We develop a tool to support design evaluation based on performance criteria: energy consumption, comfort, and cost. Results show real-time information exchange as links between architectural geometry and engineering parameters. The outcomes of this research describe workflows and methods to evaluate alternative design proposals at early stages in real time.

1. INTRODUCTION

The need for a sustainable environment demands a closer look at the existing building stock and a more integrated approach in Architectural Engineering and Construction (AEC) industry. Many environmental analysis tools have appeared in the market, to support the evaluation of a building model based on parameters assigned and choices made in the design phase. However interoperability issues between different data exchange formats hinder the iterative process between design synthesis and analysis (Eastman et al, 2008, Augenbroe, 2004). Efforts to improve interoperability between design and analysis domains are on-going. Previous research has focused on overcoming hurdles rooted in the building semantics and representation in different commercial applications (Eastman, 1999, Augenbroe et al, 2004).

The design synthesis process includes design problem description, definition, and solution (Simon, 1996). Design synthesis supported by computational means is referred to as a generative synthesis process (GSP) (Alfaris and Merello, 2008). The complexity and iterative nature of GSP renders the integration of building performance parameters dependent on interoperability between CAD and analysis tools. This typically delays the analysis step until after a design solution is fully developed. Thus real-time analysis of design synthesis cannot be fitted into the process. In this project we integrate architectural design parameters with engineering constraints within GSP. Our aim is to provide real-time feedback and support the selection of candidate solutions. The integration of design synthesis and analysis is implemented through coupling simple geometric representations generated by CAD tools with normative calculations in spreadsheets. We explore two paths for GSP: a parametric modeling environment and a procedural programming environment. Our tool supports analysis, evaluation, and selection based on the performance criteria of energy consumption, comfort, and cost. Results show real-time information exchange through links between architectural geometry and engineering parameters and constraints. The results of this research describe workflows and methods that are used to evaluate alternative design proposals in real-time thus facilitating informed design decisions.

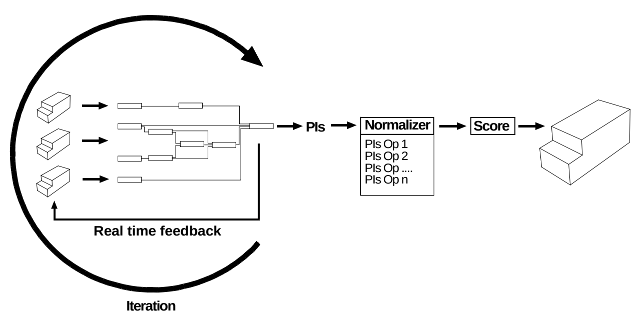


Figure 1: Real-time feedback and selection process

2. CASE STUDY: ENVELOPE DESIGN

In this case study we integrate performance-based analysis and evaluation within GSP. This approach supports the exploration of design options, ensuring compliance with the performance requirements for a particular project. The integration analysis and evaluation within an iterative loop is applied in the selection of an envelope system for a building retrofit.

The building is an existing precast concrete 6-story office in Atlanta, GA, to be converted into the headquarters of a large architecture firm. The retrofit of the building includes the replacement of the envelope system. Glazing units are replaced throughout and external shading devices added. An outdoor terrace in the top levels is to include a large canopy. Two levels of underground parking are not considered in the modeling process.

The proposed method involves the generation of a model of the existing building as a base case and three design options. A design model and its variations as candidate options are implemented in two GSP environments, procedural and parametric. The performance-based design evaluation is implemented in a spreadsheet environment as modules for normative calculations. This tool supports the evaluation of the candidate design options based on the following performance indicators: annual energy consumption, payback to quantify energy cost savings, and lighting comfort of the occupants.

2.1. Implementation Methods

Two CAD environments were explored to provide design options: Digital Project, a parametric modeling environment, and RhinoScript a scripting environment within Rhino. We acknowledge the latter as a procedural process. Figure 1 shows successive stages of this experiment and how the procedural and parametric systems are linked to the performance-based normative calculations which we refer to as calculators. To perform design option evaluations, calculators read parameters from the design option, and provide results as measures of performance.

From the calculators, designers can get two different and complimentary kinds of feedback for decision-making. First, each generation of a single candidate option is evaluated by the calculator as providing real-time feedback in the form of performance indicators (PIs) for each performance criteria. PIs constitute the aggregation and normalization of data to quantify performance (Foliente, 2000). Designers receive real-time feedback from the PI, in a one way process, to adjust and refine the proposed option. Once a set of candidates is generated, the second type of feedback occurs. Through iteration, performance indicators for each of the four performance criteria are used to evaluate, compare, and rank a pool of options to select the best candidate.

2.2. Procedural Environment

A GSP was implemented in RhinoScript, a Visual Basic scripting language. The script generates geometric representation of window panels and external shading devices (horizontal and vertical). These windows appear on all building facades. The base data needed for the script is the vertical and horizontal faces of the building mass. Once selected, the user is prompted with four input widows, once for each façade orientation (North, South, East, and West). Upon completion of geometric generation, the user is finally prompted to select the spread sheet designated to read the procedural system outputs and execute the calculations. The process is one directional, from Rhinoscript to spreadsheets. The time needed to generate final analysis results is under one minute.

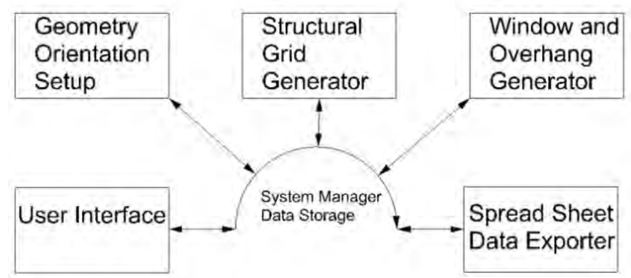


Figure 2: Architecture of the procedural environment

Figure 2 illustrates how the procedural system works. The main controlling component is the System Manager and Data Storage. This component orchestrates the process. It begins by implementing the User Interface component which deals with selection of building's mass faces and receiving user input to define main driving parameters. The user is prompted with series of data input fields to define slab-to-slab heights, column-to-column distances, window offsets from each, shading device depths, and the building's orientation angle. This allows for exploring various window designs (sizes and overhang depths). It also aids the designer in understanding the impact of various orientations on the overall energy performance. The system then orients the entire building, finds the orientation of each façade, then passes on the information back to the System manager. The System Manager triggers the Structural Grid Generator to generate a set of points that will be used in the following function to build geometric representations for the window/overhang systems based on the user-defined input. This grid generator reports back to the System Manager the areas of all windows, their orientation, and count. This information is sent to the Spreadsheet Data Exporter which asks the user select the Calculation Sheet designed to read output from the overall system. The main advantage of such a system is that it allows for fast explorations of alternative forms, and that it provides a visual feedback as well as an automated numerical feedback of their geometric properties.

2.3. Parametric Environment

The parametric model architecture derived in Digital Project is structured by driving, constant, and driven parameters that control the configuration of windows and shading devices for design exploration. The driving parameters control main building dimension and floor to floor height. Constant parameters control the dimensions of the structural components.

Driven parameters control window dimensions, spacing, and distribution. Generic shading devices are parametrically modeled and propagated on all four facades (figure 3). Each façade has its own set of driving parameters so that variation is explored for every orientation. The shading objects include both horizontal and vertical devices and are propagated through every window. The user can control the geometric depth and quantity of shading subdivisions, both horizontal and vertical, for every façade independently.

Parameters and outputs (floor, wall, and window areas for every façade, etc) are exported to a spreadsheet and passed to the normative calculators for performance evaluation. From that point, the parametric model is fully controlled by this external spreadsheet. Hence, for every change in the parameters, actualization occurs in the parametric model as well in the spreadsheet, providing real-time feedback to the user. This process allows iterative refinement, as a two-way process, for the selection of design options.

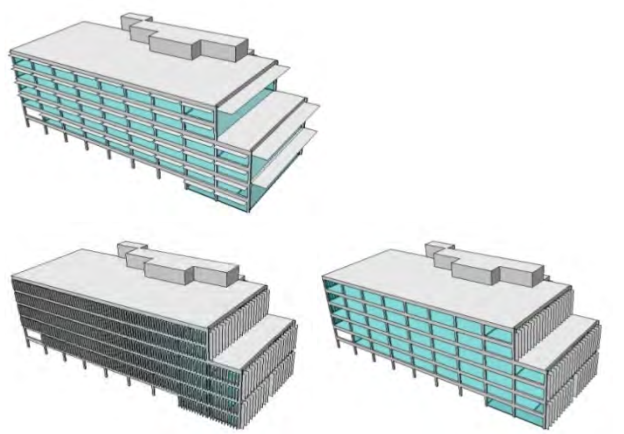


Figure 3: building envelope options generated parametrically

2.4. Normative Calculators

We refer to the performance-based analysis and evaluation process tools as calculators, based on the modular implementation of linked spreadsheets. All calculators use normative equations to obtain PIs. All general equations used in these calculations can be found in the Appendix section.

The performance-based analysis is based on a system of performance calculators that can be generalized into the following categories: Energy Consumption, Investment Payback, and Visual Comfort. The results of these calculators serve as input for another calculator: Weighted Selection. This calculator is used for the evaluation of multiple candidates and optimum option selection.

The calculators are meant to be used for general evaluation of building mass design choices. An orthogonal geometry of four walls and a roof is assumed for energy calculations. For visual comfort analysis, an orthogonal room, of a size partially determined by the user, is used, instead of an entire floor, because overall building proportions are normally out of the range of the IESNA data used to calculate daylight factors. Evaluating these variables at a room level is beneficial for internal layout arrangements. General input variables including latitude, longitude, weather data, and utility/material costs enable customized analysis of localized user requirements. Typical meteorological year (TMY3) weather data files, a representative set of actual weather data collected over a 45-year period, are used as the source of all external weather input data. TMY3 files for 1020 locations are available from the National Renewable Energy Lab at http://rredc.nrel.gov/solar/old_data/nsrdb/1991-2005/tmy3/.

2.4.1. Energy Consumption Calculator

The concept for the Energy Consumption Calculator is based on assumption of thermal equilibrium, meaning that thermal energy lost or gained due to various sources must be equalized by energy output in the form of heat or air conditioning produced from a mechanical system to maintain a desired internal temperature range and a specific 'thermostat' temperature constant. As shown in Figure 4, numerous user inputs are combined with weather data to find three primary sources of heat flow: solar radiation, heat transmitted through walls and windows, and internal gains and losses by people, equipment, lighting and heat exchanged during ventilation air changes. These inputs are all totaled to find the overall energy required of the system. Building location and orientation inputs are used to calculate hourly solar altitudes and azimuths that are combined with geometric inputs from shading device and window measurements to find percentages of direct solar radiation that penetrate into the building. This data is supplemented with the weather information from twelve design days, constructed as monthly averages from the raw TMY3 data file, to find the amount of heat gained through direct, diffuse and reflected solar radiation.

Sky, ground, ambient and internal temperatures are combined with thermal resistance and conductance values of user-specified wall materials into an equation matrix to track temperature change as it travels through the wall section. U-values and solar heat gain coefficients, provided by the manufacturer, are used in a similar manner for the window sections. These values are multiplied by their respective façade and roof areas to find the overall heat energy transmitted.

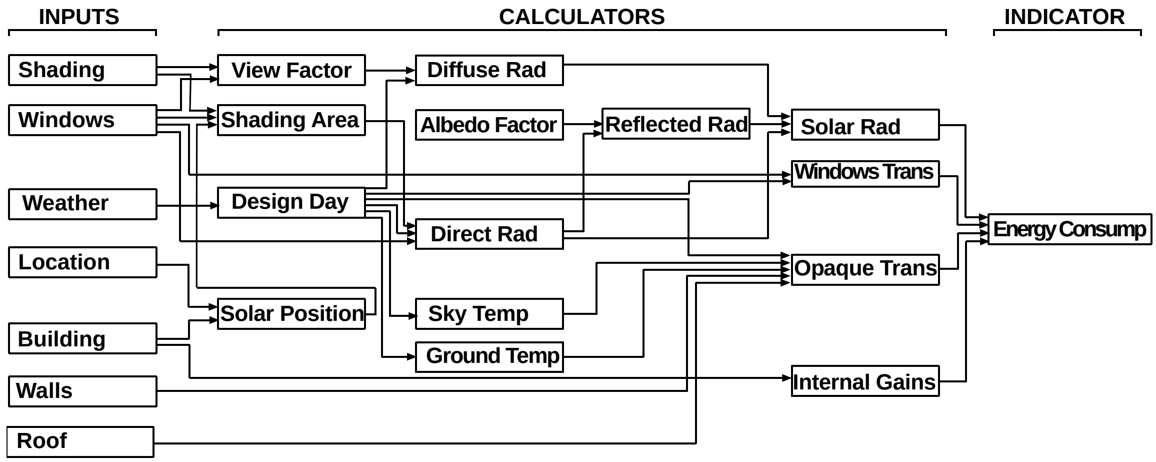


Figure 4: Energy consumption calculation process

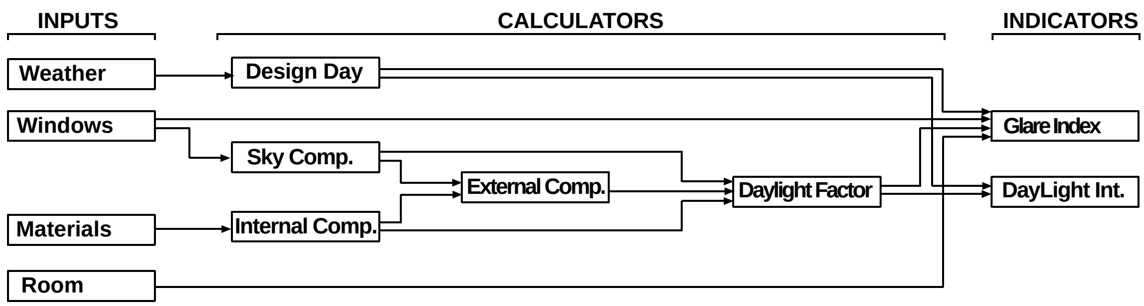


Figure 5: Visual performance calculation process

Internal gains from people, equipment and lighting are calculated based on wattage per unit area values that are derived from type of activity within. These values are obtained from ASHRAE, table 1 F 30. Ventilation exchange is determined from temperature differences and conductance values based on building volume and number of air exchanges per hour selected by the user.

2.4.2. Visual Comfort Calculator

The quantification of visual comfort is calculated using the Daylight Factor Method to verify task illuminance and visual discomfort. This method is implemented with a default value for the CIE design sky illuminance. First, window and room input parameters and default reflectance values for the floor, ceiling, and work surfaces in a standard office space are used to calculate the three components of the Daylight Factor (Figure 5). This factor is used in the calculation of the minimum expected interior daylight values and in the calculation of the Glare Index based on proximity to the window and time of day and year (refer to Appendix for detailed equations). The Glare Index is a defined measure of visual comfort based on the ratio of contrast between interior and exterior light values.

2.4.3. Investment Payback Calculator

The Investment Payback Calculator (Figure 6) is based on user-provided inputs of price per unit area of window, vertical shading system, horizontal shading system and utility costs per kilowatt hour. All values from the Energy Consumption Calculator are output as kilowatt hours which can be directly associated to the unit cost established by a local utility company. These kilowatt hours are considered as cost savings when they contribute toward maintaining the desired temperature range selected by the user and considered as a financial loss when the opposite is true. The costs of the window and shading systems are based geometric areas and the unit price. These system costs are averaged over a range of payback life spans, ranging from 1 to 50 years. For every year, the annual utility cost is added to the annual system cost that corresponds to the payback lifespan, giving a total cost of the system. These costs are only estimates and intended as performance indicators for evaluation in the decision process. Utility cost savings from daylight are not included in this calculator, but should be added in future iterations.

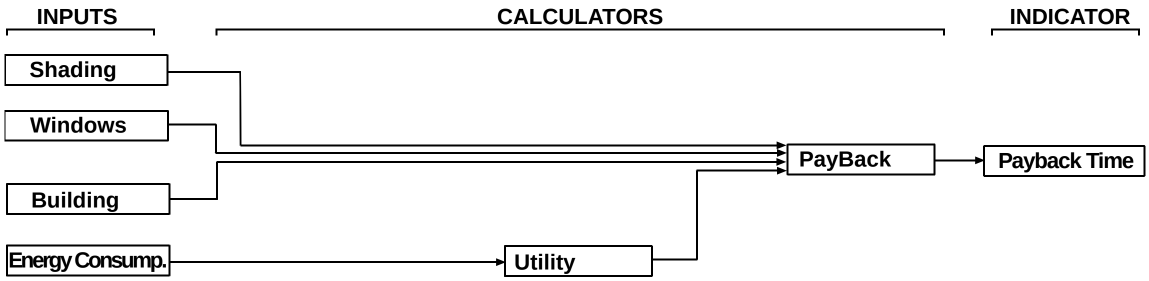


Figure 6: Payback calculation process

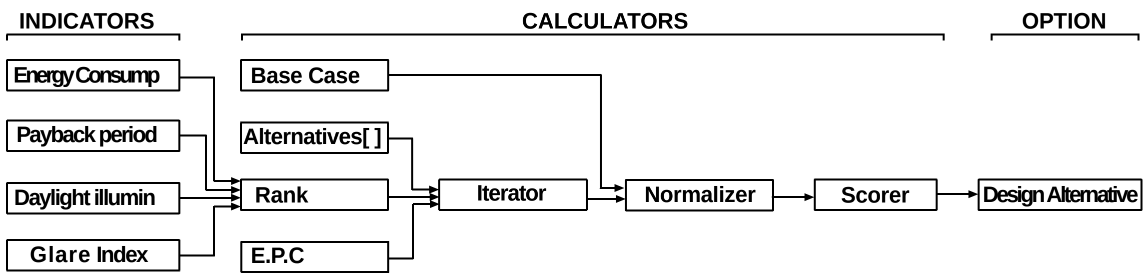


Figure 7: Design alternative selection process

2.4.4. Option Selector Calculator

The selection of a design option (Figure 7) involves normalizing and weighing the results of the four normative calculations according to preference established for the project. The process includes: a) User-established ranking of the performance criteria as an input parameter. This allows for an integrated system where each performance criterion has an embedded project-specific value, which is used in determining a weighing factor for the evaluation of result. b) Scoring of the base case: the calculations results for the existing building are used as a base case and given a default score as an average performance. This default can also be changed by the user, so that the score is project-specific. c) Normalization and scoring of the 3 design options: the calculation results for the three design options are automatically normalized to establish a comparative analysis. d) Selection: the "winning" option is based on the overall scores for the 4 performance criteria. Because this last step requires the collection of the results from the other calculators and the manual input the user's performance rankings, the implementation was done as a spreadsheet separate from the other calculators.

3. RESULTS

Both Digital Project (DP) and Rhinoscript proved effective in defining design parameters, rules, and constraints. However, the nature of generating a design option as a computer model is different. In DP, the geometric elements of the model require methods for retrieving the properties and variables needed by the calculators. This takes place through setting relationships between objects measures or driving parameters. Generating a new design option can be done by changing the variables and updating parametric relationships (figure 7). Rhinoscript allows for setting relationships among variables as part of generating the design option (Figure 8). Because the variables created within Rhinoscript are not stored in the system's memory after the generation is finished, generating new design options requires re-running the script. Both tools communicate differently with the calculator spreadsheets through import/export capabilities. DP maintains a live link to the spreadsheet leading to automatic, real time updates, when input options in the spreadsheet are changed, and regeneration of geometric models. Rhinoscript requires the user to regenerate the models manually, by rewriting and exporting the spreadsheet file.

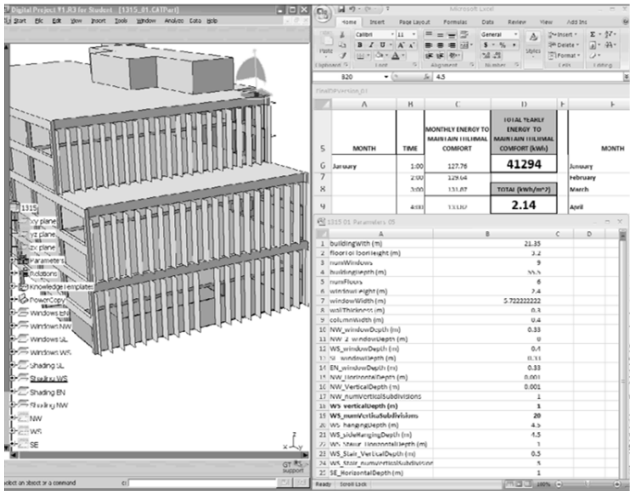


Figure 8: Parametric environment and calculator

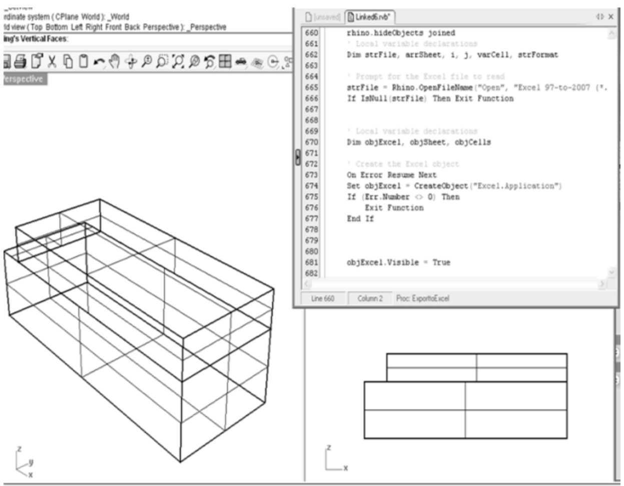


Figure 9: procedural environment

INPUT CRITERIA PREFERENCE		Ranking from 1-5 (1=least	DERIVED WEIGHT FACTORS
CRITERIA 1	Energy to keep thermal comfort	2	18
CRITERIA 2	$ Payback of Retrofit investment	5	45
CRITERIA 3	Task illuminance in the space	3	27
CRITERIA 4	Visual discomfort durign work hours	1	9
		11	99

CALCULATOR RESULTS units	base case	OPTION 1	OPTION 2	OPTION 3
kw.hr	51877.88	44200.22	40423.25	41508.16
$	-192933.25	-259268.84	-338055.09	-263083.85
$	-38485.16	-51838.34	-67831.17	-52785.89
$	-19179.15	-25909.53	-34055.18	-26498.65
$	-9526.15	-12945.12	-17164.18	-13355.03
lux	237.50	237.50	237.50	237.50
lux	93.10	93.10	93.10	93.10
lux	65.50	65.50	65.50	65.50
lux	49.50	49.50	49.50	49.50
hrs	69.00	70.00	85.00	70.00
hrs	0.00	0.00	3.00	0.00
hrs	0.00	0.00	3.00	0.00
hrs	0.00	1.00	3.00	1.00

OVERALL PERFORMANCE SCORES	OPTION 1	OPTION 2	OPTION 3
	15	16	16
	22	8	21
	20	20	20
	8	8	8
	66	52	66

Figure 10: Selector results with scoring of design options

4. DISCUSSION

4.1. Coupling of Design and Analysis Tools

This case study is concerned with performance based evaluations for a building envelope. Parameters and properties for windows and shading devices are varied to produce three candidate options. In this early design stage, three performance criteria are evaluated: energy consumption, payback time period, and visual comfort via glare index and interior daylight intensity. We propose that this methodology can be adapted to incorporate other performance criteria. The following steps are identified for the integration of performance-based analysis with early conceptual design:

Definition of Performance Indicators: Designers should analyze the context with which the design proposal interacts. Designers should decide what parts of the context should drive the solutions, and what parts will be driven. It is also important to design on the level of impact of context onto the design. This will enable the designer to devise the initial set of parameters that will drive the generation of solutions. Designers should also establish the main reasons for analyzing their design (lighting, acoustics, energy consumption, cost, etc).

Development of Normative Calculations: We suggest developing a set of normative calculations based on generic formulas, and heuristics within spreadsheet software. Nowadays, most computer software packages integrate with spreadsheet software directly or via text files.

Generative geometric representations: Designers should develop GSP to arrive at alternative solutions. These alternatives should only provide the needed geometric representation of the design proposal for visual evaluation and extraction of geometric data for analysis purposes. We suggest parametric modeling environments or scripting.

Dynamic Linking: Once the previous three requirements are met, designers should explore methods for linking the driving parameters, the spreadsheet calculator and the GSP.

Development of a selection method: For selection to take place, designers should first normalize the analysis results. Once results are normalized, a selection criteria control must be in place to give users the ability to select weighted value for each of the decision drivers.

Calibration: Finally, designers must calibrate their results by comparing the various analysis feedback outcomes with one another, to hand calculations, or to specialized analysis packages for verification. Once this is achieved, the performance of many alternative solutions can be evaluated.

4.2. Real-Time Feedback, limitations and opportunities

The main advantage of our approach is the synchronization between parametric modeling and/or scripting environment and performance-based calculations, providing feedback on geometric and material variation. We find that the procedural environment re-executes the GSP after every feedback to rebuild the model, whereas the parametric environment allows localized parameter modification. The latter approach enables direct visualization of the effect of the calculation, without intermediate steps of design modifications.

In contrast, the weaknesses of this approach are due to the limitations of the tools to support complex mathematical calculations. In addition, this approach requires an expert user to define the architecture of the design model, including parameters, rules, and constraints. However, the model itself does not need to be highly elaborated, since a basic model is normally sufficient to perform evaluation.

Based on the results of this study, we observe that the exchange of properties and attributes remains a hurdle in linking a design model from a CAD tool to a performance-based evaluation environment. The issue is rooted in the way each applications classifies the model data within their system. Comparatively, our approach resolves this issue by exchanging the design model's geometry and related properties rather than complete data structures. This data exchange is coupled with the specification of attributes within the spreadsheet environment to the support analysis, evaluation, and decision-making.

5. CONCLUSION

Unlike Engineering, Architectural design problems are commonly ill-defined, which presents both opportunities and limitations, depending on methodologies followed in defining the scope of the problem-solving investigation. In such a context, it becomes crucial to set methods to evolve the design solution progressively. Today most CAD tools support customized algorithms to generate variation and explore design intent. In these tools, parameters, constraints, and rules are mainly used for geometric variations, without verification of the outcomes. In this research, we explored the capabilities of the current tools to support performance-based evaluations by linking them to normative calculations and actualize designs in real-time. The results of our research reveal the complexity involved in the implementation process, rooted in the limitations of the current design tools, and the level of expert knowledge required to perform analysis calculations. Our study presents the work flows to support the integration of these two types of domain knowledge though detailed methods for coupling design with low fidelity analysis. A generative synthesis process, implemented in procedural and parametric environments, is used to provide an array of proposed solutions. Results show how to support designers in making informed decisions in the early design stages. This approach offer opportunities and limitations based on the nature of the design, the user's expertise, and the definition of the design objectives and performance requirements. We believe that even though the tools are linked externally with calculations outside the CAD tool, this approach expands the design environment to include quantitative analysis and improve the overall quality of each design option.

References

Alfaris, A. and R. Merello (2008). The Generative Multi-Performance Design System. Silicon + Skin: Biological Processes and Computation. Proceedings of the 28th Annual Conference of the Association for Computer Aided Design in Architecture (ACADIA) Minneapolis.

Augenbroe, A., (2004) Developments in interoperability, in Ali Malawi and God fried Augenbroe (Eds.). Advanced Building Simulation, Spon Press, 2004 pp189-216

Augenbroe, G., A, Malkawi, P. de Wilde, (2004). A Workbench for Structured Design Analysis Dialogues. Journal of Architectural and Planning Research, 21 (4), 321-330

Eastman. C. and Kutay, A, (1991) "Transaction Management in Design Databases" in Sriram, D., R. Logcher and S. Fukada (eds.) Computer-Aided Cooperative Product Development, Springer-Verlag, N.Y, 1991

Eastman. C, (1999) Building Product Models: Chapter 10: Information Exchange Architectures, pp.321-349

Eastman, C.M., Tiecholz, P., Sacks, R., Liston, K., (2008) BIM Handbook , John Wiley & Sons Inc.

Foliente, G.C., (2000). Developments in performance-based building codes and standards, Forest Products Journal, 50(7/8), 12-21

Hagentoft, C.E.(2003) Introduction to Building Physics, Studentlitteratur, Lund

IESNA (2000) Illuminating Engineering Society of North America, Mark Stanley Rea , ed. 9 edition (July 2000)

Simon, H. (1996). The Sciences of the Artificial. The MIT Press.

Szokolay, S.V, (2008) "Introduction to Architectural Science", Architectural Press, Oxfor, UK.

Appendix

Energy Consumption Equations

To find the position of the sun solar altitude and solar azimuth:
$\beta = \sin^{-1}[\ \sin(\text{latitude}) * \sin(\delta) + \cos(\text{latitude}) * \cos(ha) * \cos(\delta)\ ]$
if $ha \leq 0$: $\theta = \cos^{-1}[(-\sin(\text{latitude}) \sin(\beta) + \sin(\delta)) / (\cos(\text{latitude}) * \cos(\beta))] - \pi$
if $ha > 0$: $\theta = \pi - \cos^{-1}[(-\sin(\text{latitude}) \sin(\beta) + \sin(\delta)) / (\cos(\text{latitude}) * \cos(\beta))]$

Where, β is the solar altitude measured in radians; θ is the solar azimuth measured in radians from true south; ha is the hour angle; δ is the sun declination angle.

Profile angles are used to find shading geometry:
$\gamma = (180 / \pi)*[\tan^{-1}(\tan(\beta) / \cos(|\psi - \theta|))]$
Where, ψ is the building surface azimuth angle.
Thermal Energy through solar radiation, heat transmission, internal heat sources, and ventilation exchange are found using the equations below.
$Q_{total} = Q_{radiation} + Q_{transmittance} + Q_{internal} + Q_{ventilation}$

Where,
$Q_{radiation} = (Q_{diffuse} + Q_{direct} + Q_{reflected}) * SHGC$
SHGC is the Solar Heat Gain Coefficient of the window
Q_{direct} is Direct Normal Radiation coming through the unshaded surface of the window
$Q_{reflected}$ is the daily Global Horizontal Radiation
$Q_{diffuse}$ is the Diffuse Horizontal Radiation
$Q_{transmittance} = Q_{transparent} + Q_{opaque}$ (Hagentoft, 2003)

Where,
$Q_{transparent} = U_{value} * Area_{window} * |\text{Dry Bulb Temp External} - \text{Internal Temp}|$ [3] (Szokolay, 2008)

$Q_{opaque} = U_{value} * Area_{wall}(T_1 - T_2)$
T is the temeperature each surface of the wall construction

$Q_{internal} = Q_{people} + Q_{equipment} + Q_{lighting}$ (Szokolay, 2008)

Q_{people} = # people/unit area * Watt lvl/per person/unit area * $A_{rea\ of\ bldg}$
$Q_{equipment}$ = Watt lvl/unit area * $A_{rea\ of\ bldg}$

$Q_{lighting}$ = Watt lvl/unit area * $Area_{bldg}$
The wattage level for people, equipment and lighting are based on (table 1, F30 (ASHRAE 2005)) and found in source: www.ibpsa.org/proceedings/BS09_1857_1864.pdf

$Q_{ventilation} = q_v (T_{ambient} - T_{surface})$ (Szokolay, 2008)

Where, q_v is the ventilation conductance based on the number of air changes and the volume of the space.

Visual Comfort Equations

The equations below were used to calculate Daylight Factor

DF = SC + ERC + IRC (source: IESNA, 2000)

Where; DF is the daylight factor; SC is the sky component with uniform or CIE overcast skies; ERC is the externally reflected component; IRC is the internally reflected component.

Equations for Glare Index calculations are based on sequences and empirical charts found in Steven V Szokolay. Intro. To. Arch. Science.

Glare Index = 10 * Log10 (0.478 * $\sum$ g)

Where g is the glare constant found by
$g = L_1^a / L_2 * \omega^b / P^c$

L_1 is the luminance of the glare source, in this case the ratio of luminance to the area of the window; L_2 is the luminance of the background, based on the illuminance value, the average reflectances of the of the surrounding surfaces, and the ω the angle of view to the work plane.

Luminance = $(1+2\sin\beta) * Lz/3$
Where, β is the solar altitude; Lz is the zenith luminance.

Cost Estimates Equation

System Cost per year = [(# of beneficial kWh per year - # of detrimental kWh per year) * $/kWh]
+ [($\$_{Horizontal\ Shading}$/unit area * $Area_{Horizontal\ Shading}$) + ($\$_{Vertical\ Shading}$/unit area * $Area_{Vertical\ Shading}$)
+ ($\$_{Window}$/unit area * $Area_{Window}$)] / # of years in payback lifespan

Explorations of Agent-based Simulation for Architectural Design

Nick Puckett

The University of Kentucky
Dept. of Architecture
Dept. of Electrical and Computer Engineering

Keywords: Fuzzy Logic, Component system controls, Modeling Interfaces, Drawing-Based Interfaces.

Abstract

With most forms of computational or Information-based architectural design the 2D scale drawing is often treated as an automated output of a complex system to translate information to builders or fabricators in an understandable format. We present a case study for a new typology of scale drawing which serves as the input and design control to rule-based component systems. This type of drawing serves as a construction document between the designer and the digital construction logic setup within the system and draws on the history and robustness of the scale drawing as a means of encoding information to develop a stronger design interface with complex component systems. The cases presented examine the potential of fuzzy logic as a method of control for component systems, and specifically how the logic graphs can become an "active scaled drawing" for designing system rules.

1. INTRODUCTION

The architectural drawing is an entity which serves several purposes throughout the design process to communicate design intentions. In relationship to construction the drawing's role is to encode highly detailed sets of information into a format that can transfer design intentions into formal outcomes. It is an object Robin Evans described as "not done after nature, but prior to construction; it is not so much produced by reflection on the reality outside the drawing, as productive of a reality that will end up outside the drawing" [1]. As tools and methods have matured within parametric or component based design strategies we have come to the point where the digital act of modeling is much more closely related to an act of building in that we must manage the flows of data and the rule sets that govern it to all components simultaneously. This becomes problematic as the numbers of components grow into the thousands or tens of thousands and a designer must find the way to attain both and overview of the system and a local understanding of each rule. Much work has been done recently to simplify the creation of these types of systems with a move away from only written script languages to nodal connections, but once created there has been little advancement of the means of controlling these systems once they are in place. This is a shortfall that many proclaim in regards to parametric design as the results can become incredibly similar. Others see this as a positive force and the re-emergence of a categorical style within architecture [2]. Our argument is that while the software tools for creating these systems have matured, there have been few advancements in looking at the interface and interaction between the designer and these complex systems. At best we are relying on the "cult of the slider" which biases a notion that these systems can produce many results rather than looking to new methods which provide a truly new type of interaction in digital space. We would argue it is this lack of meaningful control tools that have led to a certain "sameness" in component-based systems as de- signers lose the ability to craft the individual components. We present that to push the designer's capabilities within these systems we must look both backward and forward. To do this We will present a series of experiments in component based design which utilize "fuzzy logic" as a means to more effectively design the interaction between components and the data that feeds them. The Theory of Fuzzy Logic was presented in 1965 by Prof. Lofti Zadeh as a means of creating modeling techniques which could more thoroughly describe sets with vague or diverse input, but precise outputs [3].

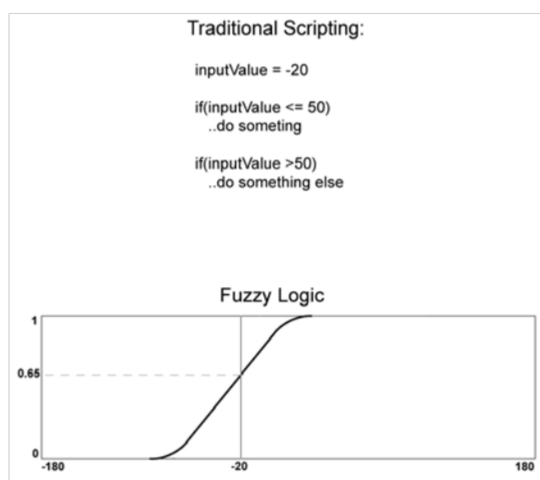


Figure 1: Comparison of fuzzy and boolean logic

The key to fuzzy logic methods are fuzzy sets which describe the "membership" of incoming data in relation to the set in terms of degree rather than a Boolean logic [4]. In this way all incoming data is described as to being a degree of true or false. This methodology has been used for decades as a means of creating control systems in devices such as elevators or traffic lights [5]. For the purposes of this paper the interest in "fuzzy logic" stems from the way of implementing these "fuzzy rules". These rules are typically implemented via "logic graphs"; 2D function curves which graphically describe the rule and determine the resultant degree of truth. In this way potential rules are not interpreted as being either True (1) or False (0) but rather a degree of truth (between 0-1) (see Figure 1).

2. CASE STUDIES – MASSIVE PRIME

The case studies undertaken were developed using Massive Prime. Massive was originally created as an agent-based crowd tool for battle scenes in the Lord of the Rings Trilogy, and later was developed into a standalone package [6]. Within Massive the user imports geometry from a 3D modeling package and develops a "brain" which controls the behavior of that geometry. Brains are created by connecting a series of nodes that typically follow the format of INPUT – FUZZ – AND/ OR – DEFUZZ (see Figure 4). The input nodes function in a very similar fashion to sensors on a physical robot in that they simply "sense" a given type of information within an area or "see" using computer vision techniques. In this way "agents" can have identical brains which result in a greatly varied outcome due to the difference in the information "sensed". The fuzz nodes are where the designer can draw the functions which determine how the agent interprets the information being sensed. Typically there is more than one rule that is associated with the input values, and one of the strengths of using fuzzy logic is calculating the results from overlapping rules (see Figure 2).

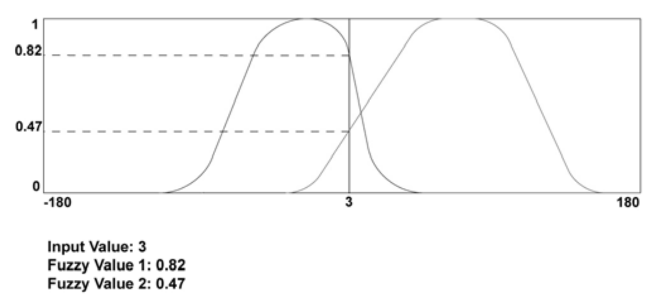


Figure 2: Membership functions and their resultant output

This in turn removes the hard edges between rules and creates a more fluid system. The AND/OR nodes allow multiple rules to be combined and computed together. The defuzz stage is where the output or "membership level" of the rules are converted to a finite value for the agent. This value could for example, control the rotation of an assembly, scale a member, deform a mesh, or choose a different geometry type. In this way the designer can define the minimum and maximum possibility of the output and the precise value is calculated by the rules drawn.

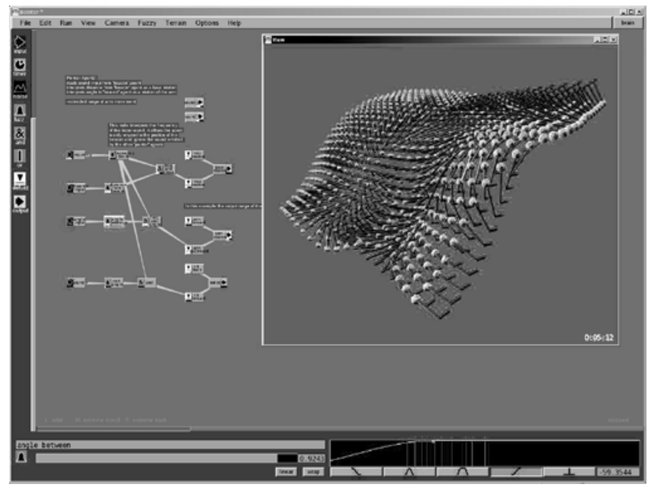


Figure 3: Massive Prime's interface between brain and geometry

2.1. Case Study: Light-based deformation

The initial studies looked to replicate relatively simple component based design techniques to begin to understand the methodologies of creating and controlling these systems in Massive and with fuzzy logic. Thus the initial studies focused on taking a single component, applying it to a guide surface, and deforming the geometry of the component based on it orientation, location, and relationship to the changing sun angle (see Figure 5). More than anything these initial studies brought to light the general working methodology within Massive for these types of systems.

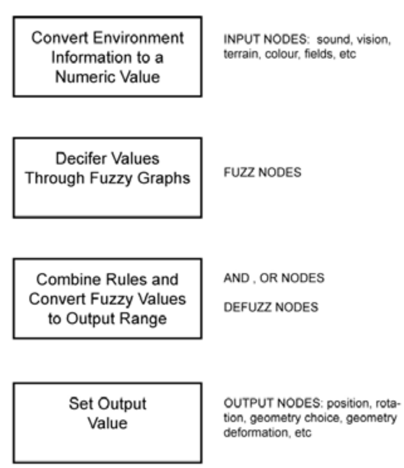


Figure 4: Calculation and Output Logic in Massive Prime

First the component is developed on its own and initial rules are set for the governing behavior. The components are then instanced in their hundreds or thousands, and then the simulation is started. As it is an animation software the behaviors of each agent are simulated as an animation frame by frame. In this case the main animating elements were A) the sun animating along its predefined sunpath and B) the geometries deforming in response. What is extremely powerful about this method is that while the simulation is calculating, the logic curves governing the components can be adjusted so that the system can be "tuned" in real time. This is made more useful by the fact that since all the agents operate off the exactly the same "brain" you can simply adjust the single "master" rule drawings and all of the components react. (Note: individual component brains can also be adjusted individually). Once the system has been tuned the component is saved with its updated rules. In this way the simulation provides a much closer link with the designer and allows for changes to be made during the calculation process rather than waiting for a final result and adjusting values.

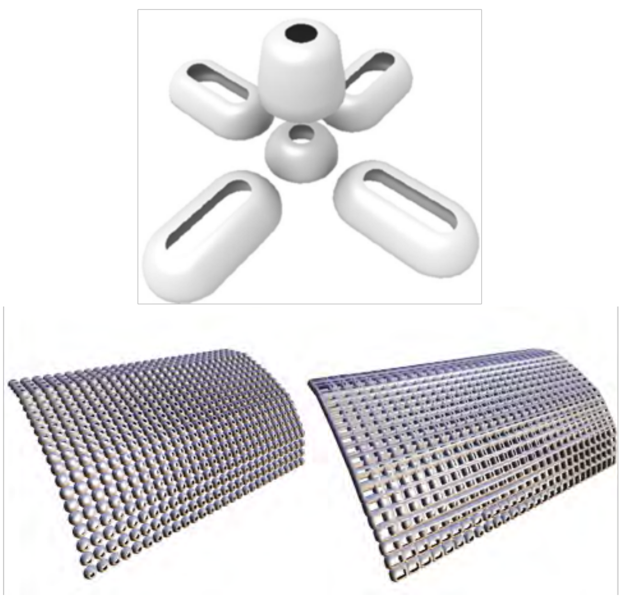


Figure 5: Facade study of component morphing based on orientations and sun angles.

2.2. Case Study: Encoding Multiple Inputs

Building upon the previous study, this case study looks at how 1 type of component can read inputs from 3 different types of component and make decisions based on all of them. This was investigated in the form of a very diagrammatic master-planning tool (see Figure 6). In this simulation one set of agents function as the input to the system and provide the conditions that the masterplan must react to. This is achieved by the agents drawing a different intensity of their respected color based upon their relationship to the other input agents. The colors represent the programmatic inputs and their level of mixing with other inputs. By having the agents encode the information into color it can be simultaneously transferred to another system and visually read by the designer, providing crucial feedback. As the color is being drawn the grid of "buildings" reads the levels of Red, Green, and Blue underneath them and adjusts its height and overall form to correspond to the formal outcome of the programmatic inputs. The way that the building reacts to the input is determined by its own rule-set which are adjusted during the simulation process. The other main difference from the previous example is that the input agents have much more intelligence of their own and it becomes very important where they begin as they interact with each other rather than simply following a path.

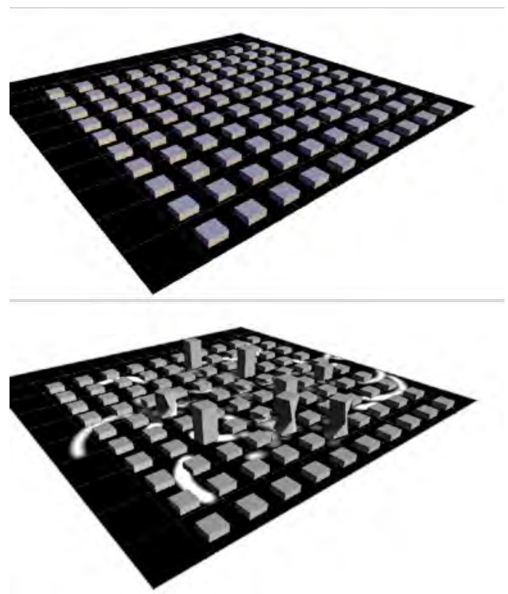


Figure 6: Masterplan Simulation using color based inputs

2.3. Case Study: Arrangement and Deformation

The next studies looked to move beyond typical surface controlled components and develop the rules which would tell the components how to arrange themselves as well as to perform local deformation of their geometries based on other inputs (see Figure 7). In this simulation all of the components start at the exact same position and by examining their local conditions arrange themselves into a net-like structure with each piece connecting with its neighbor. In this case the rule-drawings which govern the process of assembly become even more closely related to constructing documents. Though not a graphic representation of the outcome these rule-drawings hold a scalar relationship to the final form achieved. A secondary layer of rules also determines the height of each unit as well as its degree of bending. As the objects deform the information is sensed by its neighbors and the arrangement is shifted, affecting the overall form. In this case the simulation of the behaviors is used as an animation of the digital construction process and by interacting with the rules during simulation the designer can adjust properties during the process to form a much shorter feedback loop.

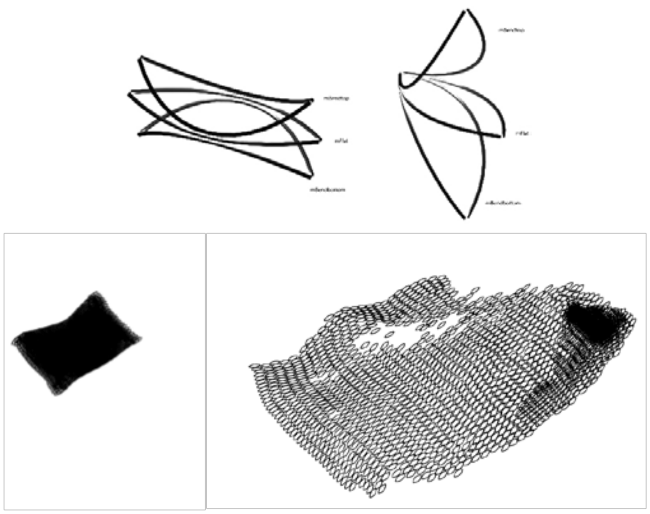


Figure 7: Global configuration and local deformation of the component

2.4. Case Study: Propagation of Intelligence

The final case study examines the ability to spawn new agents at runtime during the simulation. The typology of the "clustered tower" was used as a mechanism to investigate the possibility for a set of "builder" components to actually create the towers floor by floor based on their own rules about height and proximity to the other builders (see Figure 8). Interestingly, the floorplates created are not simply "dumb geometry" as they each have an embedded brain and rule-set, and after they are created by the builders they are continually looking at the local and global conditions to adjust their size, shape, and color. Each floorplate can also build new agents based on the overall stability of the cluster such that if the overall form becomes unstable, new "legs" begin to build from the floorplate back to the ground. This process of digital construction of intelligent geometry multiplies in complexity with each new component created and the simplicity of the rule controls are invaluable to the ability to actually craft the outcome during the simulation.

3. FUTURE WORK

The development of this work will look to push this type of drawn input to systems even further into the realm of digital constructing documents. This will be achieved by introducing other standard practices of construction documents into the logic graphs such as line-weights and hatching. Future development will also look to develop was of visualizing the "scale" of the rule in relationship to the output values it is controlling. The goal is to embed the aspects of the logic system into the graphic rules to more easily understand the relationship between rules within a component.

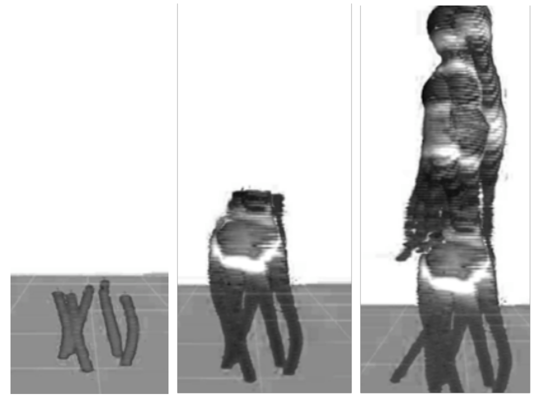


Figure 8: Tower cluster built by propagating smart floorplates from builder agents

4. CONCLUSIONS

These case studies represent the beginning stages of research into a new link between the scale drawing and the rule as a mechanism of digital design. The goal is to reinvent the scale drawing as a method of interface to production that allows designers to gain a higher degree of control over complex, component-based systems. This becomes extremely important as more information is embedded into the digital modeling process, and this method provides a simplified control mechanism as well as a simplified graphic representation of the information. These case studies also highlight new methods of working with simulations as a form of design by interacting with the simulation frame by frame forming a much tighter feedback loop between the components and the simulation data.

5. REFERENCES

[1] Evans, Robin, Translations from Drawing to Building. Cambridge,MA: MIT Press, 1997.

[2] Schumacher, Patrik, "Parametricism: A New Global Style for Architecture and Urban Design" 2009. AD, Vol 79, No 4: 14-22.

[3] Lawry, Jonathan, Modeling and reasoning with vague concepts. New York: Springer, 2006.

[4] Celikyilmaz, Asli. Modeling uncertainty with fuzzy logic. Berlin: Springer, 2009.

[5] Kluska, Jacek. Analtical methods in fuzzy modeling and control. Berlin: Springer, 2009.

[6] http://www.massivesoftware.com

Programming in the Model:
Contextualizing Computer Programming in CAD Models

Maryam M. Maleki, Robert F. Woodbury
School of Interactive Arts and Technology,
Simon Fraser University, Surrey, BC, Canada
(mmaleki,robw)@sfu.ca

Keywords: End-user programming, scripting, visual programming, direct manipulation

Abstract

Programming in the model locates programming elements and tasks contiguous with computer aided design (CAD) models. It aims to reduce the separation between acts of programming, modeling and design, using both spatial coincidence to reduce task shifting and common CAD techniques to simplify the expression of code. Using techniques from visual programming, parametric modeling and CAD selection we demonstrate how programming in the model can express the core steps in a very simple simulation algorithm.

1. INTRODUCTION

Designers use programming in computer aided design (CAD) for several reasons including exploring unconventional designs, reducing repetitive work and enabling change within the design process [Aish, 2003; Woodbury, 2010]. In recent years, such programming has become a normal part of design in innovative firms worldwide. Almost all CAD systems now provide extensive programming interfaces and such organizations as Smart Geometry conduct regular workshops and conferences that focus on programming within the design process. These designers are end-user programmers: they have little formal programming education and program as part of their job, that is to complete tasks [Woodbury, 2010]. Therefore they experience the same issues that other end-user programmers face in domains such as accounting in spreadsheets and analysis in structural engineering. One of these issues is the separation between programming and the task (design, analysis, etc.). In CAD, designers have to switch between the programming environment and the model view several times to make even the most simple modifications and evaluate their effects on the model, which results in a loss of focus and efficiency.

The literature emphasizes the importance of a close fit between the programming world and the domain to reduce novice and end-user programmers' cognitive load [Green and Petre, 1996; Pane and Myers, 1996], as well as to minimize or simplify the use of syntax and code to reduce non-programmers' difficulties in writing code [Kelleher and Pausch, 2005]. Direct manipulation of objects in the interface is a way of reducing the distance between what users think and what the system does and giving users a greater sense of engagement [Shneiderman, 1983; Hutchins et al., 1985]. Visual programming represents the program with visual elements and allows direct manipulation of the program in graphical form (e.g. Rhino's Grasshopper). However, the program is still separate from the model and in order to observe the result of any manipulation of the visual program, users need to go back and forth between the model and the visual programming environment.

2. PROGRAMMING IN THE MODEL

In this ongoing project, we propose *programming in the model* as a technique that uses designers' spatial and visual capabilities and their familiarity with the 2D/3D model and contextualizes programming concepts and subtasks directly in the model. In this approach to programming in CAD, object properties as well as dependencies between objects are represented in the model and can be accessed, modified and assigned by direct interaction in the model view. In addition, programming constructs such as functions and loops are created and modified in the model directly where they are needed and linked to the objects in the model to get their inputs [Maleki and Woodbury, 2010]. In this approach, we use both direct manipulation and visual programming. Contrary to visual programming where the program is represented visually in an independent window in the interface, programming in the model embeds the visual program in the 2D/3D model to reduce the need for designers to move their attention back and forth between multiple windows. Further, debugging is too often at yet another level removed from the design task. Effective programming in the model must include both programming and debugging.

The domain in which we work always presents a collection of objects as its model. Objects in the model are essentially global variables, even if namespaces are used. Programs tend to operate over a subset of the model. This makes practical a strategy of spatial localization of code near the subset so considered. We consider programs that (1) create, (2) change, and/or (3) report object properties (have side-effects). We aim for those constructs comprising a general programming language, i.e., constants, variables, expressions, statements (including control statements) and functions, as well as needed debugging facilities.

The product of programming in the model is a textual program, readable by machine and both readable and editable by people. The reason for this is that we envision programming in the model as only one of the productive modes of programming work. Program visualization and text provide complementary views and insights into code. Over fifty years of work on textual programming should not be ignored. Programming in the model complements, rather than replaces textual programming.

Textual programs are abstractions from the concrete domain over which they compute. In contrast, programming in the model is domain-proximate: it locates and demonstrates code working on specific objects, in this case within the CAD domain. A goal, therefore, is that programming in the model should provide a way for novice CAD programmers to progress from operations over specific objects to general functions that apply to a class.

As designers develop skill and as programs grow in size and complexity, there are tasks that will be more amenable to writing textual programs than to direct manipulation. Also, typical CAD systems provide multiple programming environments, each of greater capability and ending in a full development environment. Making textual programs available and editable at every level in a system can aid the process of laddering to more complex programming tasks and environments.

The current generation of new CAD systems are almost all parametric—they provide for the representation and maintenance of geometric and other data relationships within a model. In such systems programming both takes on a new role and become an inextricable aspect of work. No system can support the myriad of relationships that a designer might envision, so user-extension of these relationships is the norm. The current solution, provided by all systems, is programming. In a parametric system, writing small programs, distributed throughout a model is a normal aspect of the modeling process.

Following is a summary of some features of programming in the model. This is a work in progress; we present only a subset of the eventual features.

- Making lists: A designer makes lists of objects in the model using the list making feature and receives immediate visual feedback on the list in the model. The lists created by this method have the correct syntax based on system's programming language. The designer can modify the lists in the model by clicking on objects to be added to, removed from, or reordered within the list.

- Predefined operations and their inline customization: These are the operations that are applied to lists and individual objects. The system resolves indexing depending on the type of the operation and its operands.

- Implied indexing: The customization of the operations in a regular CAD system requires using the loops to access individual items of the lists. Programming in the model removes the loop by allowing implied indexing of the lists.

- Object properties and relationships: In parametric modeling, objects use other objects as inputs. These creates a network of dependencies among them. Some systems show these relationships independently from the model in a symbolic representation. Programming in the model also presents relationships in the model where the objects are. In addition, object properties are accessible in the model in short or extended modes and are editable.

We explain a sample of these features in the next section.

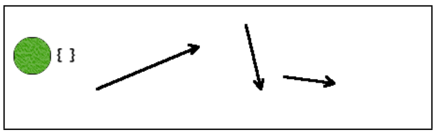
3. SIMULATION WITH PROGRAMMING IN THE MODEL

In this section, we demonstrate programming in the model through a very simple simulation system. Simulation is seldom provided in conventional CAD interfaces, yet designers often need to understand the effect of an environmental force over time or aim at design involving concepts of adaptation, in which parts of a design "respond" to other parts. For such goals, simulation is an almost-necessary tool.

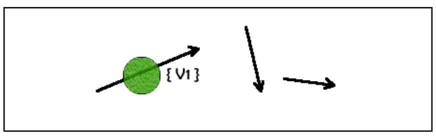
Beginning with the basic operations of vector addition, scaling and point subtraction (which themselves could be further decomposed into addition, subtraction and multiplication over real numbers), we develop all the logic needed for a simulation step. Our intention in providing an example that is both simple and complete is to show that programming in the model is a general concept, rather than a limited convention.

The model in the simulation example is a set of points called *force points* and a single point called the *target point.* The force points and the target point define a set of vectors. The simulation (explained at the end of this section) uses relaxation to move the target points so that the sum of these vectors is zero. We use a convention of variable replication [Aish and Woodbury, 2005] in which a variable of a particular type can carry either a collection or an individual object. While it would be usual in mathematical notation to distinguish between collections (using, for instance, uppercase letters) and objects (diacritics and/or lowercase letters), in the following we use single uppercase letters for all variables (P,Q,R,S), characters followed by a single number to distinguish objects destined for a collection ($P1,Q4,R3,S2$) and use the prime diacritic (P',Q',R',S') to distinguish objects clearly derived from earlier versions.

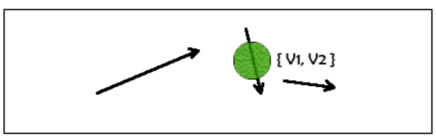
In this example, as in most parametric modeling tasks, we need to make lists of objects to use as inputs for functions or other objects. In programming in the model, we provide tools for users to make and use lists in the model and provide visual feedback in the model for each list. In addition, list objects take care of the specific syntax of the lists for the user (commas, brackets, etc.). Figure 1 shows a list of three vectors that is made by moving the list object in the model and clicking on the vectors to add them to the list. This list object then will be used as input to the operations performed on these vectors.



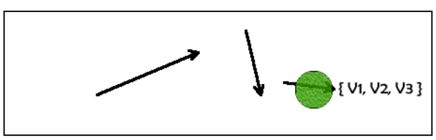
(a)



(b)



(c)



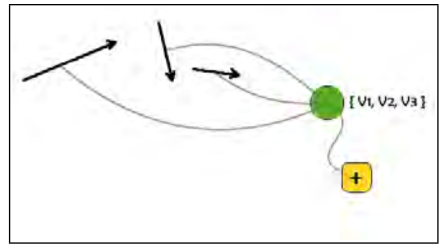
(d)

Figure 1. Making a list of three vectors.

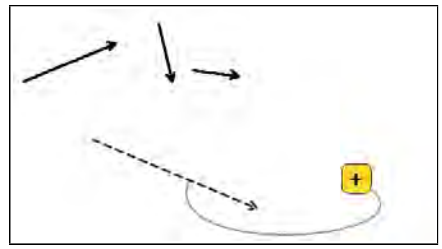
There is a set of predefined operations for each type of object in programming in the model. For vectors the operations are vector addition and negation (which yield vector subtraction), point-vector addition, point subtraction, and scaling. Using *replication* [Aish and Woodbury, 2005], each of these operations applies to lists as well as individual objects. In Figure 2 an addition operation is applied to the list of vectors by bringing the corresponding node to the model and linking the list to it. The result of this operation is a vector that is the sum of the three input vectors.

Let V be a list of vectors, and V' be a new vector, then $V' = +V$ means

```
for (int i = 0; i < V.count; i++)
{
  V' = V' + V[i]
}
```



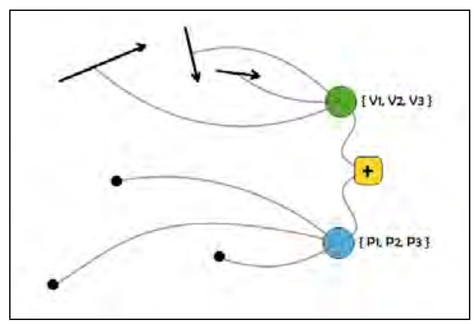
(a) Inputs.



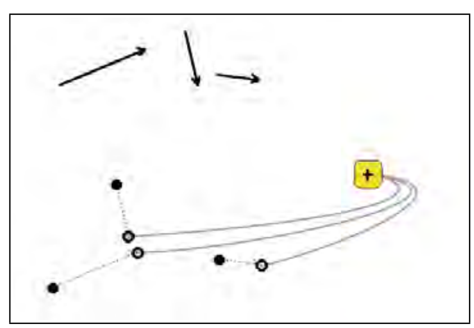
(b) Outputs.

Figure 2. Using addition operation on the vectors in a list.

Figure 3 shows the addition of a list of vectors to a list of points by making a list of points and a list of vectors and linking them to the addition node. The output of his node is a list of points that are the results of adding each vector to each point in the same order they appear in the lists.



(a) Inputs.



(b) Outputs.

Figure 3. Adding a list of vectors to a list of points using the addition operation.

Let V be a list of vectors, P be a list of points, and P' be a new list of points, then $P' = P + V$ means

```
for (int i = 0; i < min (P.count, V.
    count); i++)
{
  P′[i] = P[i] + V[i]
}
```

However, the user may want to add the vectors to the points in a different order. In most CAD systems, using a loop is the only way of doing such thing. Here, we eliminate the loop with *implied indexing* by allowing the user to define the order in a single line added to the addition operation using indexing. By doing that we relieve the user from having to deal with the syntax of a loop, as well as list counts and out of bound indices. The system forms and tests the needed conditional satements. Figure 4 shows a customized addition operation in which each member of the vector list with index i is added to the member of the point list with index $i+1$. The user can customize the operation by opening a text box and typing the desired operation.

Let V be a list of vectors, P be a list of points, and P' be a new list of points, then $P'[i] = P[i+1] + V[i]$ means

```
for (int i = 0; i < V.count; i++)
{
 if( (i+1) < P.count)
  {
   P′[i] = P[i+1] + V[i]
  }
}
```

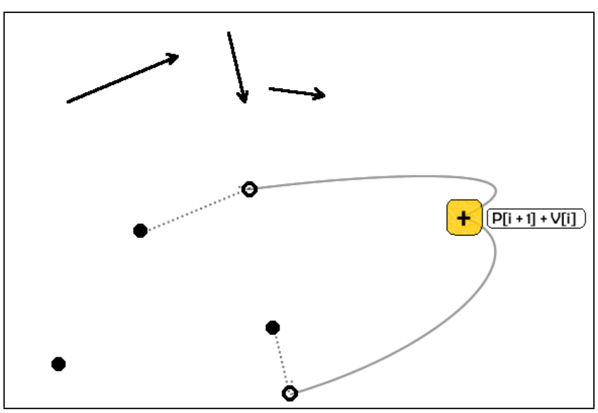


Figure 4. Addition operation is customized so that every $V[i]$ is added to $P[i+1]$.

This method of accessing list items without using a loop can be very helpful in simplifying a collection of loops and conditionals into a single line of code. Here is another example of its use.

Let L and K be lists of real numbers. Then $L[i] = K[i-1] + K[i] + K[i+1]$ is an inline operation in programming in the model that gives the user access to members in the K list by only typing a single line of code. However, in traditional programming languages, the following loop is required for such an operation.

```
for (int i = 1; i < K.count-1; i++)
{
  L[i-1] = K[i-1] + K[i] + K[i+1]
}
```

Notice that the loop starts from $i = 1$ instead of $i = 0$ and ends at $i < K.count - 1$ instead of $i < K.count$. Consequently, the sum of three members of K is assigned to $L[i-1]$ in order to start the list L from the 0^{th} member. All of these irregularities are sources of hard mental operations, confusion, and error for designers [Green and Petre, 1996].

Now we use these primitive operations to create the relaxation algorithm. A *relaxer node* is a composite node built of the above operations that applies the scaled sum of force vectors to the target point. The output of this node is a single point. First we make a list of the force points by using a list object. This list F, the target point P, and a scale factor S are inputs for the relaxer (Figure 5). As displayed in Figure 6, the first operation in the relaxer node is subtraction of the target node from each of the force points to make a list of vectors V. Then we add these vectors together to make a sum vector V'. We scale V' by multiplying it by S. At the end, we add this scaled vector to the target point to get the result point.

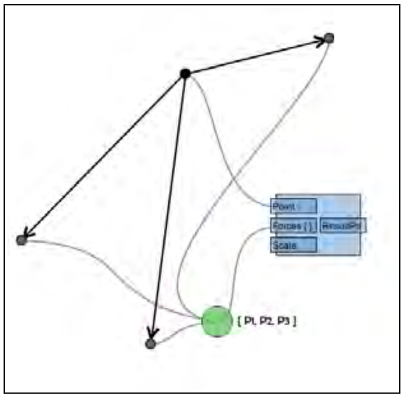
Let F be a list of force points, P be a target point, and V be a new list of vectors, then $V = F - P$ means

```
for (int i = 0; i < F.count; i++)
{
    V[i] = F[i] - P
}
```

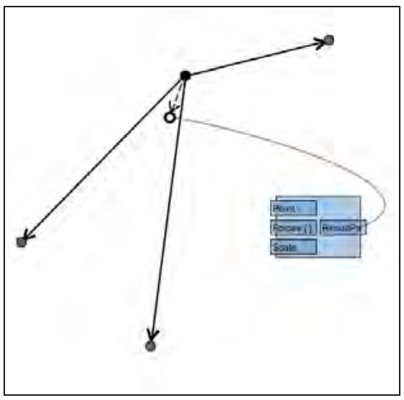
Let S be a scale factor and V' and V'' be new vectors, then $V' = +V$ and $V'' = S * V'$ mean

```
V′ = 0
for (int i = 0; i < V.count; i++)
{
    V′ = V′ + V[i]
}
V″ = S * V′
```

Let P' be a new point, then $P' = P + V''$, which is the addition of a vector to a point, creates the result point. Details of some of these operations are shown in Figure 7.



(a) Inputs.



(b) Output.

Figure 5. Feeding a list of force points and a target point to the relaxer node results in a new point made by addition of the scaled sum of the vectors to the target point.

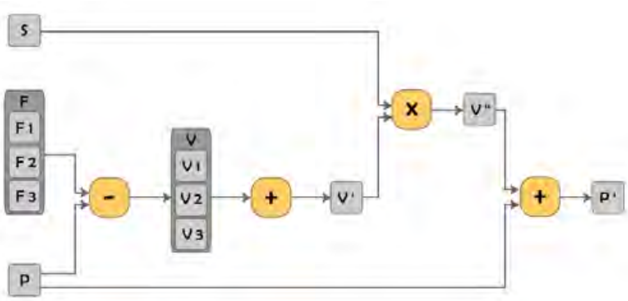


Figure 6. Inside a relaxer node.

The simulator is a node that is independent from the propagation graph and runs only once (Figure 8). The simulator takes the target point and the result point, moves the target point to the location of the result point to get the model closer to the final, balanced state. It then updates the graph. By doing that the relaxer algorithm is run and the result point moves to a new location based on the new state of the model. The internal loop in the simulator repeats these actions until the target point is coincident (within a tolerance) to the result point and the model is "balanced."

$$V = F - P \begin{cases} V1 = F1 - P \begin{cases} V1.X = F1.X - P.X \\ V1.Y = F1.Y - P.Y \\ V1.Z = F1.Z - P.Z \end{cases} \\ V2 = F2 - P \begin{cases} V2.X = F2.X - P.X \\ V2.Y = F2.Y - P.Y \\ V2.Z = F2.Z - P.Z \end{cases} \\ V3 = F3 - P \begin{cases} V3.X = F3.X - P.X \\ V3.Y = F3.Y - P.Y \\ V3.Z = F3.Z - P.Z \end{cases} \end{cases}$$

(a) Subtraction of the target point (P) from the list of force points (F) results in a list of vectors (V).

$$V' = +V \begin{cases} V'.X = V1.X + V2.X + V3.X \\ V'.Y = V1.Y + V2.Y + V3.Y \\ V'.Z = V1.Z + V2.Z + V3.Z \end{cases}$$

(b) Applying an addition operation on a list of vectors (V) results in a sum vector (V').

Figure 7. Details of subtraction and addition operations in the relaxer.

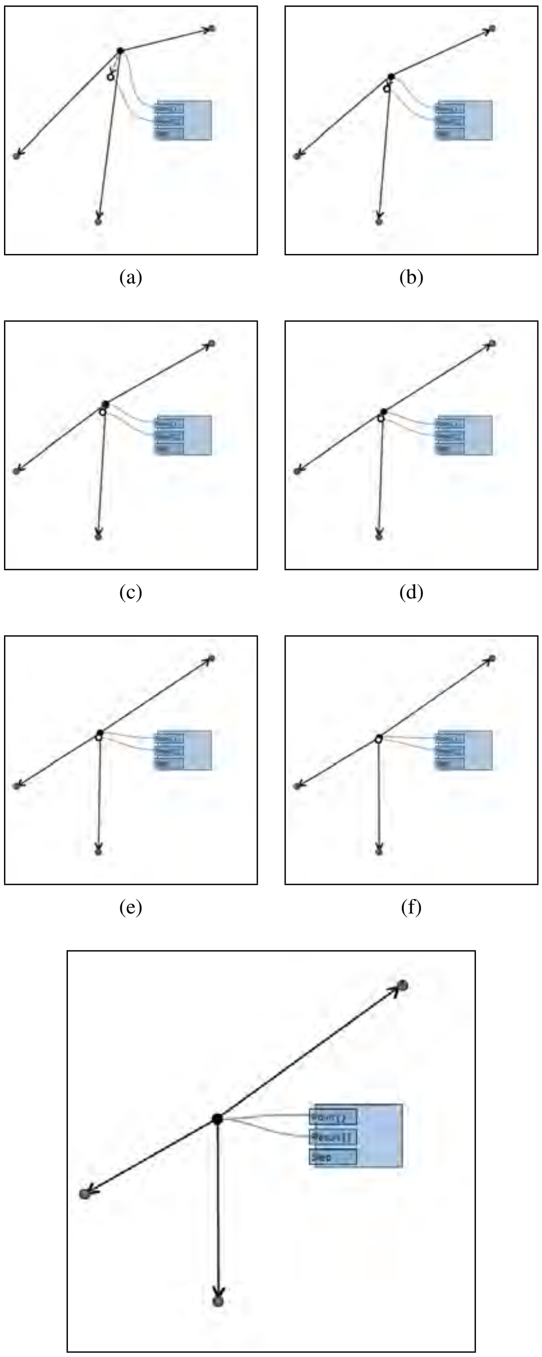
4. DISCUSSION

The simulation example above raises issues and questions that need to be addressed and provides an opportunity for future research.

An important aspect of programming in the model, inherited from the direct manipulation approach, is immediate visual feedback to user actions. This is important for the list making feature where the designer can benefit from having the list visually represented in the model by proper brushing and filtering methods. Modifying a list must immediately change its visual representation in the model. This brings up a challenge when the objects in the list are not simple points. Regular brushing techniques may not work for more complex objects such as surfaces and solids.

Representing programming elements in the model view can cause a cluttering problem, as shown in Figure 3, or may obscure the model all together. Transparency, layering, and on demand filtering can be useful but need to be tested in the context of design tasks.

Although inline indexing can eliminate the use of loops in simple customizations, it does not support every situation. For example, accessing the odd indices of a list requires every index to be checked against a condition. For situations where inline indexing is not enough and the use of a loop is required, how can we simplify the syntax and visualize it in the model for designers?



(a) (b) (c) (d) (e) (f)

(g) Final state of the model.

Figure 8. The simulator node takes the original target point and puts it in the location of the result point made by the relaxer node. This operation is repeated until the model is relaxed and the target point is in a balanced position.

The example of the relaxer node shows that a combination of a number of operations should be generalizable and abstracted into a node, so that with proper input, the desired output is achieved. This will lead to developing a method for representing functions and modules in programming in the model.

And more important of all, what do designers think about this? Does programming in the model help them program during the design task? Does it present too much interference? Is it easy to learn? Does it provide the means for learning the textual language? These questions remind us of the importance of user studies and user feedback sessions along the way.

5. CONCLUSIONS

This simple simulation example demonstrates two goals of programming in the model: simplifying the code, and bringing programming closer to the design.

Operations on some or all members of a list use loops (for, for each, and while loops) to access desired members. In programming in the model, we eliminate these loops by allowing direct access to members of a list through implied indexing as shown in Figure 4. The system takes care of the task of keeping the indices in bound that eliminates the use of conditionals (if statements) to check those indices. In doing so, we simplify the code for designers.

We hypothesize that representing object properties and relationships (visual programming) in the model and providing operations and inline customizations of those operations in the model (direct manipulation) may decrease the back and forth switching between design and programming environments, which in turn may allow the designers to focus their attention on the design task.

At the end, we need to emphasize that programming in the model is not intended to replace textual programming. Rather, the two work side by side and complement one another. Users can choose the form of programming based on their task and their programming experience and preferences.

This is a work in progress. We are currently working on other programming constructs such as functions and loops (when implied indexing is not enough) and how to present them to designers in the context of the model; as well as debugging. We are making working prototypes for future user testing of programming in the model, the results of which will be presented as they become available.

ACKNOWLEDGMENTS

This research is partially supported by the Canadian NSERC Discovery Grants program; Bentley Systems, Inc.; the NSERC Collaborative Research and Development program; and the MITACS Accelerate Canada program. This support is gratefully acknowledged.

REFERENCES

Aish, R. (2003). Extensible computational design tools for exploratory architecture. In *Architecture in the digital age: design and manufacturing*, pages 244–252. Spon Press, New York, NY.

Aish, R. and Woodbury, R. (2005). Multi-level interaction in parametric design. In *Lecture Notes in Computer Science: Smart Graphics*, pages 151–162. Springer Berlin / Heidelberg.

Green, T. R. G. and Petre, M. (1996). Usability analysis of visual programming environments: A 'Cognitive dimensions' framework. *Journal of Visual Languages & Computing*, 7(2):131–174.

Hutchins, E. L., Hollan, J. D., and Norman, D. A. (1985). Direct manipulation interfaces. *Human-Computer Interaction*, 1(4):311.

Kelleher, C. and Pausch, R. (2005). Lowering the barriers to programming: A taxonomy of programming environments and languages for novice programmers. *ACM Comput. Surv.*, 37(2):83–137.

Maleki, M. M. and Woodbury, R. F. (2010). Programming in the model: Combining task and tool in computer-aided design. In *New Frontiers, Proceedings of CAADRIA 2010, Hong Kong*. Forthcoming.

Pane, J. F. and Myers, B. A. (1996). Usability issues in the design of novice programming systems. Technical Report CMU-CS-96-132, Carnegie Mellon University, School of Computer Science, Pittsburgh, PA.

Shneiderman, B. (1983). Direct manipulation: A step beyond programming languages. *Computer*, 16(8):57–69.

Woodbury, R. F. (2010). *Elements of Parametric Design*. Taylor and Francis. Forthcoming.

Life Support

Integrating Building Information Modeling & Cell-DEVS Simulation

Ahmed Sayed Ahmed, Gabriel Wainer, and Samy Mahmoud
Department of Systems and Computer Engineering
Carleton University
1125 Colonel By Drive
Ottawa, Ontario, Canada K1S 5B6
{**asahmed, gwainer, and mahmoud**}**@sce.carleton.ca**

Keywords: Cell-DEVS, CD++, BIM, Revit, API

Abstract

We present the development of an Interactive Environment System (IES). The IES is used for simulating Cell-DEVS models built in CD++ that interact with a Building Information Modeling(BIM) system using Autodesk Revit architecture and Autodesk 3ds Max. CD++ is a modeling and simulation tool that was created to study complex systems by using a discrete event cell-based approach. It was successfully employed to define a variety of models for complex applications using a cell-based approach. The system developed has a highly modular collection of software packages designed to facilitate the creation of device independent simulation for BIM. The integration of the proposed system is investigated via the simulation of Diffusion Limited Aggregation (DLA) which represents the growth of mold in a building wall. The results affirmed the potential of the (IES) system for interactive simulation application.

1. INTRODUCTION

The Discrete Event System Specification (DEVS) formalism [1] has gained popularity to model a variety of problems in the recent years. DEVS is a framework for the construction of discrete-event hierarchical modular models, in which behavioral models (atomic) can be integrated forming a hierarchical structural model (coupled). The Cell-DEVS formalism [2] extended DEVS allowing the simulation of discrete-event cellular models. The formalism extends the traditional Cellular Automata [3] by defining each cell as a DEVS atomic model and the space as a DEVS coupled model. DEVS and Cell-DEVS were implemented in the CD++ toolkit [3].

DEVS and Cell-DEVS were used to solve different problems in Construction and Architecture projects [4]. For successful constructions projects, an enormous amount of data should be collected and analyzed in the pre-design phase of the project. Until recently, the success of this phase was dependant on the experience of experts. However, the information required to complete the project and the amount of data that must be analyzed to complete the project is now greater and more complex. Therefore, there is a necessity for tools that can directly support this pre-design phase. Building information Modeling (BIM) has been considered as a tool that can support this part of the construction project [5]. BIM has resulted in improvements to the way architects-contractors and fabricators have been working [6].

In order to solve different design and simulation problems in the field of Building Information Modeling, it is important to be able to produce prototype solutions easily and quickly. The Interactive Environment System (IES) that we present here provides a reconfigurable system to support the simulation of BIM that can run in a Cell-DEVS environment and then feedback the output result of Cell-DEVS Simulation to the BIM system. We decided use a modeling and simulation toolkit CD++ for Cell-DEVS [3], and Autodesk Revit architecture and Autodesk 3ds Max toolkits for BIM [7, 8]. We show a Cell-DEVS/BIM integration and describe a prototype implementation in the form of a BIM add-in tab for Cell-DEVS simulation, and then visualize the output simulation of Cell-DEVS on the BIM model.

We have surveyed the available systems to see if it supports our proposal, such as Bentley Architecture, Graphisoft ArchiCAD, VectorWorks Architect, Autodesk Revit Architecture, and Autodesk 3ds Max [7–11]. These systems allow us to use Building Information Modeling applications and present a 3D visualization that improves the productivity in building design and construction.

Section 2 of this paper describes the background of Discrete Event Systems Specification (DEVS) ,Cell-DEVS modeling simulation approach, and Building Information Modeling (BIM). Section 3 and Section 4 discuss the simulation of Building Information Modeling (BIM)and the 3D visualization of this simulation on a modeling example. Finally a conclusion of this research is presented in Section 5.

2. BACKGROUND

Discrete Event systems Specification (DEVS) is a system that allows us to define hierarchical modular models [2]. As shown in Figure 1, a real system modeled using DEVS can be represented as a set of atomic or coupled submodels. The atomic DEVS model is defined as:

$$M = < X, S, Y, \delta_{int}, \delta_{ext}, \lambda, ta >$$

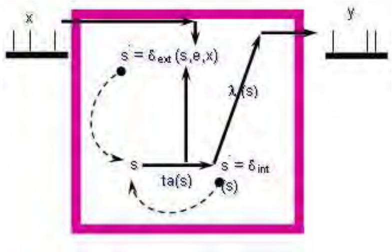


Figure 1. DEVS Atomic Model semantics

where **X** is the input events set, S is the state set, Y is the output events set, δ_{int} is the internal transition function, δ_{ext} is the external transition function, λ is the output function, and ta is the time advance function.

The atomic model can be considered as the base element in which we define dynamics of any system, while the coupled structural model consists of one or more atomic and/or coupled models. Coupled models are defined as a set of basic components (atomic or coupled).The coupled model can be defined as:

$$CM = < X, Y, D, \{M_d \mid d \in D\}, EIC, EOC, IC, select >$$

where X is the input events set, Y is the output events set, D is the set of component names, M_d is a DEVS basic model(i.e., atomic or coupled), EIC is the set of external input couplings, EOC is the set of external output couplings, IC is the set of internal coupling, and select is the tie-breaker function. The coupled model explains how to convert the outputs of a model into inputs for the other models, and how to handle inputs/outputs to and from external models.

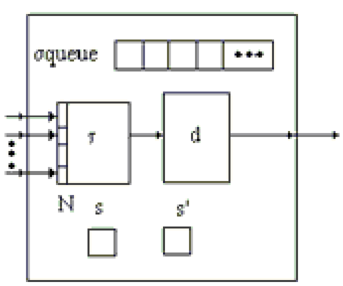


Figure 2. A Cell-DEVS atomic model with transport delay

Cell-DEVS [12] has extended the DEVS formalism, allowing us to implement cellular models with timing delays. Once the behavior of a cell is defined, a coupled Cell-DEVS can be created by interconnecting a number of cells with their neighbors. Each cell is defined as a DEVS atomic model, as shown in Figure 2.

Each cell uses N inputs to compute its next state. These inputs, which are received through the model's interface, activate a local computing function (τ). A delay (d) can be associated with each cell. A coupled Cell-DEVS model is the resulting array of cells (atomic models) with given dimensions, borders, and zones (if applicable). Cell-DEVS were implemented using CD++. CD++ has solved a variety of complex problems [13, 14].

The basic features of the CD++ toolkit can be shown by providing an example of an application. Figure 3 shows the definition of a maze solving algorithm example as an application of such models.

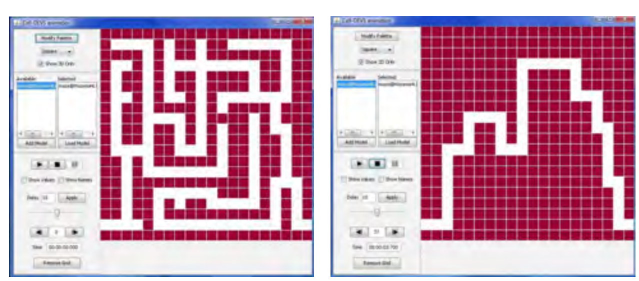
```
[top]
components : maze
[maze]
type : cell
dim : (15, 15)
delay : transport
defaultDelayTime  : 100
border : nowrapped

neighbors :                 maze(-1,0)
neighbors : maze(0,-1)  maze(0,0)  maze(0,1)
neighbors :                 maze(1,0)

initialvalue : 0
initialCellsValue : maze.val
localtransition : maze-rule

[maze-rule]
rule : 1 100 { (0,0) = 0 and (truecount = 3
        or truecount = 4) }
rule : 0 100 { (0,0) = 0 and truecount < 3 }
rule : 1 100 { t }
```

Figure 3. Definition of the maze game

In the example of the maze, the rules are as follows:

- If the cell is a wall cell, the cell remains a wall cell.
- If the number of neighborhoods of a cell is three or more, the cell becomes a wall cell.

When the maze model is executed using these rules, all non-solution paths in the maze are closed successfully. One example of the initial cell state to the maze and the final steady state of the given initial maze cells is shown in Figure 4.



(a) (b)

Figure 4. (a)Initial State of Maze (b)Final State of Maze

A Building Information Modeling (BIM) is the process of generating and managing building data during the life cycle of the project [15]. BIM uses three-dimensional, real-time, dynamic building modeling software to increase productivity in building design and construction. BIM allows us to get better and accurate constructional projects with minimized financial costs [16].

BIM software creates parametric 3D models instead of 2D perspective drawings and operates on a digital database where any change made to this database will be reflected in the whole drawing produced. BIM is often associated with IFCs (Industry Foundation Classes) which are data structures used to represent information used in BIM. IFCs were developed by the International Alliance for Interoperability [17].

BIM is considered as an important improvement to the way architects-contractors and fabricators have been working [6] in that it allows us to minimize conflicts between them, and present a 3D visualization of the building during design and fabrication. Therefore, errors made by design team can be minimized, resulting in costs reduction.

There are different simulation applications for BIM such as IDA Indoor Climate and Energy (IDA ICE) and Design-Builder Software [18, 19]. IDA Indoor Climate and Energy (IDA ICE) is a dynamic simulation application that allows us to calculate the thermal indoor climate of individual zones and the energy consumption of the entire building. Design-Builder Software is a simulation software used to check building energy consumption, CO2 emissions, and more building environmental.

3. SIMULATION OF BUILDING INFORMATION MODELING (BIM)

To use Cell-DEVS simulation for Building Information Modeling (BIM) models, we built an Interactive Environment System (IES). IES is composed of IES_Revit and IES_Max. IES_Revit is a piece of software, used to integrate BIM and Cell-DEVS, whereas IES_Max is used to visualize the output simulation results on BIM models. IES_Revit will be discussed in this section and IES_Max will be discussed in Section 5. To build IES_Revit, there are two main phases:

- Receive the required data to be simulated from the BIM model.
- Simulate the received data of the BIM model by using Cell-DEVS.

For first phase we used Autodesk Revit Architecture [7] as the implementation software for BIM. This part of the Interactive Environment System for Revit (IES_Revit) was written in Visual C# that provides a graphical user interface invoked in Autodesk Revit Architecture (Add-in tab). IES_Revit sends data to be simulated and visualized in Cell-DEVS.

In the second phase we use CD++ as an implementation software for Cell-DEVS. CD++ is a toolkit that models complex systems using a discrete event cell-based approach. CD++ provides 2D and 3D visualization using VRML and Java [20]. 2D and 3D visualization enables visualization of Cell-DEVS models so that the output of our simulation model will be visualized. Figure 5 shows how IES_Revit interfaces with CD++, Revit Architecture, and other software libraries.

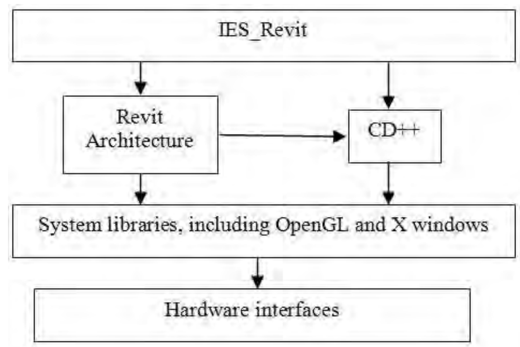


Figure 5. IES_Revit Software hierarchy

3.1. Receiving BIM data

AutoDesk Revit Architecture is a Parametric Building Information Modeling (BIM) tool, in which we can build 3D models and 2D drawings. We can develope different tasks using Revit API. The Revit API allows us to create and delete different model elements like floors, walls, and more. We also use Revit API to get different model parameter data and model model graphical data. Revit Platform API applications can be develop using Visual C# or VB.NET. We decide to use Visual C# for our system.

We create a 'SendToDEVS' class library to be invoked in the add-in tab of AutoDesk Revit Architecture. On the Add-Ins tab, 'SendToDEVS' appears in the External Tools menu-button, as shown in Figure 6. The SendToDEVS is responsible to execute the program to get the required parameters and send it to Cell-DEVS to be simulated.

SendToDEVS includes two functions:

- The GetParameter function, and
- The WriteMacro function

SendToDEVS receives the required parameters of the chosen item (e.g. a wall) in the active Revit document by calling the GetParameter function. The GetParameter function receives the parameters of the selected element in the active document of Autodesk Revit architecture. The GetParameter function sends the parameters to the WriteMacro function. The WriteMacro function is responsible to write the value of the received parameter of the chosen element of the active

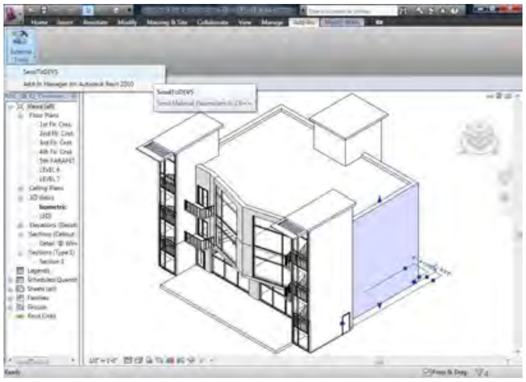


Figure 6. IES_Revit Interface

Revit document into RevitMacro.inc file. The RevitMacro.inc file now contains the new value, which will be simulated in Cell-DEVS. The RevitMacro is explained briefly in Subsection 3.2. If we choose another element in the same active document and click the SendToDEVS add-in tab in the External Tools menu-button, the GetParameter function will receive the new parameters of the new selected element. Then the WriteMacro function will write the new value of the received parameter into the RevitMacro.inc file to be modified.

Finally to invoke the application into AutoDesk Revit Architecture, we should modify the Revit.ini file to register it into Revit by adding the following code to the end of the existing code of Revit.ini file:

```
ECCount=1
ECName1=SendToDEVS
ECClassName1=Integration.CS. SendToDEVS
ECAssembly1=C:\Revit\bin\Debug\SendToDEVS.dll
ECDescription1=Send Material Parameters to CD++
```

3.2. Simulation of the received BIM data

Diffusion Limited Aggregation occurs when diffusing particles stick to and progressively enlarge an initial seed represented by a fixed object. The seed typically grows in an irregular shape resembling frost on a window [21]. We use the Cell-DEVS Diffusion Limited Aggregation (DLA) model as an example of a Cell-DEVS simulation. The DLA model is defined using CD++ toolkit. CD++ includes an interpreter for a specification language that describes Cell-DEVS models. A set of rules is used to define the specification language. Each rule indicates the output value for the cell's state after satisfying the precondition in this rule. These rules are performed sequentially until one rule produces the solution. Figure 7 illustrates the architecture of the IES_Revit. We have modified the Cell-DEVS model of DLA [22] by using a macro definition. We use the macro definition to read the parameters value received from BIM model. We have defined a revit-macro in the main .ma file (model file) of DLA. The model file (.MA) allows us to define the behavior of Cell-DEVS models. This model file reads the initialization data from a .VAL file. The val file has the initial values of each cell of the model in the simulation. The log file is generated when we ran the simulation. The output simulation could be shown in 2D visualization using CD++ modeler tool of CD++ toolkit.

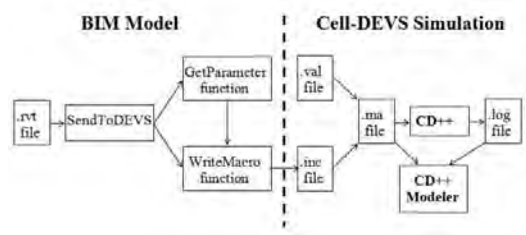


Figure 7. BIM/Cell-DEVS architecture

The DLA model is represented by two types of particles; fixed particles (seeds) and mobile particles. There are one or more seeds in each DLA Cell-DEVS model. A cell that has a seed is fixed and has a value equal to 5. There are a mobile particle percentage of the cells in each DLA Cell-DEVS model. A mobile particle can move according to its value in four directions: up (1), right (2), down (3) and left (4). We set an initial value from 1 to 4 randomly to occupy the cells in a certain concentration. This concentration is calculated and obtained in the Rivet Macro from the BIM model.

```
% initialize the cells with mobile particles
% in the range with value of concentration
rule : {round(uniform(1,4))} 100 {(0,0) = -1
        and random < #macro(Revit)}
```

The following rule presents that fixed particles remains fixed:

```
% fixed particles remains to be fixed
rule : 5 100 { (0,0)=5 }
```

The following rules present the moving of mobile particles:
- A cell has a mobile particle with value equal one can move to the above empty cell if there is no other mobile particle trying to move in to this empty cell.

```
% direction = 1 (up)
% stay and change direction when nowhere to move
rule : {round(uniform(1,4))} 100 {(0,0)=1
    and (-1,0)!=0}
rule : {round(uniform(1,4))} 100 {(0,0)=1
    and (-1,0)=0 and (((-2,0)=3 and(-2,-1)!=5
    and(-3,0) !=5 and(-2,1)!=5) or ((-1,-1)=2
    and(-1,-2)!=5 and(-2,-1)!=5 and(0,-1)!=5)
    or ((-1,1)=4 and(-2,1) !=5 and(-1,2) !=5
    and(0,1)!=5))}
% move otherwise
rule : 0 100 {(0,0)=1 and (-1,0)=0 and t}
```

- A cell has a mobile particle with value equal two can move to the right empty cell if there is no other mobile particle trying to move in to this empty cell.

```
% direction = 2 (right)
% stay and change direction when nowhere to move
rule : {round(uniform(1,4))} 100 {(0,0)=2
     and (0,1)!=0}
rule : {round(uniform(1,4))} 100 {(0,0)=2
     and (0,1)=0 and (((0,2)=4 and (−1,2)!=5
     and (0,3)!=5 and (1,2)!=5) or((−1,1)=3
     and (−1,0)!=5 and(−2,1)!=5 and (−1,2)!=5))}
% move otherwise
rule : 0 100 {(0,0)=2 and (0,1)=0 and t}
```

- A cell has a mobile particle with value equal three can move to the down empty cell if there is no other mobile particle trying to move in to this empty cell.

```
% direction = 3 (down)
% stay and change direction when nowhere to move
rule : {round(uniform(1,4))} 100 {(0,0)=3
     and (1,0)!=0}
rule : {round(uniform(1,4))} 100 {(0,0)=3
     and (1,0)=0 and ((1,1)=4 and (0,1)!=5
     and (1,2) !=5 and (2,1) !=5 )}
% move otherwise
rule : 0 100 { (0,0)=3 and (1,0)=0 and t}
```

- A cell has a mobile particle with value equal four can move to the left empty cell if there is no other mobile particle trying to move in to this empty cell.

```
% direction = 4 (left)
% stay and change direction when nowhere to move
rule : {round(uniform(1,4))} 100 {(0,0)=4
     and (0,−1)!=0}
% move otherwise
rule : 0 100 { (0,0)=4 and (0,−1)=0 and t}
```

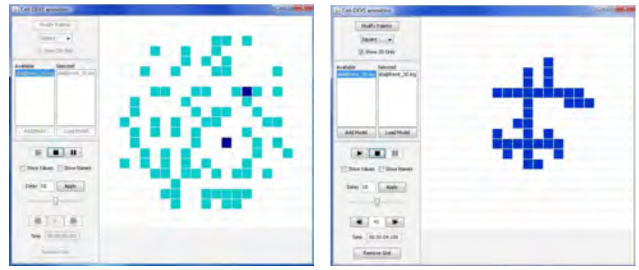
- A cell has a mobile particle becomes fixed if there is an adjacent fixed particle cell.

```
% the particle becomes fixed if an adjacent cell
% contains fixed particle
rule : 5 100 { (0,0) >0 and (0,0)<5 and
     ((−1,0)=5 or (0,−1)=5 or (0,1)=5 or (1,0)=5)}
```

We assume a DLA Cell-DEVS model with two initial seeds. The concentration percentage of mobile particles will vary due to the material parameter type value received from BIM. We ran the simulation for two different materials for the specified two seeds DLA Cell-DEVS model; one for concrete and the other for brick. We assume that the percentage of concentration will be 30% for the concrete material and 40% for the brick material. The simulation output of each run for the concrete and the brick is shown in Figure 8 and Figure 9 respectively.

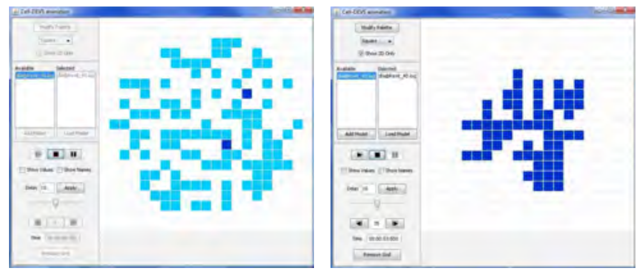
4. 3D VISUALIZATION DESCRIPTION

IES_Max is used to visualize the output simulation results of BIM models. IES_Max is implemented by using Autodesk 3ds Max [8]. Autodesk 3ds Max is used because it supports BIM and has a great 3D environment scene. Figure 10 illustrates how IES_Max interfaces with CD++, 3ds Max, and other software libraries.



(a) (b)

Figure 8. (a)Initial state for 30% concentration (b)Final state for 30% concentration



(a) (b)

Figure 9. (a)Initial state for 40% concentration (b)Final state for 40% concentration

4.1. IES_Max Configuration

We build a graphical display output using 3D visualization tools. We decided to expand our visual environment using Autodesk 3ds Max. Autodesk 3ds Max is a powerful application for 3D modeling and animation, using special effects and rendering. 3ds Max allows users to create 3D animation and visual effects. More functions can be added to Autodesk 3ds Max by using MAXScript. MAXScript is a built-in script language that facilitate the creation of functions and tools to enhance 3ds Max efficiently. We used 3ds Max modeling and animation toolkit to create 3D visual environments for Cell-DEVS Simulation of Diffusion Limited Aggregation as example. IES_Max is an application written in MAXScript that provides a graphical user interface that allows CD++ files (*.ma and *.log files) to interact with 3ds Max, and to visualize the corresponding Cell-DEVS simulation in a 3D visual environment scene of the BIM model. This BIM model, which is exported as FBX file(type of Autodesk file formats) from Autodesk Revit Architecture, is imported in the 3ds Max. Then IES_Max animates the 3D visual scene file in accordance with the CD++ files.

Figure 11 illustrates the architecture of the IES_Max. We use the ReadMa function to read the model file (.MA) of the Cell-DEVS model, which contains the definition of behavior of Cell-DEVS models and the initial data or the path of the

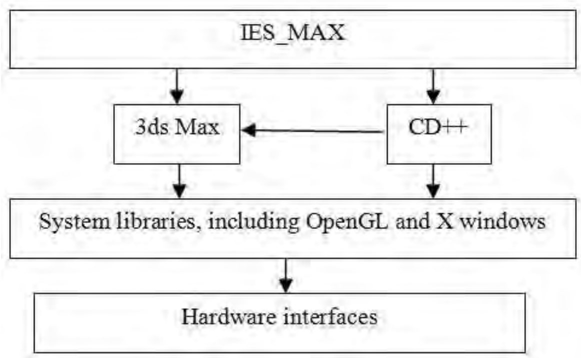


Figure 10. Software hierarchy

val file (.val). The val file has the initial values of each cell of the model in the simulation. So the ReadMa function has the initial status of the Cell-DEVS. The ReadLog function is used to read the log file (.log) of the Cell-DEVS model, which contains all the steps of output simulation results with time. So the ReadMa function has all the status during the simulation lifetime and the final simulation results. Finally, the Draw function draws the collected simulation data in the 3d Max visualization scene.

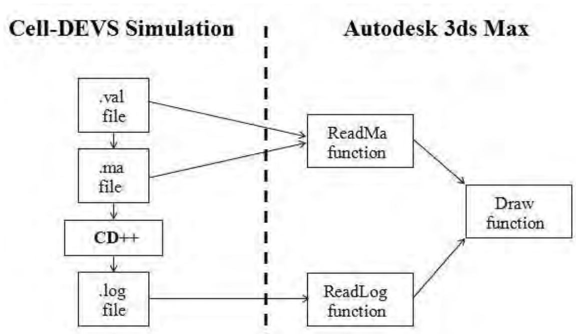


Figure 11. IES_Max architecture

IES_Max consists of Graphical User Interface (GUI). This GUI, as shown in Figure 12, is the graphical interface that requests the user to select a particular file. The GUI consists of 4 buttons. The"Select Ma File" button is used to open the *.ma file corresponding to the file name provided for reading the dimension of the Cell-DEVS and the val file which contains the initial values for the simulation, and "Select Log File" button is used to open the *.log file corresponding to the file name which is provided for reading the output of the Cell-DEVS simulation. The"ExecuteLog" button is used to execute and display the 3D visualization scene in which 3D models are created to be animated due to the data loaded from the ma and val files. Finally, the"Clear" button is used to remove all objects in the Display Window to run another simulation model if needed.

Figure 12. IES_Max 3ds Max Interface

IES_Max allows us to create a 3D visualization from the CD++ files Created by CD++ toolkit. 3ds Max has implicit support for hardware accelerated rendering. The 3ds Max visualization tool provides basic services that enable simple visualizations, including design and implementation of a graphical user interface based on the MAXScript within 3ds Max toolkit.

4.2. IES_Max Implementation

We use the MAXScript language which is the 3ds Max Toolkit script for writing the program to initialize the GUI interface window for the 3D visualization scene. The GUI consists of four buttons:

- The"Select Ma File" button is designed for selecting and opening the *.ma file.
- The"Select Log File" button is designed for selecting and opening the *.log file.
- The "ExecuteLog" button is designed for executing the 3D visualization in the Display Window.
- The "Clear" button is designed for clearing the 3D visualization in the Display Window from any object if you need to run another simulation model.

We use the ReadMa function and the ReadLog function for reading output simulation data of CD++ file, and Draw function for displaying the 3D visualization. The ReadMa function is used to read the *.ma file which contains the dimensions of the simulation model and the val file name, then reformat it to be used in the required argument that passed to the ReadVal function. The ReadVal function is used to read the *.val file which contains the initial values of each object from the simulation model and reformat it to be used in the required argument that passed to the Draw function. The ReadLog function is used to read the *.log file which contains time and position of each object from the simulation model and reformat it to be used in the required argument that passed to the Draw function. The Draw function is used to create objects and display the 3D visualization of the CD++ simulation model in the Display window of 3ds Max. The Draw function receives the dimension of the simulation model which controls the size of the drawing area and the position of each cell to be drawn in the specified location.

5. CONCLUSIONS

This paper describes our implementation of the Interactive Environment System (IES) as an integration of Cell-DEVS formalism into Building Information Modeling (BIM). Cell-DEVS approach can be applied to improve the development of Building Information Modeling. CD++ is used as a toolkit for Cell-DEVS models. We use Autodesk Revit Architecture and 3ds Max as toolkits for BIM. We implemented IES_Revit as an integration between BIM into Cell-DEVS simulation. The output simulation result of the Cell-DEVS model has been visualized in 3ds Max using IES_Max.

The Diffusion Limited Aggregation (DLA) model example was used to verify the feasibility of combining these two technologies by using IES.

REFERENCES

[1] Bernard P. Zeigler, Tag Gon Kim, and Herbert Praehofer. *Theory of Modeling and Simulation: Integrating Discrete Event and Continuous Complex Dynamic Systems*. Academic Press, 2000.

[2] Gabriel A. Wainer and Norbert Giambiasi. *Timed Cell-DEVS: modelling and simulation of cell spaces*. Springer-Verlag, 2001.

[3] Gabriel A. Wainer. CD++: a toolkit to define discrete-event models. *Software, Practice and Experience*, 32(3):1261–1306, November 2002.

[4] Hong Pang, Cheng Zhang, and Amin Hammad. Sensitivity analysis of construction simulation using cell-devs and microcyclone. In *WSC '06: Proceedings of the 38th conference on Winter simulation*, pages 2021–2028, 2006.

[5] Nam-Hyuk Ham, Kyung-Min Min, Ju-Hyung Kim, Yoon-Sun Lee, and Jae-Jun Kim. A study on application of bim (building information modeling) to pre-design in construction project. In *Third International Conference on Convergence and Hybrid Information Technology, 2008. ICCIT '08.*, volume 1, pages 42–49, November 2008.

[6] Brad Hardin. *BIM and Construction Management: Proven Tools, Methods, and Workflows*. Wiley, 2009.

[7] AutoDesk. Autodesk Revit Architecture, November 2009. `http://en.autodesk.ca/adsk/servlet/index?id=12436708&siteID=9719649`.

[8] AutoDesk. Autodesk 3ds Max, November 2009. `http://en.autodesk.ca/adsk/servlet/index?id=12341323&siteID=9719649`.

[9] Bentley. Bentley Architecture, November 2009. `http://www.bentley.com/en-US/Products/Bentley+Architecture/`.

[10] Graphisoft. Archicad 13, November 2009. `http://www.graphisoft.com/products/archicad/`.

[11] VectorWorks. Vectorworks Architect, November 2009. `http://www.nemetschek.net/architect/index.php`.

[12] Gabriel A. Wainer. *Discrete-Event Modeling and Simulation: a Practitioner approach*. Taylor and Francis, 2009.

[13] Javier Ameghino, Alejandro Troccoli, and Gabriel A. Wainer. Modeling and simulation of complex physical systems using cell-devs, 2001.

[14] Gabriel A. Wainer. Modeling and simulation of complex systems with cell-devs, 2004.

[15] Ghang Lee, Rafael Sacks, and Charles M. Eastman. Specifying parametric building object behavior (bob) for a building information modeling system. *Automation in Construction*, 15(6):758–776, 2006.

[16] S. Mihindu and Y. Arayici. Digital construction through bim systems will drive the re-engineering of construction business practices. In *International Conference Visualisation*, pages 29–34, July 2008.

[17] IAI/IFC. International Alliance for Interoperability, Industrial Foundation Classes, November 2009. `http://www.iai-tech.org/`.

[18] EQUA. Equa simulation, simulation software, simulation consultancy, November 2009. `http://www.equa.se/`.

[19] DesignBuilder. Building design, simulation and visualization, November 2009. http://www.designbuilder.co.uk/.

[20] Gabriel A. Wainer and Wenhong Chen. A framework for remote execution and visualization of Cell-DEVS models. *Simulation: Transactions of the Society for Modeling and Simulation International*, 79(11):626–647, 2003.

[21] Tommaso Toffoli and Norman Margolus. *Cellular Automata Machines: A New Environment for Modeling*. MIT Press, 1987.

[22] Wei Ding, Changshuang Lin, Laurentiu Chechiu, X. Wu, and Gabriel A. Wainer. Definition of cell-devs models for complex diffusion systems, 2005.

A Method for Simulating NO_x Dispersion in an Urban Area Using ENVI-met

Francisco Bemquerer Costa Rasia
chico.ix@terra.com.br
Eduardo Leite Krüger
ekruger@utfpr.edu.br

Keywords: urban air quality, ENVI-met, pollutant dispersion, traffic emissions

Abstract

This paper explores a method for the simulation of oxydes of nytrogen (NO_x) in an urban environment. A section of downtown Curitiba, Brazil has been modeled, using data gathered from city plans, satellite photographs and visual surveys; weather data at street level have been gathered by means of field monitoring and used for calibration of the model, focusing on the wind speeds within the urban canyon. Roadside noise level measurements and vehicle counting have been used to determine hourly traffic patterns. NO_X emissions from vehicle exhaust were modeled using ENVI-met's dispersion modelling functions. The model construction and calibration workflow is discussed in detail, and the results of three different simulation scenarios are analyzed and discussed in relation to Brazilian air quality standards.

1. INTRODUCTION

In urban areas, wind conditions affect the ventilation of buildings, the human outdoor thermal comfort and the dispersion of airborne pollutants. This paper is part of an ongoing research project that addresses the impacts of urbanization on outdoor comfort, including air quality and airflow in urban canyons, by means of microscale climate simulation.

The objectives of this research were to evaluate the suitability of ENVI-met and to develop a workflow for simulation of pollutant dispersal in urban areas, using it to predict the effects of alternative land use and building regulations scenarios. It is a readily accessible free software tool which provides visually effective 3D microscale climate analysis.

ENVI-met was used to simulate the micro and toposcale climatic phenomena that take place in two streets in downtown Curitiba, Brazil. Data gathered from city plans, satellite photographs and visual surveys were used to build the model; weather data at street level was gathered by means of field monitoring and used for calibration of the model. The effects of construction typology, green areas and presence of trees were considered as well. Great attention was given to the development of the model workflow and model calibration, which focused on the average wind speeds (measured versus simulated) within the urban canyon. These aspects comprise sections 2 through 4 of this paper.

Sections 5 and 6 of this paper discuss the dispersion of oxydes of nitrogen (NO_x) from vehicle exhaust, simulated with ENVI-met's dispersion modelling functions. The results of three different scenarios (no wind, northerly and easterly winds) are analyzed and discussed, in relation to Brazilian air quality standards.

2. BACKGROUND

The influence of the built environment on atmospheric processes has been a matter of research for several decades. T. R. Oke's seminal work – "Boundary Layer Climates" [Oke 1978] – was a reference in this field of study for the last three decades, with essential concepts such as Urban Boundary and Urban Canopy Layers. It is understood that the built environment "involves the transformation of the radiative, thermal, moisture and aerodynamic characteristics, and thereby dislocates the natural solar and hydrologic cascades" [Oke 1978] within urban regions.

Subsequent work by the same author details the intricacies of vertical layers and horizontal scales, including the notion of a Roughness Sub Layer (RSL) closest to the ground, within which airflow is most disturbed by natural and human-built features [Oke 2006].

In addition to the physical (both natural and built) features of a city, Monteiro [2003], with his notion of an Urban Climate System, proposes the inclusion of social and economic phenomena, which take place in invisible and relative spaces: "even though the social relations have no concrete meaning in themselves, the behaviour of city dwellers and subsequent socio-economic expressions, materialized into the urban structure acquire significance" [Monteiro 2003].

2.1. About the ENVI-met model

ENVI-met is a microclimate simulation tool designed to evaluate plant-surface-air interactions at 0.5 to 10m spatial scale and 10s temporal scale. It is under continuous development by Bruse et al.

According to Bruse [2009], the model is capable of handling aspects of the flow around and between buildings, turbulence and particle and substance dispersion of interest for this study. Model workflow was found to be simple and streamlined, as will be discussed later on this paper.

As examples of applications, Ali-Toudert [2005] has thoroughly reviewed the mathematical model behind ENVI-met in his simulations of urban canyons in hot and dry climates; Spangenberg et al. [2008] have analysed the influence of vegetation in thermal comfort in the city of São Paulo, by means of weather data measurements and simulations. Temperature data gathered at street level (1m above ground) was used to verify the accuracy of the model.

Despite the growing adoption of the model and its userfriendliness, there is still a lack of studies that compare field data to simulation results.

2.2. Air quality within the urban canopy layer

Atmospheric pollution may occur due to natural (such as forest fires, volcanic eruptions) or anthoprogenic processes. Whatever their origin,

> Air pollutants are substances which, when present in the atmosphere under certain conditions, may become injurious to human, animal, plant or microbial life, or to property, or which may interfere with the use and enjoyment of life or property. [Oke 1978]

In the urban environment, wind flow (and therefore the conditions for vertical and horizontal diffusion and mixing) is disturbed by the presence of buildings, vegetation and other features. The cumulative effects from the convergence of several pollutant sources and the existence of stagnation areas may lead to pollution hotspots which ellude mesoscale pollution measurements, and the vertical dimension of pollutant dispersion is often overlooked [Wang et al. 2008].

Substances such as nitrogen monoxide (NO), sulphur dioxide (SO_2), carbon monoxide (CO), which originate from combustion, are considered primary pollutants, while ozone (O_3) and nitrogen dioxide (NO_2) are formed in the atmosphere from chemical reactions of precursor substances under the effects of solar radiation[*]. Particulate matter (PM) may be the result of the burning of diesel fuels, vehicle tire and brake wear, as well as natural processes.

[*]For a comprehensive explanation of these processes, see Oke 1978.

Sulphur dioxide is associated with respiratory illnesses and acid rain. Carbon monoxide is a result of the incomplete combustion of fuels and is extremely toxic, causing deaths by asphyxiation. In the lower atmosphere, ozone is generated in a photochemical reaction involving nitrogen oxides, carbon monoxide and hydrocarbons; it is an irritant and powerful oxidizing agent, also associated with respiratory conditions. Particulates under 10µm in diameter can be inhaled and have been associated with asthma, cardiovascular problems, lung cancer and premature deaths.

Nitrogen monoxide (NO) is involved in various pathological and physiological processes in mammals, both beneficial and detrimental, and may become toxic in high concentrations; in the atmosphere, it rapidly combines with available O_3, generating NO_2 (which is toxic by inhallation and reduces visibility) and O_2. Determining the NO_2/NO_x ratio is beyond the capabilities of the chosen model. Roadside measurements in the Baden-Württemberg region of Germany showed the ratio to be up to 25% [Kessler et al. 2006].

National air quality regulations[**] set two standards for the concentration of pollutants. The primary standard is the upper legal limit for the presence of contaminants; it ensures the safeguard of public health but it does not take into account the needs of fauna and flora. The secondary standard represents the currently agreed threshold below which minimal ill effects on public health, plant and animal life, property and the environment occur [IAP 2009].

Table 1 – National standards for air quality, adapted from [IAP 2009]

Substance	*Sampling time*	*Primary standard ($\mu g/m^3$)*	*Secondary standard ($\mu g/m^3$)*
Particulate Material (PM10)	24 h	150	150
SO_2	24 h	365	100
CO	24 h	40.000	40.000
O_3	1 h	160	160
NO_2	1 h	320	190

3. AREA OF STUDY

Curitiba is a city of 1.8 million inhabitants in Southern Brazil, located at 25º25’ S, 49º16’ W and 934m altitude, in a region of temperate oceanic climate

[**]Air quality in Brazil is regulated by IBAMA no. 348/1990, CONAMA 03/90. State law SEMA 054/2006 replicates the national standards which, however, are not as stringent as WHO regulations for the same substances.

(Cfb under Köppen-Geiger classification), with typically dry Winter season [IPPUC 2009]. Official weather measurements are taken by INMET (Brazilian official weather service) at a meteorological station in Centro Politecnico, Eastern Curitiba (Figure 1). Average maximum temperature is 26.0ºC, and average minimum is 7.4ºC. The prevailing winds blow from East, Northeast and Southeast, with average wind speed between 2 and 3 m/s [Lamberts et al. 1998].

This study focuses on the XV de Novembro Street, a pedestrian street in downtown Curitiba. Its historical importance is undisputed, being the first pedestrian street established in Brazil, in the 1970s. Marechal Deodoro Street, one block to the south, is a four-lane street with very intense traffic during business hours (Figure 1).

Buildings along the XV de Novembro span several decades: terraced houses built in the late 19th and early 20th Centuries, 5- and 6-storey art-deco buildings fron the 1940s and 1950s, and high-rise buildings from the 1960s and 1970s modern movement. This highly varied organization defies simple H/W ratio as well as sky view factor analysis.

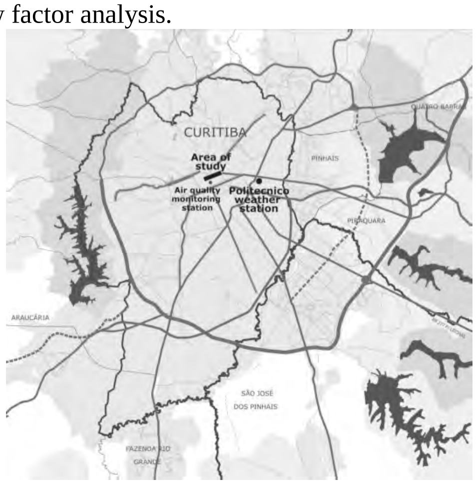


Figure 1 - Location of measurement stations and area of study

3.1. Air quality measurements in Curitiba

Danni-Oliveira (2003) has studied the influence of geo-ecological aspects and urban attributes of the city of Curitiba on the dispersion of pollutants in the Winter by means of sample collection and analysis, between 1996 and 1998. Samples were collected in areas representing different urban structures: the high-density CBD (central business district), residential district and and rural district and, though restricted to a single season, clearly showed the influence of urbanization and land use on the concentrations of the analysed pollutants.

IAP (state environmental agency) currently monitors air quality in the metropolitan area of Curitiba on thirteen measuring stations. Reports for 2008 showed good air quality throughout most of the city; however, a few instances of high concentrations of NO_2 and O_3 were recorded [IAP 2009] in the downtown monitoring station.

4. METHOD

4.1. Overview of the model workflow

The model was built within Eddi, a dedicated model editor included in the ENVI-met suite. A section of 1440mx600m of downtown Curitiba was modeled from city plans, aerial photographs and visual surveys of the area of interest.

City plans spanning 70 blocks were obtained from the Municipality; each street was walked through and the number of floors on each building was counted visually. Approximations were necessary due to different aged buildings, and an average 3m floor-to-floor height was assumed. The maximum number ofvertical grid units allowed in the model is 30, so buildings with more than 25 storeys were limited to 25 in order to allow sufficient space from the top of the model. Ground vegetation and pavements were also included. For many applications of ENVI-met, the geometrical (building, groundcover, vegetation) information suffices.

Simulation of pollutant dispersion, however, requires additional data: vehicle traffic volume (vehicle/hour), hourly distribution of traffic, fleet composition, individual rates of emission (per vehicle) and the number of roadway lanes.

Weather data is input in a configuration (`.CF`) file – temperature, wind speed and direction, and several other parameters. Input wind speed in the configuration file can be adjusted for calibration of the model. After the calibration phase, alternative simulation scenarios were run; the output data was finally plotted for analysis.

4.2. Model domain and resolution

Different model domains (from 240x100 to 40x40 grid units) and resolutions (6x6x3m and 3x3x3m) were evaluated in order to determine the optimal settings (adequate balance between model complexity and processing time). The model versions were run using the same input data and the results were compared. Version 3.1 beta 4 of ENVI-met was used under MS Windows Vista. A model domain with 230x60 grid units at 6x6x3m was ultimately chosen; higher resolutions and wider domains did not improve the accuracy of the model despite the added processing time and model complexity.

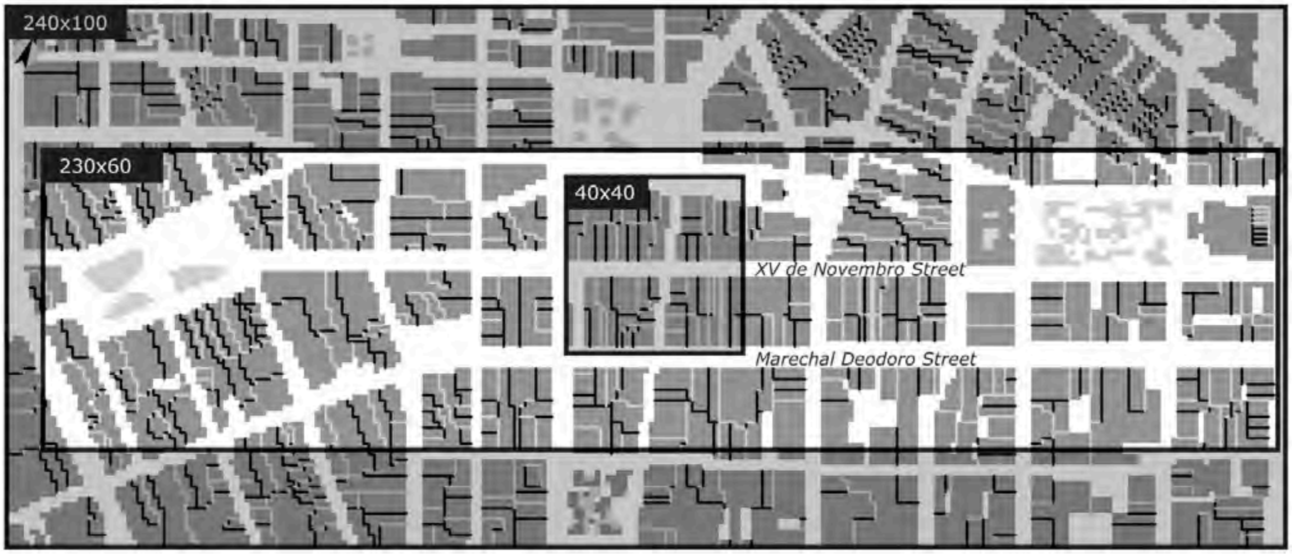


Figure 2 - Comparison between model domains

4.3. Field measurements

Field measurements were used for verifying the model accuracy and for the development of a procedure of adjustment of input data.

Figure 3 - Field measurements at XV de Novembro Street

Thirteen measurement points were choosen along the XV de Novembro Street and nearby squares. Measurements were taken at two points each day, spanning up to five hours (typically from 10h00 to 15h00 local time). Two HOBO Onset weather stations were used, equipped with a three cup anemometer (at approximately 2,1m height), air temperature and relative humidity sensors, two Copper gray-colored globe thermometers and a silicon pyranometer. Data from all sensors were recorded every five seconds, and sampled for each hour. Measurements were taken between January and August 2009, over a wide range of air temperatures, wind conditions and solar angles and height.

4.4. Model calibration

Most attempts at verification of the accuracy of ENVI-met have focused on air temperatures [Spangenberg et al. 2008; Hedquist et al. 2009]; however, since this research focuses on the dispersion of pollutants at street level, mean wind speed at 2.1m above ground has been chosen as the variable of interest. Due to the turbulent nature of wind flow within the Roughness Sublayer, wind direction at said height was ignored.

Virtual receptors were added to the model at locations corresponding to the measurement points. Predicted mean wind speed data at 2.1m height on each receptor at the end of every simulated hour was then compared to the average measured speeds. Calibration of the model was done through several iterations, with changes being made to model geometry, presence of vegetation and adjustment of input data.

Since one is attempting to link the data from a standard weather station in rural-like conditions to data measured at street level in a urban context, some assumptions were necessary; namely, wind direction in downtown Curitiba was assumed to be the same as in the Politecnico station, ignoring any effects of mesoscale urban structures.

Wind speed at 10m above ground (standard WMO measuring height) is used as input into ENVI-met. Early results – from inputting the unadjusted wind data from Politecnico station – showed very poor convergence of measured and simulated mean wind speeds. Equation 1 was used in order to adjust the logarithmic wind profile for urban conditions [Allard and Santamouris 1998, p. 90]:

$$\frac{U_1}{U_0} = K \cdot Z_1^{\ a} \quad (1)$$

Where U_0 is the wind speed in the meteorological station, U_1 is the wind speed at the point of interest, Z_1 is the measurent height (10m), and K and a are adjustment coefficients with values of 0.21 and 0.33 for urban environments, respectively. Hourly mean wind speeds measured at the Politecnico station where adjusted using this equation and input into the ENVI-met simulation; it was found that ENVI-met tends to superestimate wind speeds within the canyon for input wind speeds over 2 m/s (Figure 4a). For input wind speeds under 2 m/s, values predicted with ENVI-met strongly converge with the field measurements (Figure 4b).

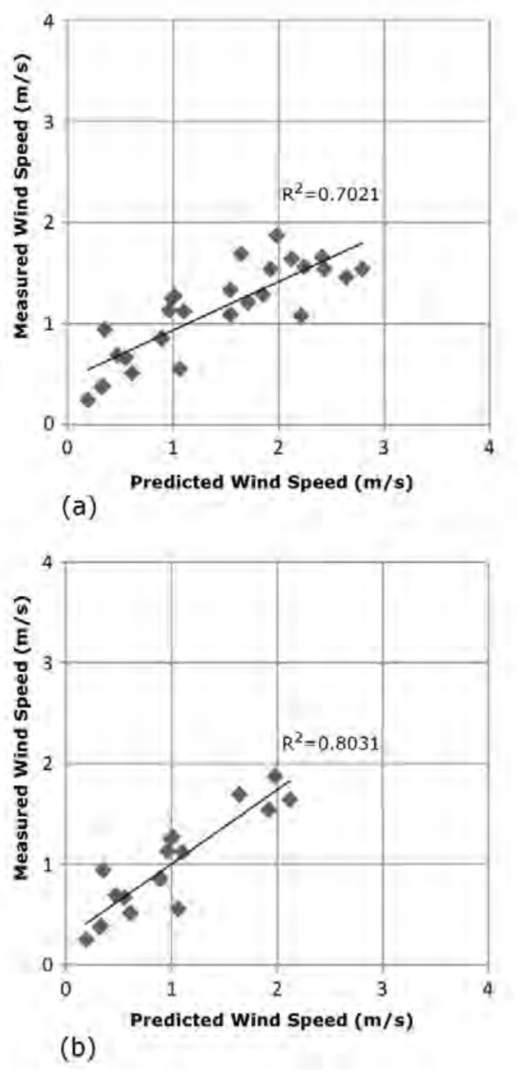


Figure 4 – Comparison between measured data and values predicted with ENVI-met (a) Input wind speed under and above 2 m/s (b) Input wind speed under 2m/s

4.5. Pollutant dispersion modeling

In 2008, instances of high concentration of NO_2 were recorded for downtown Curitiba [IAP 2009], an area of intense heavy vehicle (mass transport busses) traffic. Therefore, the substance chosen for this study was NO_x. ENVI-met requires source information (rate of emission, type of substance, size of particle) to be typed into the `SOURCES.DAT` file. Point, linear and area sources can be modeled, and rates of emission must be adjusted accordingly.

Detailed road traffic data for downtown Curitiba were not available; an indirect method for determining the daily traffic profiles was adopted, from historical noise level data [Bortoli 2002], traffic noise prediction models [Calixto 2002] and short period vehicle counting. Traffic of motorcycles, light and heavy vehicles was visually counted in the rush hours (between 17h and 19h), in three locations, for periods of 15min and extrapolated over one hour, according to the local road management authority recommendations. Average traffic per lane, in vehicles/hour, were 58.3 (motorcycles), 10.6 (light diesel-powered vehicles), 443.3 (light vehicles, not diesel-powered) and 14.2 (heavy diesel-powered vehicles), for a grand total of 526.4 vehicles/hour per lane[*]. The percentage of heavy vehicles was found to be 2.70%.

According to the local transit authority, the average age of vehicles in Curitiba is 10.49 years [Bertotti 2009]; the upper limits of emissions for May 2003, from national standards [Brasil 1993 and 2002], were adopted for motorcycles and light vehicles. The rate of emission of NO_x for heavy vehicles was obtained from Rabl [2002] (Table 2). The composite rate of emission (R) is defined by Equation 2:

$$R = \sum \frac{E \cdot V}{3600} \quad (2)$$

Where R (rate of emission in mg/m.s) is the weighted sum of values of E (emission per vehicle in mg/m) and V (in vehicles/hour).

Table 2 - Emissions of NO_X for motorcycles, light and heavy vehicles

	E (mg/m)		
Substance	*Motorcycles*	*Light vehicles*	*Heavy vehicles*
NO_x	0.18*	0.25 / 1.40**	21.11†

*up to 151cc engine (Brasil, 2002)
**Diesel powered light vehicles (Brasil, 1993)
†Rabl (2002)

[*] The main streets in the model domain have either 3 or 4 lanes; effective traffic on these streetways were 1579 (for 3 lanes) and 2105 (4 lanes) vehicles/hour. Since ENVI-met accepts a single value for the rate of emission, adding pollutant sources on a lane basis allows greater flexibility in model construction.

Daily traffic patterns were extrapolated from Bortoli's [2002] research on street noise in Curitiba, which showed maximum traffic in the downtown area between 17h00 and 19h00. Equivalent noise level values were input into Calixto's [2002] model for prediction of traffic noise (Equation 3):

$$L_{eq} = 7.7 \cdot \log[I \cdot (1 + 0.095 \cdot VP)] + 43 \quad (3)$$

Where L_{eq} is the equivalent noise level, I is the intensity of traffic (in vehicles/h) and VP the percentage of heavy vehicles. From this equation, an hourly pattern of traffic can be estimated, as well as the hourly emission adjustment factor. The composite rate of emission (0.1211 mg/m.s) was multiplied by the emission adjustment factor (Table 3) and the values for R_{NOX} were input into the `SOURCES.DAT` file.

Table 3 - Hourly traffic profile for downtown Curitiba and hourly rates of emission for NO_x

Time of day	*Sound Pressure dB(A)*	*Percentage of mean traffic*	*Emission adjustment factor*	*R_{NOx} (mg/ms)*
15h00	65.0	74%	0.74	0.0898
16h00	65.0	74%	0.74	0.0898
17h00	65.0	100%	1.00	0.1211
18h00	66.0	100%	1.00	0.1211
19h00	66.0	100%	1.00	0.1211
20h00	64.5	64%	0.64	0.0773
21h00	63.0	41%	0.41	0.0494

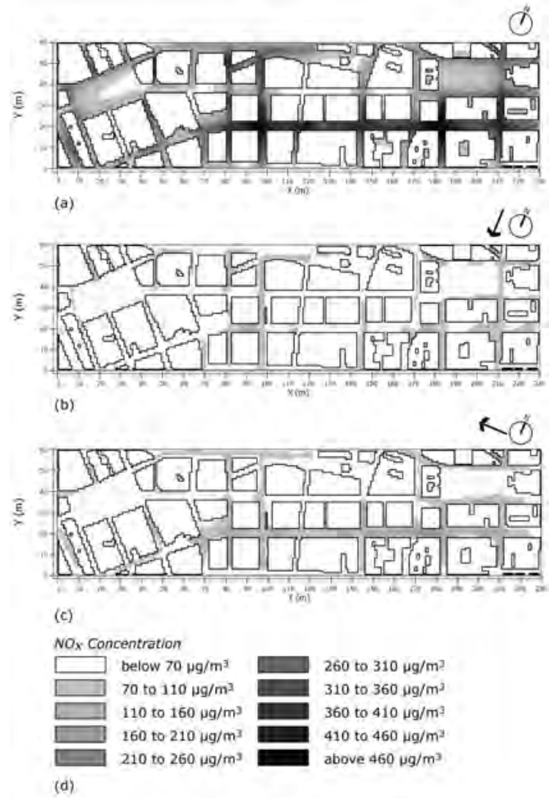


Figure 5 - NO_x concentration, 2.1m above gound: (a) Scenario 1; (b) Scenario 2; (c) Scenario 3 (d) Key

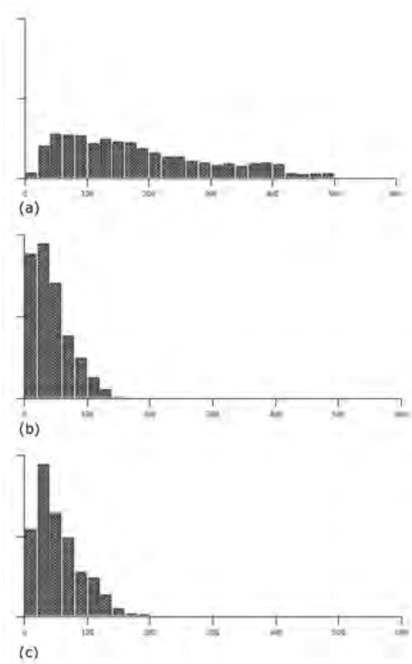


Figure 6 – NO_X Concentration histogram ($\mu g/m^3$), 2.1m above ground; (a) Scenario 1; (b) Scenario 2; (c) Scenario 3

5. SIMULATION SCENARIOS AND RESULTS

On this section, results from the simulation of NO_x dispersion are presented and analysed. Three simulation scenarios were run: Scenario 1 corresponded to stagnation conditions (no wind); Scenarios 2 and 3 corresponded to Northerly and Easterly winds, respectively. Wind speed for scenarios 2 and 3 was 1.36 m/s, adjusted from local mean wind speed of 3.1 m/s.

Each scenario was run for 6 hours, from 15h00 to 21h00, to allow the model to "spin up". Simulations showed that NO_x concentration peaked at 19h00. Concentration data at 2.1m above ground at 19h00 were plotted in Figure 5, and concentration histograms (for the entire model domain) can be found in Figure 6.

5.1. Analysis and discussion

Simulation results showed significant differences between stagnation and windy conditions (Figures 5 and 6a). NO_x dispersion was generally better under northerly rather than easterly wind conditions, as can be seen on the histograms (Figures 6b and 6c).

Within the XV de Novembro street canyon, simulated NO_x concentration with northerly wind was under 30μg/m^3 throughout much of the street. Under easterly wind conditions, NO_x concentration was slightly higher, in the 40-60 μg/m^3 range throughout most of the street. NO_x concentrations for Marechal Deodoro Street were considerably higher. In the first Scenario, NO_x concentration peaked at 499 μg/m^3, and was found to be above 350 μg/m^3 throughout most of the canyon; once again, northerly winds favoured the dispersion of pollutants, with NO_x concentration below 120 μg/m^3 for Scenario 2 through most of the street (except for some hotspots) and within the 60-210 μg/m^3 range virtually along all of the canyon in Scenario 3.

As discussed, ENVI-met is not capable of determining the NO_2/NO_x ratio, but the high concentration of precursor substances offers a hint to NO_2 hotspots. The work of Kessler et al. [2006] has shown evidence of increasing NO_2/NO_x ratio, from 5% (in 1995) to 25% in urban areas. For the first scenario, if the higher ratio of 25% is assumed, NO_2 concentrations could peak at 125 μg/m^3, still within secondary air quality standards. Determining the actual NO_2/NO_x ratio in the area of this study would allow better interpretation of these results.

6. LIMITATIONS OF THE MODEL

For the purposes of this study, four limitations of ENVI-met became apparent:

- Regarding wind speed, the model accuracy is reduced if the input wind speed is greater than 2 m/s;
- It is not possible to input background substance concentrations; simulated concentrations represent only those originating from the sources within the model domain;
- Only a single substance can be simulated at a time, demanding several successive runs in order to simulate different substances;
- Finally, the profiles for trees and vegetation included in the software suite may not accurately represent local species.

7. CONCLUSIONS

In this paper, a method for using ENVI-met for the simulation of pollutant dispersion in a high density urban area has been explored. A section of downtown Curitiba was modeled, from readily available data – city plans and satellite photographs – and visual surveys.

Different model domains and resolutions have been tried; expansion of the model domain beyond 1 block on each side of the urban canyon and the increase in resolution – with the added expense of computation time – had no significant effect on the accuracy of the model.

It was found that, when representative wind data from within the UCZ is not available, standard wind speed measurements can be used as long as an adjustment is made to the input data, to account for the differences between the roughness contexts. This suggests that it is possible to use the model to predict conditions at the urban canyon from standardized climatic data.

NO_x dispersion from vehicle emissions was then modeled, from traffic volume, fleet composition and Brazilian emission standards. Three simulations scenarios were run: stagnation, northerly and easterly wind. NO_x concentrations were found to be significantly higher in the stagnation scenario; northerly wind conditions showed marginally better NO_x dispersion than the easterly wind scenario.

Notwithstanding its limitations, ENVI-met was found to be a viable tool for the study of airflow and pollutant dispersion in highly complex urban areas; this workflow could be effortlessly adapted to simulate other substances, such as SO_2 and particulate material. From this base scenario, alternative geometries – i.e., different land use and zoning regulations – and alternative traffic solutions could also be explored. Detailed analysis of the simulation results may highlight potential pollution hotspots and their relation to urban morphology and local weather characteristics.

We hope that by comparing on site measurements with predicted data, improvements can be made to the software model. As this research project progresses, future work will use the ENVI-met model and the workflow presented on this paper to analyze ariflow and pollutant dispersion in high density linear development areas in Curitiba.

Acknowledgments

Fellow researchers Francine A. Rossi, Flavia Minella, Cintia Tamura, David Q. de Sant'Anna and Simone C. Dias. We are specially thankful to the Brazilian funding agencies CNPq and CAPES.

References

1. Ali-Toudert, F., 2005. *Dependence of Outdoor Thermal Comfort on Street Design in Hot and Dry Climate*. Meteorologischen Institutes der Universität Freiburg. Universität Freiburg, Freiburg.
2. Allard, F. Santamouris, M. (Editors), 1998. *Natural Ventilation in Buildings*. James & James, London.
3. Bertotti, J. N., 2009. "40% dos veículos de Curitiba têm mais de 10 anos". *Gazeta do Povo*, Curitiba, 9th July 2009.
4. Brasil, 1993. "Lei Federal no. 8723, 28 de Outubro de 1993. Dispõe sobre a redução de emissão de poluentes por veículos automotores e dá outras providências".

Diário Oficial [da] República Federativa do Brasil, Brasília.

5. Brasil, 2002. Conselho Nacional do Meio Ambiente. “Resolução CONAMA no. 297/2002, 26 de Fevereiro de 2002. Estabelece os limites para emissões de gases poluentes por ciclomotores, motociclos e veículos similares novos”. *Diário Oficial [da] República Federativa do Brasil*, Brasília.
6. Bortoli, P. S., 2002. *Análise da poluição sonora urbana em zoneamentos distintos da cidade de Curitiba*. Master’s Dissertation – Programa de Pós-Graduação em Tecnologia. Federal Technologic University of Paraná, Curitiba.
7. Bruse, M., 2009. *ENVI-met website*. http:// www.envi-met.com.
8. Calixto, A., 2002. *Vehicle traffic noise on highways and major avenues situated within the urban perimeter of Curitiba, analyzed from the standpoint of acoustic parameters and their environmental impact*. Master’s dissertation in Mechanical Engineering, Federal University of Paraná, Curitiba.
9. Danni-Oliveira, I. M., 2003. “A Cidade de Curitiba e a Poluição do Ar: Implicações de seus atributos urbanos e geoecológicos na dispersão de poluentes em período de inverno”. In: Monteiro, C. A. F.; Mendonça, F. *Clima Urbano*. Contexto, São Paulo.
10. Hedquist, B. C. Brazel, A. J. Di Sabatino, S. Carter, W. Fernando, H. J. S., 2009. “Phoenix Urban Heat Island Experiment: Micrometeorological Aspects”. In: *Eigth Symposium on the Urban Environment. 89th American Meteorological Society Annual Meeting, Phoenix. Joint Session 12*, Available from: <http://ams.confex.com/ams/89annual/techprogram/paper_145712.htm>
11. Instituto Ambiental Do Paraná (IAP), 2009. *Relatório Qualidade do Ar na Região Metropolitana de Curitiba, Ano de 2008*. Paraná.
12. Instituto De Pesquisa E Planejamento Urbano De Curitiba (IPPUC), 2009. *Website*. http://ippuc.org.br.
13. Kessler, C. Niederau, A. Scholz, W., 2006. “Estimation of NO_2/NO_x relations of traffic emissions in Baden-Württemberg from 1995 to 2005”. In: 2nd conf. Environment & Transport, incl. 15th conf. Transport and Air Pollution. Reims, France, 12-14 June 2006. *Proceedings* no. 107, Vol. 2, Inrets ed., Arceuil, France.
14. Lamberts, R. Goulart, S. Firmino, S., 1998. *Dados Climáticos Para Projeto e Avaliação Energética de Edificações Para 14 Cidades Brasileiras*. Federal University of Santa Catarina, Florianópolis.
15. Monteiro, C. A. F., 2003. “Teoria e Clima Urbano: um projeto e seus caminhos”. In: Monteiro, C. A. F.; Mendonça, F. *Clima Urbano*. Contexto, São Paulo.
16. Oke, T. R., 1978. *Boundary layer climates*. Methuen, New York.
17. Oke, T. R., 2006. *Initial guidance to obtain representative meteorological observations at urban sites*. WMO , Geneva.
18. Spangenberg, J. et al., 2008. “Simulation of the influence of vegetation on microclimate and thermal comfort in the city of São Paulo”. *Revista SBAU*, Piracicaba, v. 3 n. 2.
19. Rabl, A., 2002. “Environmental benefits of natural gas for buses”. *Transportation Research Part D: Transport and Environment*, Volume 7, Number 6.
20. Wang, G. Bosch, F. H. M. Kuffer, M., 2008. “Modelling Urban Traffic Air Pollution Dispersion”. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*. Vol. XXXVII, Part B8.

Space Perception and Luminance Contrast: Investigation and Design Applications through Perceptually Based Computer Simulations

Nan-Ching Tai
University of Washington, Department of Architecture
Seattle, WA, 98195, USA
tai@u.washington.edu

Mehlika Inanici
University of Washington, Department of Architecture
Seattle, WA, 98195, USA
inanici@u.washington.edu

Keywords: Space Perception, Depth Cue, Lighting Simulation, Psychophysical Experiments, High Dynamic Range Imagery.

Abstract

Pictorial cues are the visual information gathered from 3D scenes; and they provide depth perception in the physical world. Pictorial cues are also used to create the illusion of depth on planar media. Planar media are a common platform for architects to visually examine the spatial qualities of their designs. Therefore, knowledge of pictorial cues can be used as a design strategy to enrich the spatial experience. In this paper, luminance contrast is proposed as an effective depth cue and design strategy. Lighting based perceptual studies are challenged by the dynamics of the luminous environments in physical experimental settings. Computer simulation allows the study of lighting variability throughout the day and year in a systematic manner. This paper utilizes a computational framework to simulate perceptual reality. Psychophysical experiments are conducted in this alternative environment. 3D scenes and the resulting 2D imagery are utilized to investigate the impact of lighting patterns and luminance contrast on depth perception. The results of the study demonstrate that luminance variations within a space impacts the perceived distance as much as they impact the luminance contrast between the task and the background. Application of this pictorial cue is demonstrated through architectural and urban design examples.

1. INTRODUCTION

The human visual system processes three-dimensional (3D) physical environments based on the two-dimensional (2D) images projected on to retinae [Palmer 2000]. Conversely, the process and products of 2D representation systems are used to design and develop 3D un-built proposals, and to provide instructions for constructing the projects [Ching 1998]. Pictorial cues gathered from the physical world supplement the 2D retinal images with additional visual information to perceive the spatial depth. These pictorial cues are also used to create the illusion of depth on planar media. Therefore, knowledge of pictorial depth cues has been used to enrich spatial experience within given structural boundaries. However, the ability to examine the spatial quality from design alternatives is limited by the perceptual realism offered by the representation media.

Although visual perception is a complex process that is yet to be fully understood, it is generally agreed that the visual system has two distinct functions: i) identification (i.e. color perception, object and face recognition) and ii) localization (i.e. perception of motion, spatial relation and depth). The localization system is insensitive to color and responds only to luminance differences [Ferwerda 2001; Livingstone 2002].

In space perception, edge detection provides conceptually meaningful visual information, and facilitates the understanding of the geometrical properties of the 3D scene. The geometrical properties are governed by size-distance relationships, i.e., the perceived size of an object is derived from its perceived distance, and supplemented with the pictorial cues of size perspective that include relative size, familiar size, linear perspective and texture gradient.

Reproduction of depth cues based on detected edges is straightforward and does not require high levels of perceptual realism. Therefore, it is relatively easy to demonstrate the effect (Figure 1). In Figure 1a, three cylinders located in the hallway have exactly the same projected size. However, the pictorial cues of cylinders' relative size, convergent lines, and foreshortening of texture pattern inform the viewer that the three cylinders are located at different locations. The further the cylinder is located, the larger it appears. In Figure 1b, the effect of the pictorial cue of shadow is illustrated. The three cylinders are floating at the same distance from the viewer, and the presence and the length of the shadows inform their vertical positions. The cylinders appear the same size.

Lighting has been identified as a pictorial depth cue. However, its application is often limited to the resultant shadow to locate the objects in a 3D context. Schwartz and Sperling [1983] promoted perceived luminance as a pictorial depth cue. O'Shea, Blackburn and Ono [1994] further pointed out that target with higher luminance contrast would appear closer on a 2D setting. Authors [Tai and Inanici 2009a] previously investigated this effect in a 3D context. In a computer-simulated environment, physical structures were manipulated to introduce different lighting conditions in an architectural scene. Effect of luminance distributions on depth perception was studied through psychophysical experiments. It was concluded that lighting distributions can affect the perceived distance of the visual target. Figures 1c and 1d illustrate this effect in the hallway illusion. With the introduction of additional lighting, the perceived distance of the farthest cylinder in Figure 1d is increased in comparison to the same configuration in Figure 1c.

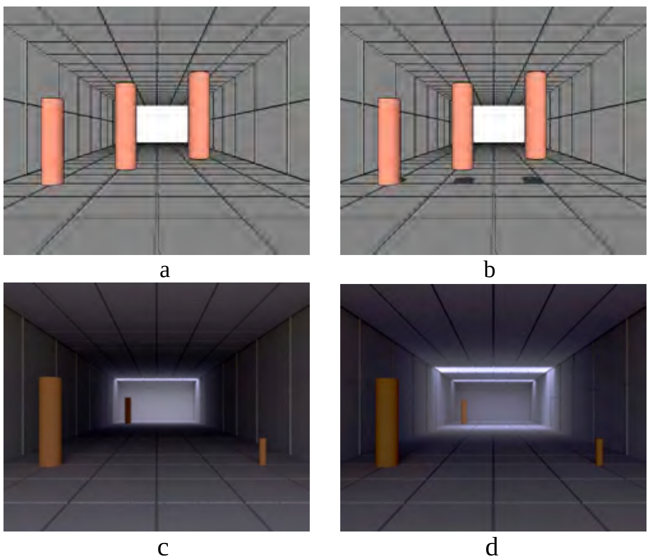


Figure 1. Hallway illusion illustrating a) the size-distance relationship, b) the pictorial cue of shadow, c) and d) the pictorial cue of luminance distributions

Luminance distributions in a scene determine the lighting patterns around the visual target, and the luminance contrast between the target and its background. The objective of this research is to study the individual contributions of lighting patterns and luminance contrast on depth perception. This study extends the research framework previously developed [Tai and Inanici 2009a and 2009b] by taking advantages of two-dimensional nature of the computer imagery. The lighting patterns and luminance contrast were isolated using computational methods. Psychophysical experiments were conducted to investigate the particular contribution from each component.

2. METHOD

Using computer simulations to demonstrate and study depth perception is not a new approach. Meng and Sedgwick [2001] used a computer environment to manipulate the presence of shadows in their research to investigate its impact on depth perception. Determining the hard shadows of objects in a computer environment is a fairly simple mathematical calculation that is based on the geometric relationship of the object, light source and the viewer. Photorealistic rendering methods used in Meng and Sedgwick's study provided the necessary realism of representing forms and casted shadows. However, the study of lighting distribution and the investigation of its impact on depth perception require a representation that matches the intended perceptual reality.

The Radiance lighting simulation and visualization system is a physically based rendering program that models the light transport and material properties based on their governing physical equations [Ward and Shakespeare 1997]. The rendered image is a luminance map that encompasses numerically accurate lighting data [Mardaljevic 2001; Ruppertberg and Bloj 2006]. Current display devices cannot display high dynamic range lighting information that is simulated through physically based rendering tools. Tone mapping operators were developed to compress the full range of lighting data that can be processed by human visual system (14 logarithmic units in cd/m^2 from starlight to sunlight) to the range of conventional display devices (2 logarithmic units). Using appropriate tone mapping operators, physically accurate, high dynamic range images can be displayed as low dynamic range perceptually realistic scenes on conventional display devices. Physically and perceptually accurate images have previously been utilized in experimental studies to investigate various aspects of lightness, color, shape, and depth on visual perception [Boyaci et al. 2003; Doerschner et al. 2004; Delahunt and Brainard 2004; Fleming et al. 2004; Tai and Inanici 2009b].

Experimental scenes used in this study were generated by the Radiance lighting simulation and visualization system and tone mapped by the Photographic tone mapping operator [Reinhard et al. 2002]. The Photographic tone mapping operator was selected based on the research that compared perception between physical scenes and tone mapped imagery [Cadik 2008].

3. PSYCHOPHYSICAL EXPERIMENTS

The effect of a pictorial depth cue is often studied through psychophysical experiments. Psychophysics is a field that investigates the relation of psychological sensation and physical stimuli [Gescheider 1984]. In general, subjects are asked to judge the perceived distance of the visual targets from the test and reference scenes. By comparing the measured differences, the effect of a depth cue controlled in the experimental scenes can be systematically studied.

The method of constant stimuli is employed in our study to measure depth perception. In this method, both the standard stimulus and comparison stimulus are presented to the subjects simultaneously. While the standard stimulus remains constant, the comparison stimulus changes from trial to trial. Subjects' responses are recorded based on their response to the varied stimuli. In this method, the comparison stimulus usually has five, seven or nine values, separated by equal distances in physical scale. The comparison stimulus is varied such that variation of greatest magnitude is almost always perceived greater than the standard stimulus, and the least magnitude is perceived less than the standard stimulus. Pairs of standard stimulus and comparisons are presented in a random order; and subjects are asked to determine which stimulus produces a greater magnitude of sensation [Gescheider 1984]. The design of the binary responses allows the data to be analyzed with probit regression models to derive psychometric functions, as illustrated in Figure 2. The 0.5 point in the regression model is called the point of subjective equality (PSE). It represents the value of comparison stimulus that is subjectively perceived equal to the standard stimulus. In other words, at the point of subjective equality, measured sensory perception is equal to the physical stimuli [Finney 1971].

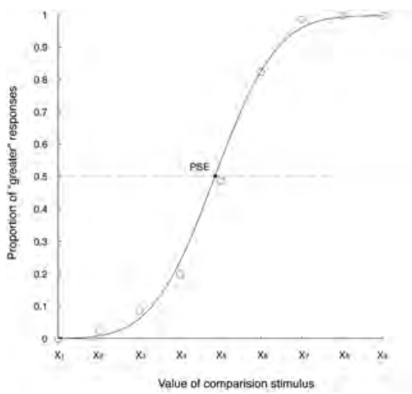


Figure 2. Typical psychometric function fitted with Probit regression curve [Gescheider 1984]

3.1. Experiment Design

As the size perspective is a dominant depth cue for the space perception of physical space, it is essential to ensure that the size-distance relationship can be observed effectively in the computer-generated pictorial space. The authors [Tai and Inanici, 2009b] previously conducted a computer based study to reproduce a classic experiment [Holway and Boring 1941] that investigates the size-distance relationship in a physical environment. In the Holway and Boring experiment, circular disks are used as visual targets. By acquiring similar results in a computer generated study in comparison to the original one, it is demonstrated that the utilized computational framework is a reliable alternative environment to study depth perception.

The use of circular disks as visual targets is also adopted for the current study. Figure 3 illustrates the setup, where architectural configurations are modified to study the impact of lighting distributions on spatial depth. The test and reference scenes are presented simultaneously. One of the hallways serves as a test scene and the other as a reference scene. A different lighting condition is introduced into one of the hallways. For each particular condition, seven images were generated. In each of these images, the test target (standard disk) remains constant at a fixed distance, while the reference target (comparison disk) is rendered at one of the seven different locations (Figure 4). Subjects are asked to report which visual target is closer. The underlying concept of the experiment design is to identify when the test and reference targets are perceived equal in depth. By comparing the perceived distance and its actual location, the effect of lighting distribution on depth perception can be revealed.

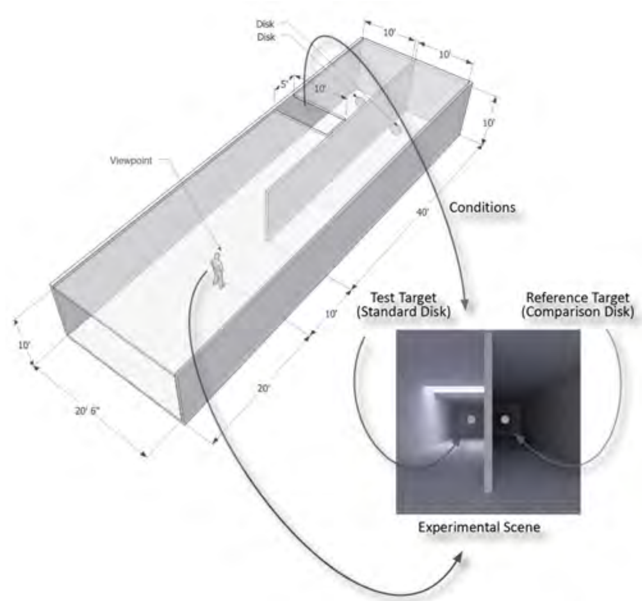


Figure 3. Experimental setup

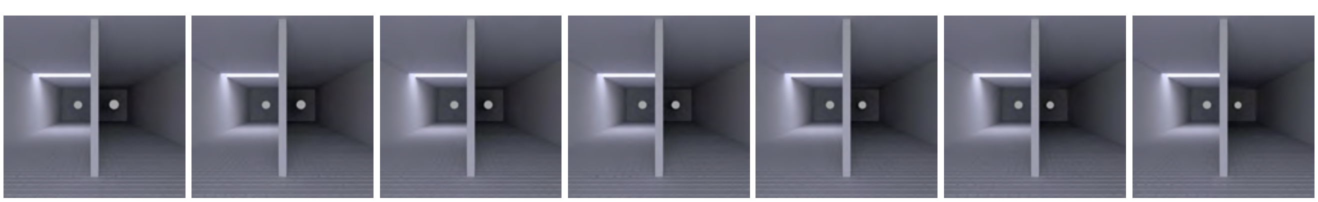


Figure 4. Experimental scenes with different configurations of visual targets under the same lighting condition

Two experiments are conducted in this study. In the first experiment, the setup is manipulated to create different lighting conditions (Figure 5). Each scene is illuminated with the daylighting admitted from the rear end of the corridor. Three different configurations are rendered (single skylight, two skylights and three skylights) with the two different sky conditions (intermediate sky with / without sun patch) to generate total of six different setups: single skylight (1sky), single skylight with sun patch (1skyS), two skylights (2sky), two skylights with sun patch (2skyS), three skylights (3sky), and three skylights with sun patch (3skyS).

The second experiment is conducted to investigate the individual contributions of the lighting patterns around the visual target and the luminance contrast between the target and its immediate background. The renderings of single skylight (1sky), single skylight with sun patch (1skyS) and three skylights with sun patch (3skyS) from the first experiment are adopted, and image processing techniques are used to generate six different setups as shown in Figure 6. In Figures 6a, 6b, and 6c, the portion of the disk and the back wall are edited so that the luminance contrast of the disk and its background are identical between the right and left corridors. This set equalizes the luminance contrast between the target and its background, leaving lighting distribution patterns as the variable. In Figure 6d, 6e, and 6f, the presence of the skylight and the resulting lighting distribution patterns surrounding the disk were removed. The luminance contrast between the disk and its background is the only variable between the right and left hallways of each scene.

3.2. Procedure

In all conditions, the location of the left disk is fixated at 40’ away from the viewpoint and it is referred as the standard disk. The location of the right disk is varied to create 7 different alternatives (located 34’, 36’, 38’, 40’, 42’, 44’ and 46’ away from the viewpoint). The right disk is referred as the comparison disk. Various disk locations are simulated under twelve different lighting conditions. To avoid lighting effect always coming from one side (i.e. left hallway), the rendered scenes were flipped horizontally to create another set.

Images were displayed on the center of a LCD display in a dark room. Eight subjects participated in each experiment. Subjects were aged between 21 and 36, had normal or corrected-to-normal vision. They were given enough time to adapt to the dark environment. They sat in front of the display at a normal viewing distance and angle. They were instructed that in each scene there were two identical disks floating at the center of left and right hallways; and were asked to provide a quick response. In each session, images were shown in a random order ten times. Participant’s responses were recorded as the number of times that the standard disk is reported closer.

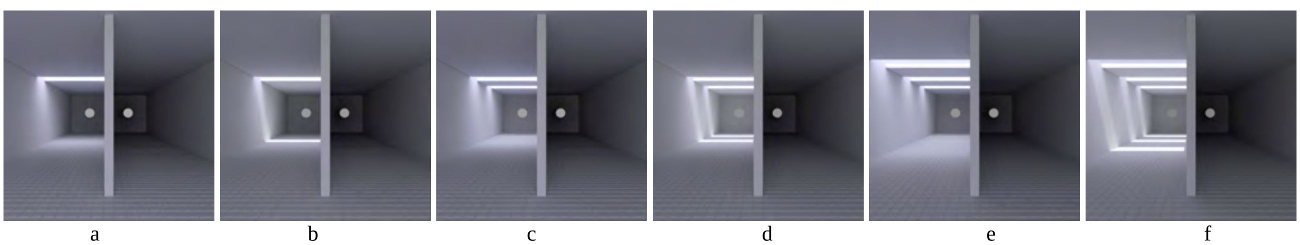


Figure 5. Experimental scenes with right and left disks configured at same locations under different conditions for experiment 1: (a) Single skylight; (b) Single skylight with sun patch; (c) Two skylights; (d) Two skylights with sun patch; (e) Three skylights; (f) Three skylights with sun patch.

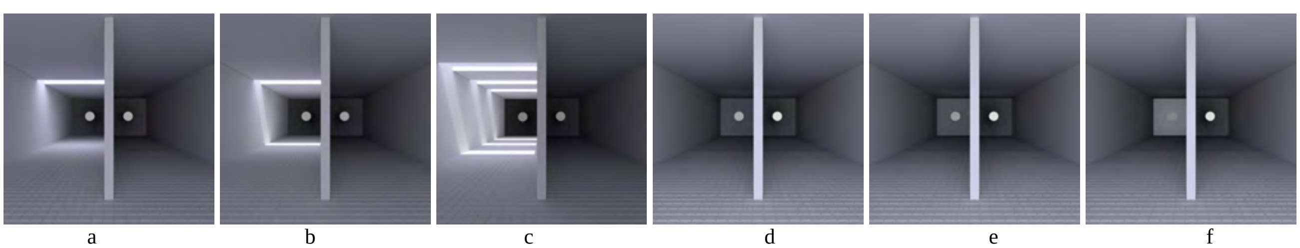


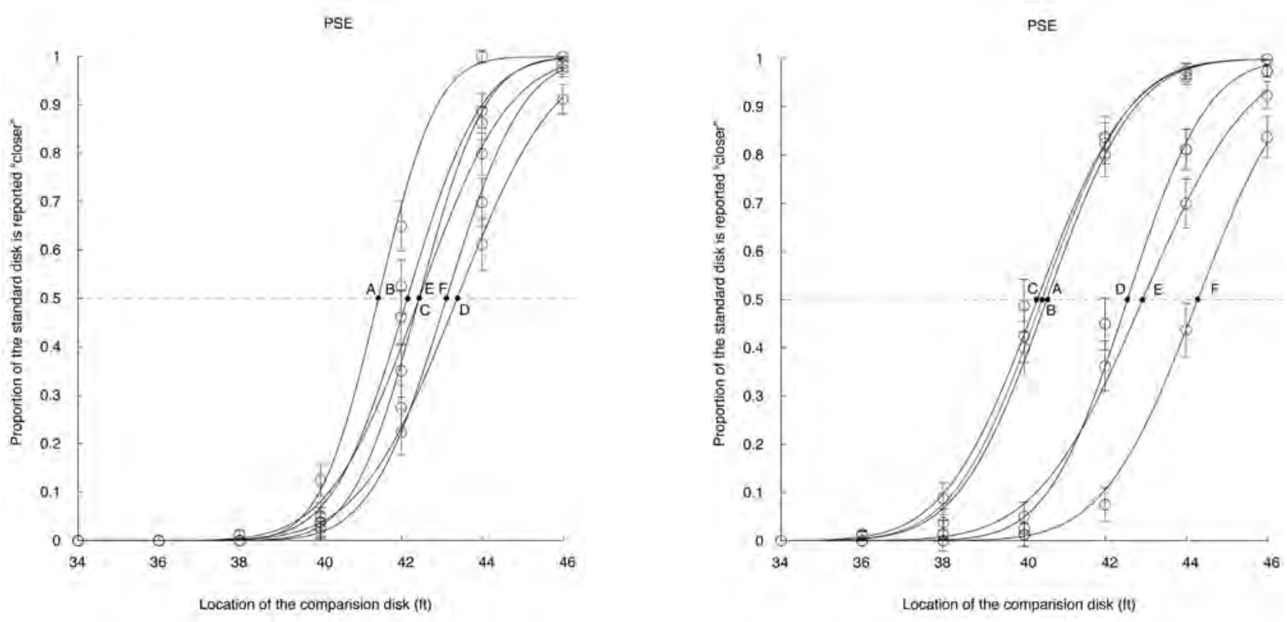
Figure 6. Experimental scenes with right and left disks configured at same locations under different conditions for experiment 2. Contrast between disk and the background is equal, variable is the luminance distribution patterns in the hallway: (a) Single skylight; (b) Single skylight with sun patch; (c) Three skylight with sun patch. Lighting distribution patterns are eliminated; variable is the luminance contrast between the disk and background: (d) Single skylight; (e) Single skylight with sun patch; (f) Three skylight with sun patch.

3.3. Results and Analysis

The experiment results are illustrated in Figures 7a and b for experiments 1 and 2, respectively. Subjects' responses are plotted as the proportion of scenes on which the standard disk is reported "closer" as a function of the comparison disk's location. The probit regression curve is fitted to each data set. The intersection points of each curve with the 50% proportion line are taken as the PSE. The PSE represents when the right and left disks are perceived equal in depth. In other words, it is the measurement of the perceived distance of the standard disk. In experiment 1, the perceived distance of standard disk was increased from 40 feet to 41.432 ± 0.125 (5a), 42.150 ± 0.141 (5b), 42.438 ± 0.133 (5c), 43.377 ± 0.165 (5d), 42.445 ± 0.156 (5e) and 43.109 ± 0.143 (5f) feet, respectively. In experiment 2, the perceived distance of the standard disk of the setups shown in Figure 6a, 6b, 6c was increased slightly from 40 feet to 40.440 ± 0.154 (6a), 40.554 ± 0.154 (6b) and 40.324 ± 0.158 (6c) feet respectively. For setups shown in Figure 6d, 6e, 6f, the perceived distance of the disk was increased from 40 feet to 42.547 ± 0.146 (6d), 42.926 ± 0.171 (6e) and 44.284 ± 0.166 (6f) feet respectively.

The results indicate that luminance contrast significantly increases the perceived distance of the visual target. These results are in agreement with O'Shea et al. study [1983], which argued that the higher the luminance contrast, the closer the target appears to be on a 2D context. Our study demonstrates that the same effect can be observed in 3D environments. In the experimental setup, daylighting introduced by the skylight increases the overall dynamic range of the luminance distribution; as a result, it decreases the luminance contrast between the visual target and its immediate background. Thus, it increases the perceived distance of the visual target and the overall spatial depth of the hallway.

Figure 8 further illustrates the impact of the luminance contrast on perceived distance. The dashed line represents the ratio of the luminance contrasts from original HDR images, and the straight line represents the same ratio from the tone mapped images. It is important to note that tone mapping conserved the contrast within the scenes. In both cases, the perceived distance increases as the ratio decreases.



a b

Figure 7. Probit analysis. (a) PSE is given for A "single skylight", B "single skylight with sun patch", C "two skylights", D "two skylights with sun patch", E "three skylights", F "three skylights with sun patch". (b) PSE is given for A "single skylight with distribution variable", B "single skylight with sun patch and distribution variable", C "three skylights with sun patch and distribution variable", D "single skylight with contrast variable", E "single skylight with sun patch and contrast variable", F "three skylights with sun patch and contrast variable".

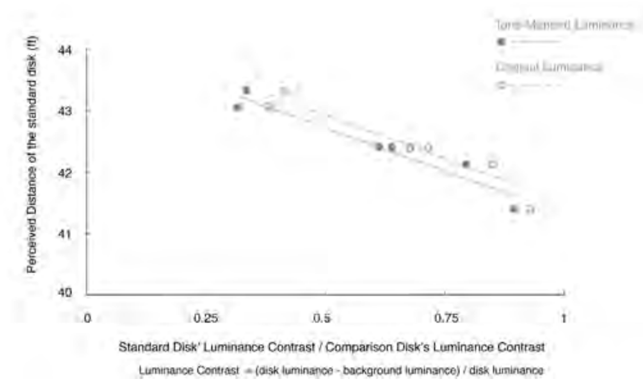


Figure 8. Impact of luminance contrast ratio on perceived distance

The results of the experiment also indicate that lighting patterns alone has insignificant effect on increasing the perceived distance of the visual target. Comparison between the measured perceived distances between the conditions shown in 5b (single skylight with sun patch: 42.150 ± 0.141 ft.), 5f (three skylight with sun patch: 43.109 ± 0.143 ft.) and the same conditions with the sun patch removed shown in Figure 6e (42.926 ± 0.171 ft.), 6f (44.284 ± 0.166 ft.), suggests that subjects tend to overestimate the perceived distance of the visual target when the lighting patterns are absent. A possible interpretation is that the patterns of sun patches provide pictorial cue of linear perspective to correctly judge the depth. This effect is a compounding factor. A further study is planned to study the impact of luminance distribution patterns that impact other pictorial cues such as the linear perspective.

4. DISCUSSIONS & APPLICATIONS

Computer simulations with appropriate algorithms can provide pictorial representations with adequate realism to study the perceptual qualities of a space. In this study, we have demonstrated its utilization to identify the effect of the luminance contrast on spatial depth in an experimental setting. Two examples are discussed to demonstrate the application of research findings in architectural design and urban planning problems.

Figure 9 illustrates a planning conflict between street expansion and the preservation of a vernacular Taiwanese temple. In the route of ritual ceremony in a vernacular Taiwanese temple, three particular scenes are expected to be perceived as part of the spatial experience (Figure 9b, c, d). Each scene is a composed perspective view rendered with repeated patterns of light and dark. However, the light pattern distributions in particular scene are not intended to faithfully inform the visitor about the actual distance of the temple at each viewpoint. Instead, it is intended to exaggerate the sense of depth [Tai 2003].

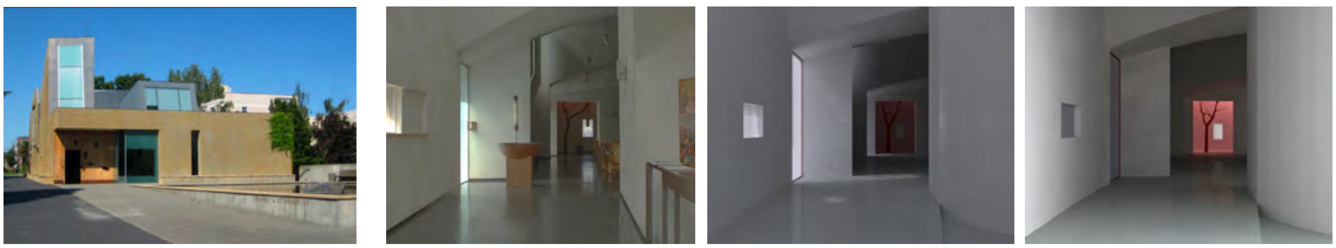
The three scenes presented in Figure 9b, c, d have similar lighting distribution patterns investigated in this study. In the first view, the visual target of the deity is rendered with repeated patterns of light and dark, similar to the two skylights setup in the first experiment. In the second view, a single skylight pattern is experienced. In the last standpoint, skylight is not visible. According to the experiment results, the spatial depth is exaggerated in the first two views, but the factor of exaggeration is removed in the last scene. As a result, the “sense of the depth” is created by repeating lighting patterns in earlier scenes. Conversely, the “sense of depth removal” pulls the visitor away from the physical site context into the ritual realm as one continues to move from one scene to the other in a very short walk.

The spatial experience in the temple is a result of the daylighting scheme introduced through the courtyard. Unfortunately, the front hall and courtyard of the temple were demolished to accommodate the growing traffic, leaving only the main hall preserved at its original location. The demolished parts eliminated the scenes that provide the repeated patterns of light and dark. If the resultant luminance contrast introduced by the courtyard has been recognized as the factor that enriches the spatial depth and its value as an integral part of a vernacular temple’s spatial experience, the decision may have been made differently for this planning conflict.

The research findings can also be applied as design parameter. The chapel of St. Ignatius in Seattle (designed by Steven Holl) is known for its rich lighting variation and spatial qualities (Figure 10). The spatial dimension in the chapel is enriched by dynamic lighting distributions. Figure 10b illustrates the very first view upon entry, and demonstrates the utilization of luminance contrast to enrich the spatial depth. Figures 10c and 10d demonstrate a simulation of the chapel with a tree that is framed by an opening at the end of the space. The tree serves as a visual target. In this view, daylight coming from the side renders the framed picture of tree with patterns of light and dark, and creates a sense of ‘deepness’. Comparison of the computer simulations in Figure 10c and 10d demonstrate the effect of luminance contrast on spatial depth. In Figure 10d the side window is covered, and therefore, the side lighting is removed from the scene. The luminance contrast between the tree and its background is increased. The increased contrast diminishes the perceived distance. As a result, the visual target (tree) in Figure 10d appears to be closer than Figure 10c. Conversely, brightness in the foreground reduces the luminance contrast between the visual target and its immediate background in Figure 10c, and thus increases the overall spatial depth.

a b c d

Figure 9. Planning conflict between the street expansion and the perseveration of a historical temple, a: outside view of the main hall; b and c: the eliminated views that demonstrate the sense of deepness; d: the surviving view of the main hall



a b c d

Figure 10. Chapel of St Ignatius. a: exterior view; b: interior view; c: computer rendering of entrance view with the side window; d: computer rendering of entrance view without the window light

5. CONCLUSIONS

This research demonstrates the utilization of a simulation approach to study the perception of spatial depth in pictorial spaces. This approach is useful as a methodology for studying space perception and for evaluating the perceptual qualities of spaces throughout the design phases. The contributions of this study can be summarized at practical and theoretical levels:

1) At the practical level, luminance contrast is identified as an effective pictorial cue. Occupants navigate built environments through a planned circulation, resulting in sequential experience of one scene after another. Careful manipulation of the physical configuration can create scenes with intended lighting distributions and luminance contrasts that can enrich the spatial experience of built environments.

2) Luminance variations within a space impacts the perceived distance as much as they impact the luminance contrast between the target and the background. However, a further study is planned to study the impact of luminance distribution patterns on other pictorial cues such as linear perspective.

3) In the theoretical level, physically and perceptually accurate simulation techniques bridge the gap between perceptual qualities of built environments (physical space) and their representation in pictorial spaces. The perceptually accurate simulations provide an alternative environment that offers flexibility and precise parameter control of scene components for 3D scenes and the resulting 2D imagery.

Acknowledgements

The authors would like to thank to Dr. John Palmer at the University of Washington, Department of Psychology for his suggestions and discussions on conducting the experiment. The authors also thank to the students and staff of the College of Built Environments at the University of Washington who have participated in the experiments.

References

Boyaci, H., T. Maloney, and S. Hersh. 2003. “The Effect of Perceived Surface Orientation on Perceived Surface Albedo in Binocularly Viewed Scenes.” *Journal of Vision,* 3: 541-553.

Cadik, M., M. Wimmer, L. Neumann, and A. Artusi. 2008. “Evaluation of HDR Tone Mapping Methods Using Essential Perceptual Attributes.” *Computer & Graphics,* 32: 35-44.

Ching, Francis D.K. 1998. *Design Drawing*. Van Nostrand Reinhold, New York.

Delahunt, P. and H. Brainard. 2004. “Color Constancy under Changes in Reflected Illumination.” *Journal of Vision,* 4: 764-778.

Doerschner, K., H. Boyaci, and T. Maloney. 2004. “Human Observers Compensate for Secondary Illumination Originating in Nearby Chromatic Surfaces.” *Journal of Vision,* 4: 92-105.

Ferwerda, J. 2001. "Elements of Early Vision for Computer Graphics." *LEEE Computer Graphics and Applications*, 21: 22-33.

Finney, D. 1971. *Probit Analysis*. University Press, Cambridge.

Fleming, R.W., A. Torralba, and E.H. Adelson. 2004. "Specular Reflections and the Perception of Shape." *Journal of Vision*, 4: 798-820.

Gescheider, George A. 1984. *Psychophysics: Method, Theory, and Application*. Lawrence Erlbaum, NJ.

Holway, A.H. and E.G. Boring. 1941. "Determinants of Apparent Visual Size with Distance Variant." *American Journal of Psychology*, 54: 21-37.

Livingstone, Margaret. 2002. *Vision and Art: The Biology of Seeing*. Harry N. Abrams, New York.

Mardaljevic, J. 2001. "The BRE-IDMP Dataset: A New Benchmark for the Validation of Illuminance Prediction Techniques." *Lighting Research and Technology*, 33(2): 117-136.

Meng, J.C. and H.A. Sedgwick. 2001. "Distance Perception Mediated through Nested Contact Relations among Surfaces." *Perception & Psychophysics*, 63(1): 1-15.

O'Shea, R.P., S.G. Blackburn, and H. Ono. 1994. "Contrast as a Depth Cue." *Vision Research,* 34: 1595-1604.

Palmer, S.E. 1999. *Vision Science: Photons to Phenomenology*. MIT Press, Cambridge.

Reinhard, E., M. Stark, P. Shirley, and J. Ferwerda. 2002. "Photographic Tone Reproduction for Digital Images." *ACM Transactions on Graphics*, 21(3): 267-276.

Ruppertberg. I. and M. Bloj. 2006. "Rendering Complex Scenes for Psychophysics using RADIANCE: How Accurate Can You Get?" *Journal of Optical Society of America*, 23: 759-768.

Schwartz, B.J. and G. Spering. 1983. "Luminance Controls the Perceived 3D Structure of Dynamic 2D Displays." *Bulletin of the Psychonomic Society*, 21: 456-458.

Tai, N. 2003. *Continuity and Innovation: A Reinterpretation of the Traditional Taiwanese Temple.* Master of Architecture Thesis, University of Washington.

Tai, N. and M. Inanici. 2009a, "Depth Perception as a Function of Lighting, Time, and Spatiality." *Illuminating Engineering Society (IES) 2009 Conference*, Seattle.

Tai, N. and M. Inanici. 2009b, "Depth Perception in Real and Pictorial Spaces: A Computational Framework to Represent and Simulate the Built Environment." *CAADRIA* 2009, Yunlin, 543-552.

Ward, G. and R. Shakespeare. 1997. *Rendering with Radiance*. Morgan Kaufman Publisher, San Francisco.

Architect-oriented

Towards 'Architect-friendly' Energy Evaluation Tools

Lieve Weytjens[1,2], Griet Verbeeck[1]
[1] Department of Arts and Architecture - University College Limburg [PHL] - Belgium
[2]Hasselt University - Belgium
Lieve.Weytjens@phl.be

Keywords: survey, interviews, design process, building performance

Abstract

Several studies demonstrate the importance of integrating energy analyses in the architectural design process. For this, a wide variety of building performance simulation programs are currently available. However, most of these tools are not in tune with the architects' approach and are not suitable for early design stages, when major decisions are made. Despite recent developments addressing the use of simulation in the design process (DP), the uptake by architects is still limited. Further, the importance of user's needs has repeatedly been reported in literature, but has rarely been investigated from an architect's point of view. However, researchers and design tool developers need this information to improve the compatibility of existing tools with the architectural DP and to develop new tools.

This research focuses on the use of energy evaluation tools by architects, and aims at a better understanding of the architect's preferences for these tools. The concept of 'architect-friendliness' was studied via interviews with 9 Flemish architects, and via a large-scale survey among 629 Flemish architects. The results provide important information for the development of energy tools that better fit the DP and better meet the architect's expectations. Finally, a conceptual framework specifying the concept of "architect-friendliness" is proposed.

1. INTRODUCTION

Different studies have demonstrated the potential role of building performance simulation (BPS) in making informed design decisions [1,2], and many researchers focus on the integration of BPS in the design process (DP) in various ways [3-7]. However, in practice, the uptake of simulation by architects is still limited [8,9] and many early decisions are entirely based on the designer's experience and intuition [10, 11]. Especially in small projects that lack engineering support due to limited budgets, BPS might be a valuable support to architects in the design of low-energy buildings. Previous studies have demonstrated reasons for not using BPS by architects. These covered among other things the architects' perception that the usage is beyond the scope of their work, professional shortcomings, a lack of know-how, and steep learning curves [8,9]. Major limitations of existing BPS tools included lack of user-friendliness, lack of integration with CAD software, extensive data-input, and the fact that they do not fit the DP.

Accordingly, most BPS tools are not adapted to the architects' approach. However, given the urgent question of climate change and stricter energy regulations, the use of BPS by architects is expected to be increasingly important. Therefore, to improve the uptake of BPS by this (growing) users group, an in-depth understanding of their needs and requirements is crucial. Better integration with CAD software, limited data-input in the architect's language, easy interpretable output, graphical representation of input and output, and use of defaults are among other frequently identified criteria for BPS tools [8,12,13]. The user-friendliness is probably the most often discussed tool criterion [13,14], but has rarely been investigated from an architect's point of view. Attia et al. [12] conducted a survey among simulation users in the U.S.A. to investigate the most important criteria for *'architect-friendly' [12]* BPS tools, but the context might be different in other countries.

To obtain a more elaborated view on the 'architect-friendliness' of energy simulation tools and the users' needs and preferences, semi-structured interviews were performed with 9 Flemish architects, and a large-scale survey was conducted among 629 architects in Flanders, Belgium. In both the interviews and the survey, a broad view was adopted for energy simulation tools, also including other types such as checklists, and is therefore referred to as 'energy (evaluation) tools'. A conceptual framework is proposed, providing tool developers important information for the development of energy tools that better fit the DP and better fulfill the architect's requirements.

2. INTERVIEWS

2.1. Method

Semi-structured interviews were performed with 9 Flemish architects, between June 2008 and October 2008, enabling the architects to actively contribute aspects that seem important to them and allowing them to elaborate more on their views, perceptions, and motivations.

The interviews were structured among three major themes, namely the DP, the integration of sustainability aspects in the design and the usage of and requirements for energy evaluation tools. This paper only reports on the last theme. Several topics were related to the EPB software ("Energy Performance and Indoor Climate"), a steady-state simulation program provided by the Flemish government to calculate compliance with regulations related to the European 'Energy Performance of Buildings Directive' (EPBD) [15]. The EPB legislation enforces a maximum energy performance level (E-level) and insulation level (K-level). The E-level stands for the level of primary energy use, calculated for standard climate conditions on a monthly basis and standard occupants' behavior. This primary energy use is compared to a reference value. The current legal requirement sets the maximum E-level at E80. The K-level takes into account the mean U-value of the building envelope and the volume compactness. The maximum is set at K45, representing a U_{mean}= 0,45 W/m²K, for a compactness of 1m. As an example, a passive house represents a K10 to K15 level and a E30 to E40 level, whereas a low-energy house with a yearly energy consumption of ca. 50 kWh/m² has an insulation level around K35 and an E-level of E60 to E70. Since most Flemish architects are familiar with the EPB software, this tool was used as a reference in the interviews. Consequently, architects were able to discuss experienced limitations of the software and future needs for energy tools.

2.2. Results

The interviews revealed that tool simplicity, intuitive tool usage, and limited time to operate a tool are extremely important aspects to architects. However, differences were identified for the latter, ranging from 5 minutes to 2 hours to operate a tool. One respondent stated that a truly user-friendly tool should be usable without consulting a manual, and explained the concept of 'simplicity' by referring to the 3D sketch tool "Sketch-Up". Sketch-Up is perceived as a very easy and intuitive to use tool, with easy navigation and a simple and clear interface, allowing to quickly create and compare different design ideas.

Apart from the ability to quickly create, test and compare alternative designs, the possibility to provide feedback on the impact of design parameters also appeared to be a very important criterion for energy tools. In particular, the impact of glazing percentage on the risk of overheating and the impact of insulation on the building's energy use were issues for which architects require additional feedback when designing.

Regarding the need for integrating energy tools with 3D CAD software, approximately half of the respondents considered an integration advantageously, while the other half perceived it to be unnecessary and preferred a simple manual input. There was also no clear agreement among the respondents on the need to provide design guidelines.

The data further identified a limited and quick data-input, and easy data review as important issues for an 'architect-friendly' energy tool. The laborious and time-consuming process of data-input and the difficult data review were the most frequently expressed limitations of the current EPB software, and hinder the usage of the software when designing since most design projects are often subject to change. Several ideas on reducing the input were elaborated during the interviews, such as using glazing percentages per orientation or standard constructions for walls, floors and roofs. The latter could easily be introduced by providing adaptable default values per project type.

Considering the output, the architects interviewed showed a particular interest in clear, simple and visual output, especially to support the communication with clients. In this frame, a cost-benefit analysis and payback time appeared to be important factors to convince clients, as many decisions are determined by the client's budget.

Other preferences are the visual aspect of input, output and interface, the transparency of the tool, the ease of navigation, and the conformance with regulations. Tools should also be non prescriptive and flexible in use.

Finally, other tool criteria introduced by the interviewees included an extensive component library, easy interpretable output and ease of learning. Regarding the usability of tools in the DP, some respondents emphasized instant feedback on design decisions and changes, and data-input adapted to (early) design phases. Accuracy of the results appeared to be of less importance than ease of use.

3. SURVEY

3.1. Introduction

In addition to the interviews, a survey was conducted, evaluating the architects' needs for energy tools among a larger group of respondents to corroborate the former results and to obtain more quantitative data.

In the past, several surveys have been conducted to investigate the architects' use of and requirements for BPS tools [8,9,12], but the context might be different in other countries. In Flanders, Belgium, energy legislation is very recent with the first energy code implemented in 1992. Further, in Flanders, architects often work independently or in small firms, mainly focusing on the design of small projects for private clients. However, tool developers may benefit from input from multiple views and contexts, as already indicated by Mahdavi [9].

3.2. Method

A self-administered questionnaire was distributed among 984 architects in Flanders. The questionnaire was structured around 4 major parts. This paper reports on the results of the first part, being the architects' use of and preferences for energy tools. Several questions were related to the EPB software, because most Flemish architects are familiar with this software. The survey was conducted during the course "Energy Conscious Architect", organized by the Flemish government in cooperation with the two main Flemish architects associations, and comprised three evenings of lectures on energy efficiency.

After a pilot study in June 2008, the final questionnaire was distributed to 319 architects attending the first series of the course in September and October 2008. The response rate was 70%. For this series, the number of participants was limited and at that time no extra courses were planned. Due to the success of the first series of the course, a second series was organized from December 2008 until February 2009. For this series more participants could subscribe and this time the subscription limit was not reached. A slightly updated version of the questionnaire was distributed to all 665 participants, with a response rate of 61%.

In total, 629 questionnaires were completed and returned, representing a response rate of 64%. 70% of the respondents was male and the median age was 36. A comparison to statistical data of all Flemish architects (6985 architects, of which 71,5% male)[1] indicates that the current sample is sufficiently representative for the Flemish architects' population. The questionnaire consisted of multiple choice questions for which the respondents could select several options and also provide additional options. In addition, the questionnaire included one open question on the user-friendliness of energy tools. Some questions are partly based on Lam et al. [8] and Mahdavi et al. [9], with adaptation to the Belgian context. The statistical program SPSS 16 was used for the analyses.

3.3. Results: Multiple Choice Questions

3.3.1. Use of Energy Tools in Daily Architectural Practice

The data revealed that the EPB software is used by 64% of the architectural practices of the respondents surveyed. This high usage can be explained by the legislative character of the software. Other energy tools (such as checklists, simulation software, software provided by construction firms, etc.) are only used by 10% or less. Of all tools, BPS software clearly scored lowest, with only 2% of usage.

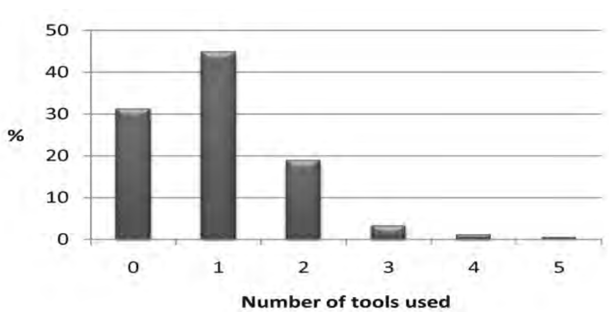


Figure 1. Number of energy tools used

[1] Data provided by the Order of Architects (Belgium)

As Figure 1 shows, the average number of tools used in the architectural practices of the respondents is 1 (45%), whereas 31% of the respondents does not use energy tools at all. Only 24% uses more than 1 tool to evaluate the energy performance of the design.

The data further revealed that the EPB software clearly dominates for single tool usage (92%), which can be explained by its legislative character. As a result, most of the architectural practices either use no tool or use the EPB software. Consequently, the current architects' use of energy evaluation tools is almost entirely limited to the EPB software, indicating that Flemish architects only evaluate the energy performance of the design for compliance with legal requirements.

Further, the size of the architectural firm appeared to be statistically significant. The EPB software is more often used by larger firms (6 associates or more), probably because larger firms often have a permanent availability of specialists for particular knowledge domains. For all other energy tools no statistically significant differences were found. Further analysis of the data indicated that there are also no statistically significant differences in the current use of energy tools for architects designing only residential buildings compared to architects also designing other project types such as public buildings, except for checklists, which are more often used by the latter.

3.3.2. Reasons for Not Using Energy Tools

As previously stated, 31% of the architectural practices of the respondents does not use energy tools. Figure 2 summarizes the reasons for this. The results are presented for both series of the survey separately due to statistically significant differences for two aspects, marked by a dotted rectangle in the figure.

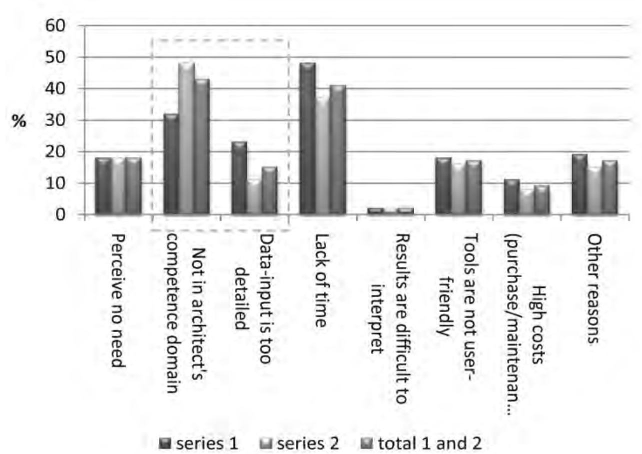


Figure 2. Reasons for not using energy evaluation tools

The major reasons for not using energy tools are the opinion that it is outside the architect's competence domain and lack of time. Consequently, the most frequently indicated reasons for not using tools are non tool specific. Detailed data-input and lack of user-friendliness appeared to be the most important tool specific reasons. Other cited reasons were mainly related to the outsourcing to specialists.

Architects of the second series seem to be more skeptical towards the usage of energy tools in general, since "the use of tools is outside the architect's competence domain" was an important reason for not using energy tools for almost 50%, compared to 32% in the first series. For the first series, lack of time seemed to be an important reason for not using energy tools for the majority of the respondents (48%). These architects also tend to be more concerned about tool-specific functions, compared to the respondents of the second series. Almost 25% of the respondents of the first series considered a too detailed data-input as an important reason for not using tools, compared to only 11% of the respondents of the second series. These observed differences might be explained by the fact that the architects of the first series, who subscribed first for the course, may be more interested in energy use.

3.3.3. Energy Evaluation in the Design Process

The results in figure 3 show a tendency of increasing use of energy tools during design development. Energy tools are mainly used in detailed design phases. Less than 50% of the respondents who use energy tools, use them in the conceptual design phase (CDP).

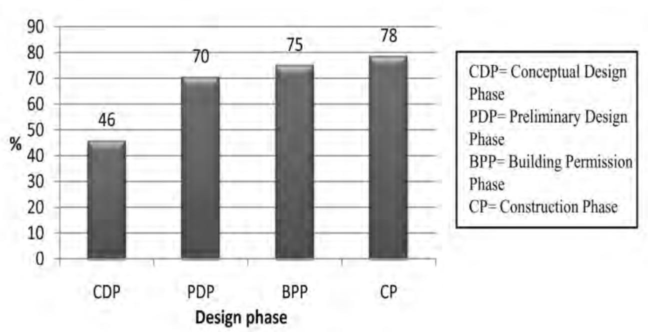


Figure 3. Use of energy tools in the design phases

When only considering the respondents who only use the EPB software, slightly different results are obtained. They use the EPB software considerably less in early design phases, with only 36% in the CDP and 62% in the preliminary design phase (PDP) (compared to 46% and 70% in figure 3). This might be explained by the fact that tools such as checklists may be easier to use in early design phases. Nevertheless, these results indicate that energy tools are rarely used to support early design decisions.

Figure 4 assesses the roles energy tools play in the DP. The results show that energy tools are mostly used to meet regulatory requirements in all design phases (over 70%), with a peak in the BPP (93%), but are much less used for all other roles. The response for all other roles is more or less evenly distributed among the design phases. Only small differences are present. Assessing the impact of design decisions is more important in early than in detailed design phases. Also, tools are used slightly less to fulfill the client's wishes during design development. Again, these results imply that energy tools are currently mainly used for code compliance, but rarely for design purposes.

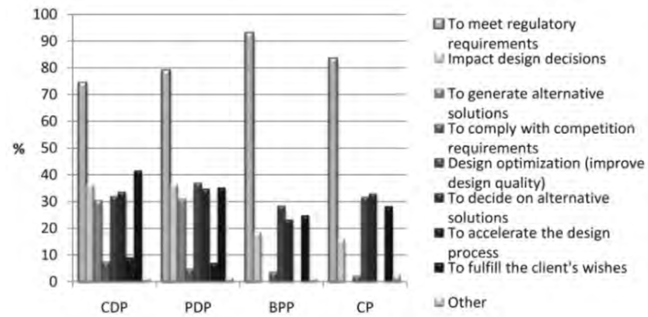


Figure 4. Reasons for using energy tools in the DP

3.3.4. Limitations of and Preferences for the EPB

59% of the respondents uses or has used the EPB software themselves, meaning that ca. 40% has never used it. The most frequently stated reason for not using the software is the outsourcing to specialists or to colleagues. Other reasons included among other lack of time, complexity of the program, lack of user-friendliness, and extra work.

There is also a clear tendency of decreasing use of the software with increasing age of the respondent (figure 5). 67% of the respondents younger than 31 uses or has used the software, compared to 58% of respondents between 31 and 40, and 54% and 53% of respondents between 41 and 50 years old, and older than 50 respectively. A possible explanation might be that younger architects are probably more motivated to learn new methods and to use energy tools. It is therefore likely that this tendency will continue in the future, indicating an increased uptake of energy tools.

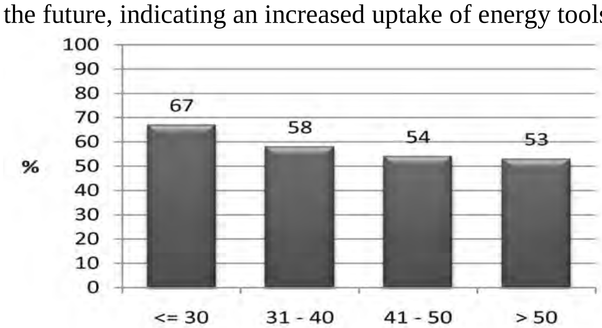


Figure 5. Use of the EPB software in relation to the respondents' age

The design supportive character of the EPB software was further examined among the respondents who use or have used the program. The results point out that 53% of these respondents believe that the software does not provide sufficient support when designing. Further analysis of the data revealed statistically significant differences between both series of the survey, as summarized in table 1.

	Series 1	Series 2
Sufficient support	39%	52%
Insufficient support	61%	48%

Table 1. Design supportive character of the EPB software. Series 1 versus series 2.

61% of the respondents of the first series who use or have used the EPB indicated that the software does not provide sufficient support during the DP, compared to only 48% of the respondents of the second series. This difference is possibly due to the fact that architects of the first series might be more interested in energy, since they subscribed first for the course "Energy Conscious Architect".

Subsequently, other experienced limitations of the EPB software were examined for the respondents who use or have used the software (figure 6). The major experienced limitations are clearly lack of an extensive component library (67%), too complex software to provide instant feedback (62%), and detailed data-input (61%). Around 30% of the respondents perceives lack of integration with CAD software as an important limitation, whereas only 13% considers the lack of support for design decisions as a limitation. Other limitations mentioned are lack of user-friendliness and of transparency. For 2% of the respondents the EPB software has no limitations.

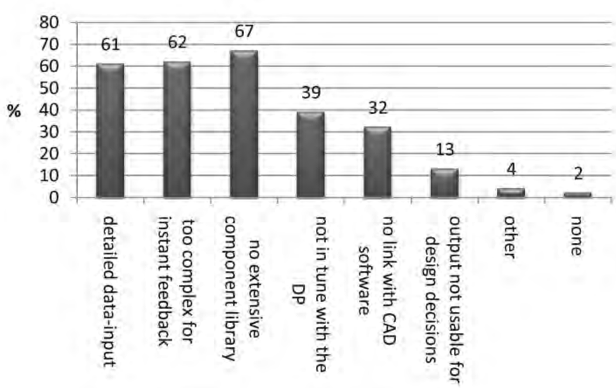


Figure 6. Limitations of the EPB software

Regarding the need for an energy design tool, facilitating intermediary design evaluations, if regulations become more severe in the future (E60), 88% of all respondents responded positive. Moreover, in this case 90% of the respondents currently not using energy tools in their architectural practice also expressed a need. These results indicate that if Flemish regulations become tightened, the need for a design supportive energy tool will be considerably high. Related to the size of the architectural firm, there was no statistically significant difference for this aspect, nor for the type of design projects. However, since 86% of the respondents focusing on residential projects only indicated a need, architects also need design support for the energy efficiency aspects of small projects.

When asked to provide an order of priority regarding the most important aspects, ease of use was ranked as the most important tool aspect. 'Provide guidelines for a specified E and K-level' and 'possibility for intermediary evaluations' had a shared second ranking. The next important tool aspect, ranked at the fourth place, was to provide feedback related to an expected E-level. The least important aspects were generation of design alternatives and a link with existing CAD software. However, further analysis of the data revealed that even the least important aspects (the lowest ranked criteria, i.e. link with CAD software and generation of design alternatives) were considered to be important by almost 60% of the respondents surveyed. Therefore, these aspects should also be taken into consideration when developing new or improving existing energy design tools.

3.3.5. Criteria for Energy Evaluation Tools

When considering the important criteria for energy evaluation tools in general (figure 7), ease of use is important for most of the respondents (90%), followed by conformance with standards and regulations (68%). Ease of learning, easy interpretation of the results, and price are all important tool criteria for about half of the respondents. For 31% of the respondents, a link with CAD software is important.

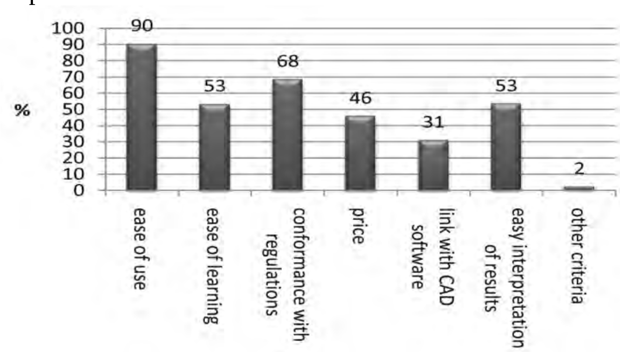


Figure 7. Criteria for energy evaluation tools

Considering specifically the aspects for which the respondents of the second series require feedback during the DP (figure 8), feedback on the impact of glazing percentage (80%) and insulation (76%) on energy use is clearly important to most architects, followed by feedback on the impact of glazing percentage on overheating (63%). The impact of thermal mass and orientation on energy use and the impact of shading devices on overheating risks are all of average importance, as indicated by about 50% of the respondents of the second series. Finally, around 40% of these respondents expressed a need for design feedback on the impact of thermal mass, orientation and insulation on overheating.

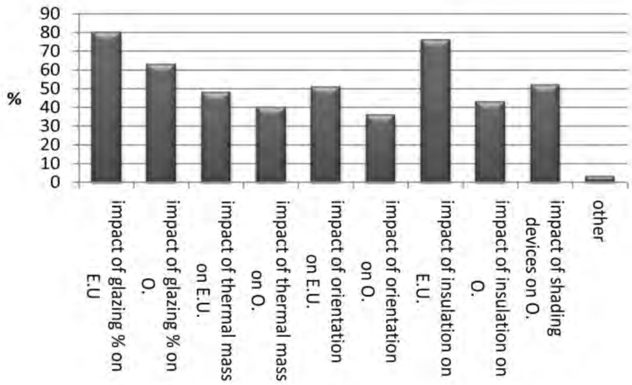


Figure 8. Feedback aspects for the respondents of the second series. (E.U.= energy use, O= overheating)

3.3.6. Factors for Design Decisions

23% of the respondents surveyed have realized a low-energy project with a maximum E-level of 60. When considering the factors influencing design decisions, 50% of these respondents usually decide on experience for the design of low-energy projects. Intuition is taken into account by 34% of the respondents. 44% makes intermediary EPB calculations and 38% appeals on specialists. Comparing these results to the factors that determine design decisions in general, it seems that architects rely less on experience for the design of energy efficient projects than for general design aspects (50% versus 85%). This further suggests that architects need more adequate support for the design of low-energy projects.

3.3.7. Current Use of 3D Modeling Software

57% of the respondents of the second series using 3D modeling software, uses Sketch-Up. All other programs (Revit, Archicad and Autocad) are more or less equally used (by around 10%), with the exception of Vector Works (17%).

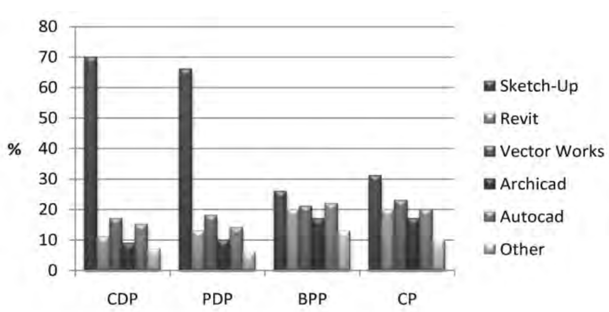


Figure 9. Most often used 3D tool in each design phase (CDP=conceptual design phase, PDP=preliminary design phase, BPP=building permission phase, CP=construction phase)

As figure 9 shows, Sketch-Up is clearly the most frequently used 3D tool in the CDP and the PDP among the respondents of the second series who indicated using 3D tools in these phases. During the BPP and the CP all 3D tools are more or less equally used. The clear domination of Sketch-Up in early design phases is important to keep in mind when considering the compatibility of BPS tools with existing CAD software.

A more detailed analysis showed that 3D modeling software is more widely used among younger architects (younger than 50). It is therefore likely that the usage of 3D tools will increase in the future.

3.4. Results: Open Question

The respondents of the second series of the survey were also asked to explain the user-friendliness of energy tools and their compatibility with the DP in an open question. 251 respondents had completed this question.

The data revealed that tool simplicity, intuitive tool usage, and time to use the tool are the most important aspects determining the user-friendliness of energy tools. Again, 'Sketch-Up' was sometimes referred to as example. Further, many respondents considered easy, fast, and minimum data-input, an extensive component database, a clearly structured interface with a restrained set of options, ease of learning, and the ability to quickly create, test and compare alternative designs to be important aspects for the user-friendliness of energy tools. Other frequently discussed aspects are: ease of interpretation of the results, a link with CAD software, and ease of data review. Finally, other aspects mentioned are: adaptability of library components, conformance with regulations, visual qualities of the tool, reliability of the output, providing impact degrees of parameters, transparency of the tool, input and output in the language of the architect, providing guidelines, and input from general to detail. User-friendly tools should also enhance the communication with clients. Considering the usability of energy tools in the DP, real-time feedback on design changes, limited data-input, speed of working, and ease of data review were frequently identified criteria. Other issues included among other: input from general to detail, providing guidelines, and usability in early design phases.

4. DISCUSSION

This paper discusses the results of interviews with 9 and a survey among 629 architects in Flanders, Belgium. Despite differences between architects' populations in other countries these results may provide valuable information for tool developers, as they might benefit from multiple views.

The results showed that the current usage of energy tools in Flanders is mainly limited to the EPB software. Other tools, in particular BPS, are rarely used. This might partially be explained by the fact that Flemish architects are often not aware of the existence of other tools, as turned out from the interviews. The current uptake is primarily focused on fulfilling legal requirements, but rarely on addressing deeper design issues, and is mainly concentrated in detailed design phases. 31% of the respondents currently does not use any energy tool in the architectural practice. Similar findings were observed in previous studies [8,9].

Given the importance of integrating BPS in the DP and the possible increased uptake of energy tools by architects in the future (as indicated by the survey), it is of major concern to thoroughly take into account the needs of this (growing) user group. By combining multiple sources of information and providing the respondents the opportunity to elaborate on their views, this study yielded a comprehensive understanding about the concept "architect-friendly" in a Belgian context. Moreover, when comparing these results with previous studies in different geographic locations a number of similar findings emerge: a) detailed data-input is one of the major limitations of current BPS tools [8,9], b) the tools do not match the existing DP [8], c) important tool criteria include among other easy interpretable output, easy data-input, user-friendliness, better ease of learning [8,9], graphical input and output, simple navigation, use of default values, extensive component library, and ease of creating and comparing alternatives [12]. While the study of Lam et al. [8] highlighted the importance of an integration with CAD software, the study of Mahdavi et al. [9] showed more divided results. In the current study, the respondents' opinions were also more divided. The interviews revealed that this largely depends on the architect's preferred design media. Further, conformance with regulations was identified as a very important tool criterion. This was also observed in other studies [9,12]. Regarding the usability of tools in the DP, Attia et al. [12] stressed the importance of a quick energy analysis and the ability to examine the sensitivity of key design parameters. Similar findings were done in the current study. In addition to this, both the interviews and the survey stressed the importance of intuitive tool usage and limited time to operate the tool. This implicates that tools should easily be employable in other design media and activities, which might be implemented by introducing energy evaluation in 3D modeling software. For early design, Sketch-Up seems most appropriate considering its common usage in the early design stages. Sketch-up's popularity also applies to other countries [12].

Based on the results of the interviews and the survey a conceptual framework is synthesized, defining the concept of "architect-friendliness" of BPS tools (figure 10). The data revealed that the criteria considered by the respondents and contributing to the concept of "architect-friendliness" could be classified into four major themes, namely data-input, data-output, graphical user interface (GUI), and usability in the DP. An extra general theme is included, to incorporate important issues that could not directly be assigned to one of the other themes.

BPS tool	1 2 3 4 5
DATA-INPUT	
Limited data-input	
Quick data-input (time to create model is less than 1h)	
Input in the language of the architect	
Use of defaults to limit and facilitate data-entry	
simple and intuitive input process	
easy data review/change	
easy create alternative designs/options	
extensive library/database of building components	
input consistent with early design phase (basic input)	
from general to detail	
graphical representation of building geometry:	
3D modeler in simulation tool	
Possibility to import 3D CAD files	
Possibility to import from Sketch-Up	
Input via drawing software (for instance Sketch-Up)	
OUTPUT	
Easy interpretation (language of architects)	
graphical representation of output	
conformance with building codes and regulations	
impact of decisions/parameters (uncertainty/sensitivity)	
Simple but supportive information for design decisions	
Convincing output, to communicate with clients	
clearly indicate problem area(s)	
benchmarking	
output displayed in 3D building model	
generate reports for alternative designs/options	
reliability of the output	
adapted for different design phases	
INTERFACE	
visual communication of GUI	
clear, intuitive, and flexible navigation	
clearly structured with a restrained set of functions (simplicity)	
USABILITY in DP	
minimally interrupt the DP	
Data-input in tune with DP	
simplicity	
minimal time required to operate the tool	
adapted for use in early design	
quickly obtain solutions	
quickly and easily create, test and compare alternatives	
real-time feedback on design decisions and changes	
provide guidelines	
GENERAL	
adaptable default values	
highly visual	
transparency of the tool	
ease and intuitive in use	
short calculation time	
easy to learn	
adequate for local usage (units/materials/…)	
easy to use after long time of non-use	

Figure 10. Architect-friendliness: Conceptual framework.

Overall, similar findings were done between this study and studies in other countries, despite the different contexts. Certain aspects, such as integrating case study databases [12], were not observed in this study. Conversely, several issues such as clear output to facilitate communication with clients, though important in this study, were not addressed in the others. This may partially be due to the fact that the current study combined a survey with interviews and that the respondents could also actively contribute aspects they considered important. Several of these issues have also repeatedly been recognized by other researchers. Hence, the proposed checklist might possibly be generalizable to other regions, and might be a good base to better evaluate "architect-friendliness" and to develop future tools that fit the architect's expectations.

The scheme comprises two columns. The left column includes the five themes with corresponding criteria. The right column provides space to evaluate a particular tool. The assessment is based on a five point scale, 1 being "not present" and 5 representing "extremely good". The scheme primarily focuses on energy simulation tools, but its usage might be extended to BPS tools in general. The first section of the scheme includes important aspects to adjust the input to an architect user, such as limited and quick data-input. Considering the graphical representation of the geometry, four items were incorporated, of which 'input via modeling software' is considered to be the most architect-friendly, as it minimally interrupts the DP. In the output section of the scheme, the item "benchmarking" was added, though not explicitly derived from the study. Summarized, the results stressed the importance of easy interpretable output. By comparing the output of a project with the performance of a well-known example, such as a low-energy house, architects get a better understanding of the consequences of their results. The third section stresses the importance of a clear GUI. Considering the usability in the DP, aspects such as simplicity and easily create alternatives are important. The last section includes a number of general criteria.

5. CONCLUSION

This paper analyzed the user-friendliness of energy evaluation tools from an architect's point of view via interviews and a survey. The data showed that these tools might play a potential role in design support for low energy projects, also for small projects. Addressing the needs of the architectural design community is crucial to encourage the uptake of BPS tools by architects. Future research will focus on evaluating existing BPS tools according to the proposed scheme, and on ways to concretize the results. The latter will be done organizing focus groups with architects.

6. ACKNOWLEDGEMENTS

The authors would like to thank the architects for their participation in this research, and Erik Nuyts for his statistical support. Research funded by a Ph.D. grant of the Institute for the Promotion of Innovation through Science and Technology in Flanders (IWT-Vlaanderen).

7. REFERENCES

[1] Hobbs, D.; C. Morbitzer; B. Spires; P. Strachan; J. Webster. 2003. "Experience of Using Building Simulation within the Design Process of an Architectural Practice." In Proceedings of the 8th International IBPSA Conference. Eindhoven.

[2] Tuohy, P.G.; C. Johnstone; L. McElroy. 2006. "Sustainable Housing Standards: Use of Simulation in Design." In Proc. of Eurosun, Glasgow.

[3] Hensen, J.L.M. 2004. "Towards More Effective Use of Building Performance Simulation in Design." In Proceedings of the 7th International Conference on Design and Decision Support Systems in Architecture and Urban Planning. Eindhoven.

[4] Bleil de Souza, C., and I. Knight. 2007. "Thermal Performance Simulation from an Architectural Design Viewpoint." In Proceedings of the 10th International IBPSA Conference. Bejing, China. http://www.ibpsa.org

[5] Schlueter, A., and F. Thesseling. 2009. "Building Information Model Based Energy/exergy Performance Assessment in Early Design Stages." Automation in Construction, 18, no.2: 153-163.

[6] Riether, G., and T. Butler. 2008. "Simulation Space. A New Design Environment for Architects." In 26th eCAADe Conference Proceedings. Muylle M., ed., Antwerp, Belgium.

[7] Bambardekar, S., and U. Poerschke. 2009. "The Architect as a Performer of Energy Simulation in the Early Design Stage." In Proceedings of the 11th International IBPSA Conference. Glasgow, Scotland.

[8] Lam, K.P.; N.H. Wong; H. Feriady. 1999. "A Study of the Use of Performance-based Simulation Tools for Building Design and Evaluation in Singapore." Proceedings of the 6th IBPSA Conference. Kyoto, Japan.

[9] Mahdavi, A.; S. Feurer; A. Redlein; G. Suter. 2003. "An Inquiry into the Building Performance Simulation Tools Usage by Architects in Austria." Proceedings of the 8th International IBPSA Conference. Augenbroe and Hensen, eds. Eindhoven, The Netherlands. 777-784.

[10]Pedrini, A., and S. Szokolay. 2005. "The Architects Approach to the Project of Energy Efficient Office Buildings in Warm Climate and the Importance of Design Methods. In Proc. of the 9th International IBPSA Conference. Montreal, Canada. http://www.ibpsa.org

[11]De Wilde, P.; G. Augenbroe; M. Van der Voorden. 1999. "Invocation of Building Simulation Tools in Building Design Practice." In Proc. of the 6th International IBPSA Conference '99. Kyoto. http://www.ibpsa.org

[12]Attia, S.; L. Beltrán; A. De Herde; J. Hensen. 2009. "Architect-friendly": a Comparison of Ten Different Building Performance Simulation Tools. In Proc. of 11th International IBPSA Conference. Glasgow. Scotland.

[13]Ellis, M.W., and E.H. Mathews. 2001. "A New Simplified Thermal Design Tool for Architects." Building and Environment 36, no. 9. 1009-1021.

[14]Morbitzer, C. 2003. "Towards the Integration of Simulation into the Building Design Process." Ph.D. Thesis. University of Strathclyde.

[15]http://www.epbd-ca.org

Schedule-Calibrated Occupant Behavior Simulation

Rhys Goldstein, Alex Tessier, Azam Khan
Autodesk Research
210 King St. East, Toronto, Ontario, Canada, M5A 1J7
Azam.Khan@autodesk.com

Keywords: Building performance simulation, energy demand prediction, occupant behavior, model calibration

Abstract

Building performance simulation promises to reduce the future impact of buildings on the environment by helping architects predict the energy demand associated with different design options. We present a new method for simulating occupant behavior in buildings, a key phase in the prediction of energy use. Our method first inputs the recorded activities of actual building occupants, then randomly generates fictional schedules with similar behavioral patterns. The main contribution of this work is a mathematical technique in which an arbitrary set of factors can be used to select plausible activity types, durations, and numbers of participants during a simulation. A prototype model was implemented to test the method, and results obtained to date suggest that the generated occupant schedules are believable when compared both qualitatively and quantitatively to real occupant schedules.

1 INTRODUCTION

It has been estimated that residential and commercial buildings account for 35 percent of global energy demand [1], as well as a substantial fraction of greenhouse gas emissions. At the same time, it is widely believed that design improvements could dramatically reduce the impact of buildings on the environment. Clarke identifies "ineffective decision-support" as the key factor impeding the possible realization of 50-75% reductions in the energy consumption of new buildings, and 30% reductions in that of existing buildings [2]. Our vision is that advances in building performance simulation will address the need for improved decision-support in building design, allowing architects to conveniently and accurately predict the energy use associated with different design options.

One of the more daunting aspects of building performance simulation is the number of different interacting subsystems that need to be modeled. These subsystems include the equipment in the building, the HVAC system, and the outdoor environment, among others [3]. Because a building's energy consumption patterns are largely dependent on the activities of its occupants, we chose to start our investigation of energy demand prediction by looking at occupant behavior. Quantifying this behavior is a prerequisite for predicting when a building's equipment is likely to be in use, and assessing the adequacy of its lighting conditions, air temperature, and air quality. As part of a long-term collaborative modeling project [4], our vision is to integrate realistic models of occupant behavior with those of other subsystems.

Existing building performance simulation tools typically use fixed schedules or relatively simple algorithms for modeling occupant behavior. In pursuit of more detailed and accurate predictions, researchers have proposed more sophisticated methods. It has been suggested that the trend towards flexible work hours will complicate occupant schedules and compound the need for these new methods [5]. Also, the emerging focus on sustainable buildings, incorporating passive air conditioning systems, will likely increase an occupant's influence over his or her surrounding indoor environment [6]. Simulations will need to capture these occupant-building interactions.

Here we present a novel occupant behavior simulation method. We describe our method as "schedule-calibrated", meaning that it first inputs the recorded activities of actual building occupants, then generates fictional schedules while striving to reproduce typical patterns of behavior. Several similar schedule-calibrated methods already exist. We contribute a mathematical technique that can be used to generate various activity "attributes". The specific attributes we focus on in this paper are the task performed, the number of participants, and the duration of each activity. Unlike pre-existing methods, ours allows each generated attribute to depend on an arbitrary set of "factors". Examples of factors include the time of day, the previously-performed task, and the time elapsed since each task was last performed. Any activity attribute may also be used as a factor. The generation of various attributes accommodates more detailed descriptions of human behavior, whereas the use of multiple factors promises to help reproduce behavioral patterns found in existing data.

Section 2 describes key concepts and reviews related work. Our proposed occupant behavior simulation method is explained in Section 3 using a simple example demonstrating the random selection of plausible tasks. In Section 4 we describe the actual prototype model. It was developed using the method of Section 3, but generates multiple attributes, uses a greater number factors, and discretizes those factors at higher resolutions. The prototype model was implemented, and simulation results are presented in Section 5. Finally, Section 6 discusses future work.

2 BACKGROUND

We use the phrase "occupant behavior simulation" to refer to a computer simulation that generates fictional occupant schedules. An "occupant schedule" is a description of the behavior of a building occupant over the course of a single day. Each of these schedules takes the form of a chronological sequence of consecutive activities, with an "activity" being a description of an occupant's behavior during a specific block of time. Each activity has several associated attributes. At very least, these attributes should include the "task", which identifies the type of activity. We expect each task to be selected from a pre-defined list of possible tasks. For a simulation of an office building, there would likely be one possible task representing desk work, another possible task for meetings, etc.

Several methods have been proposed to randomly generate plausible sequences of periods during which an occupant is either present or absent at a particular location in a building. While our interest lies in more detailed models of human behavior, these methods still satisfy our definition of "occupant behavior simulation" if the list of possible tasks is to include only "being present" and "being absent". Wang proposed that the durations of presence and absence be exponentially distributed [7], a convention which assumes that the remaining time to be spent in a location is independent of the time already spent there. Both Yamaguichi's method [8] and Page's method [9] also feature this "memoryless" property, but they differ from Wang's in that time is advanced by fixed time steps. Page introduced what we describe as a single influencing "factor"; specifically, the time of day. With this method, a simulation is calibrated using real schedules of presence and absence. If the real schedules tend to include a lunch break around noon, then the time of day factor allows that pattern of behavior to be reproduced.

Likely the most sophisticated occupant activity simulation developed to date is Tabak's User Simulation of Space Utilisation (USSU) System, described in his 2008 Ph.D thesis [10]. In USSU there are many different tasks, and occupants can interact via shared activities such as meetings and presentations. Unlike Page's method, USSU is not schedule-calibrated. Tabak instead conducted an extensive survey, using questionnaire results to calibrate his model.

The method developed for USSU [10] appears to have a few disadvantages. One possible concern is that a model calibrated using survey data might produce less realistic results than one calibrated using recorded schedules. The collection of the survey data itself is likely to be cumbersome (Tabak reported that 50 of 166 respondents failed to complete the questionnaire, presumably due to its length). That said, obtaining real occupant schedules for over 100 people would almost certainly be even more difficult. The drawback that concerns us most is the complexity of the USSU method. Based on the task performed, Tabak classifies activities as "Skeleton Activities" (eg. "give a presentation" or "do research"), "S-curve Intermediate Activities" (eg. "get a drink"), and "Probabilistic Intermediate Activities" (eg. "receive unexpected visitor"). Each type of activity is handled by a different algorithm, complicating the overall method.

3 PROPOSED METHOD

It seems plausible that the large amount of input data required by a schedule-calibrated method like Page's allows one to get away with relatively simple algorithms. Because the information contained in the real occupant schedules implicitly distinguishes one task from another, one can avoid explicit distinctions like those exhibited by USSU. But while Page's schedule-calibrated method is compelling for its simplicity, we decided at the outset of our work to avoid the memoryless property. Intuitively, an occupant's future behavior should be influenced in part by his or her past behavior. If one has taken a lunch break only three minutes ago, as opposed to three hours ago, one ought to be less likely to have a meal in the next three minutes. We therefore adopted the goal of developing a schedule-calibrated method like Page's, but with more detailed activities and an enhanced ability to reproduce observed patterns of behavior.

Recall that activities include attributes like the task, number of participants, and duration. In our proposed occupant behavior simulation method, attributes are generated one at a time using the same mathematical technique. Each attribute depends on an arbitrary set of factors, which for the time being must be chosen based on intuition. Intuitively, an occupant's next task should depend in part on the time of day, so like Page we might choose the time of day as a factor. Our method is novel in that it allows for the inclusion of other factors, the previous task being a notable example. Once one attribute is generated from a set of factors, that attribute can itself become a factor for the generation of another attribute. If one uses the time of day to generate the task, for example, one may then use the task as a factor that influences the activity duration.

Comparing our method to Page's and its predecessors, it is through the use of multiple factors that we expect an enhanced ability to reproduce behavioral patterns. Also, the generation of various activity attributes leads to a more detailed representation of occupant behavior. Striving to retain the simplicity of pre-existing schedule-calibrated methods, we generate each attribute using the same mathematical technique. This technique can be broken into four distinct phases: the population of a set of histograms using real occupant schedules; the smoothing of those histograms; the normalization of the smoothed histograms; and finally the extraction of attribute values. Each phase is explained in detail below.

TOD \ PVT	$A_0[...]$				$B_0[...]$["off"]				$B_0[...]$["work"]				$B_0[...]$["meet"]				$B_0[...]$["eat"]			
	off	work	meet	eat	off	work	meet	eat	off	work	meet	eat	off	work	meet	eat	off	work	meet	eat
7 AM																				
8 AM	1								1											
9 AM	1								1											
10 AM		2												2						
11 AM			2								2									
12 PM		1																1		
1 PM		1		2								2						1		
2 PM																				
3 PM																				
4 PM		2	1			1					1			1						
5 PM		1				1														
6 PM																				

Figure 1. Histograms after the shaded bins were populated with the Table 1 data (empty bins represent values of 0).

3.1 Histogram Population

Throughout Section 3, an ongoing example will be used in which we generate a single activity attribute: the task. To simplify matters, we will have only four possible tasks, named "off", "work", "meet", and "eat". We will also use only two factors, the time of day (TOD) and the previous task (PVT).

Observe the two real occupant schedules in Table 1. Each row describes a separate activity. An activity begins at its associated time, and ends at the time when the subsequent activity begins. The "NPO" column lists the "number of participating occupants". On Day 0, for example, the occupant had a meeting with four other occupants at 10:02 AM.

Table 1. Two real occupant schedules.

Day 0			Day 1		
Time	**Task**	**NPO**	**Time**	**Task**	**NPO**
8:45 AM	work	1	9:13 AM	work	1
10:02 AM	meet	5	10:41 AM	meet	2
11:17 AM	work	1	11:02 AM	work	1
12:10 PM	eat	3	1:16 PM	eat	1
1:01 PM	work	1	1:55 PM	work	1
5:47 PM	off	1	4:32 PM	meet	2
			4:51 PM	work	1
			4:59 PM	off	1

Our goal is to automatically generate tasks during a simulation, producing schedules that resemble those in the table. As mentioned earlier, the first step is to populate a set of histograms. Figure 1 shows five such histograms after they are populated using the data in Table 1. The histograms are all two-dimensional, as we have selected two factors for this example. There is one column for each of the four possible previous tasks, and the time of day factor is discretized such that each hour of the day has its own set of histogram bins.

Every activity in Table 1 contributes a value of 1 to both A_0 and one of the other four B_0 histograms. Each B_0 histogram is associated with a single "feature", and in this specific example we happen to have one feature per possible task. The general mathematical technique we present allows one to choose any number of features, however, and each feature can be associated with any quantity. For the first activity of Day 0, the time (8:45 AM) fits into the "8 AM to 9 AM" slot and the previous task is assumed to be "off". Thus $A_0[\text{“8 AM”}, \text{“off”}] = 1$, as shown just under the top-left corner cell of Figure 1. Because the task performed at 8:45 AM is "work", we also add 1 to the corresponding bin of the "work"-specific feature histogram ($B_0[\text{“8 AM”}, \text{“off”}][\text{“}\textit{work}\text{”}]$). It so happens that on both Day 0 and Day 1, the occupant transitions from the "work" task to the "meet" task between 10 AM and 11 AM. We therefore have $A_0[\text{“10 AM”}, \text{“work”}] = B_0[\text{“10 AM”}, \text{“work”}][\text{“meet”}] = 2$. On Day 1, the occupant transitioned away from the "work" task at 4:32 PM and 4:59 PM, giving us $A_0[\text{“4 PM”}, \text{“work”}] = 2$. Note that the 4:32 PM activity contributes to $B_0[\text{“4 PM”}, \text{“work”}][\text{“meet”}]$, whereas the 4:59 PM activity affects $B_0[\text{“4 PM”}, \text{“work”}][\text{“off”}]$.

3.2 Smoothing

The sparseness of Figure 1 is a problem, as we will eventually need to extract information from the bins that are currently empty. This problem will occur in practice, even if hundreds of real occupant schedules are used to populate the histograms, for the use of additional factors discretized at

TOD \ PVT	A₁[...]				B₁[...]["off"]				B₁[...]["work"]				B₁[...]["meet"]				B₁[...]["eat"]			
	off	work	meet	eat	off	work	meet	eat	off	work	meet	eat	off	work	meet	eat	off	work	meet	eat
7 AM	0.07								0.07											
8 AM	0.98	0.02	0.02	0.02					0.98	0.02	0.02	0.02								
9 AM	0.98	0.16	0.02	0.02					0.98	0.02	0.02	0.02		0.14						
10 AM	0.11	1.95	0.18	0.04					0.07		0.14		0.04	1.95	0.04	0.04				
11 AM	0.04	0.25	1.95	0.04					0.04	0.04	1.95	0.04		0.14				0.07		
12 PM	0.02	0.98	0.16	0.16							0.14	0.14					0.02	0.98	0.02	0.02
1 PM	0.06	0.99	0.06	1.95					0.04	0.01	0.04	1.95					0.02	0.98	0.02	0.00
2 PM		0.07		0.14								0.14						0.07		
3 PM		0.14	0.07			0.07					0.07			0.07						
4 PM	0.06	1.96	0.97	0.06	0.02	0.98	0.00	0.02	0.02	0.00	0.97	0.02	0.02	0.98	0.00	0.02				
5 PM	0.02	0.99	0.09	0.02	0.02	0.98	0.02	0.02			0.07			0.01						
6 PM		0.07				0.07														

Figure 2. Histograms after 1 smoothing iteration, with initially-populated bins shaded ("0.00" values are small but positive).

higher resolutions will increase the total number of bins. It is therefore necessary to "smooth" the data, propagating values across neighboring bins to reduce the sparseness.

Here we describe an iterative smoothing algorithm requiring one "smoothing parameter" $\alpha_{\langle factor\rangle}$ for each factor $\langle factor\rangle$. These smoothing parameters are all non-negative, and their sum is at most 1. The larger a smoothing parameter for a certain factor, the smoother the final data across neighboring bins along the axis associated with that factor.

The first step in each smoothing iteration i is to define a set of coefficients C_i, one for each histogram bin (identified by $\langle bin\rangle$) and each factor.

$$C_i[\langle bin\rangle][\langle factor\rangle] = \frac{\frac{\alpha_{\langle factor\rangle}}{n_{\langle factor\rangle}} \cdot \sum_{\langle factor\rangle} \alpha_{\langle factor\rangle}}{\sum_{\langle factor\rangle} \alpha_{\langle factor\rangle} + \sqrt{A_i[\langle bin\rangle]}}$$

If a bin has been heavily populated, then trusting its information, we lessen the effect of the smoothing by including $\sqrt{A_i[\langle bin\rangle]}$ in the denominator. The variable $n_{\langle factor\rangle}$ is the number of neighboring bins along the axis associated with $\langle factor\rangle$. For factors with continuous values like the time of day, the neighboring bins are the two adjacent bins (or the one adjacent bin if we are at the edge of a histogram). For discrete factors like tasks, all other bins along the axis are neighbors.

With c_i we record the fraction of each bin's value that will be preserved through the smoothing iteration.

$$c_i[\langle bin\rangle] = 1 - \sum_{\langle factor\rangle} \left(n_{\langle factor\rangle} \cdot C_i[\langle bin\rangle][\langle factor\rangle]\right)$$

In a single iteration, convolutions are performed separately for each histogram using the same set of coefficients. Below, X_i represents any of the histograms for iteration i. We use $\langle bin^*\rangle$ to denote the neighboring bin located at a displacement $\langle offset\rangle$ along the axis associated with $\langle factor\rangle$.

$$X_{i+1}[\langle bin\rangle] = c_i[\langle bin\rangle] \cdot X_i[\langle bin\rangle] + \sum_{\langle factor\rangle} \sum_{\langle offset\rangle} (C_i[\langle bin\rangle][\langle factor\rangle] \cdot X_i[\langle bin^*\rangle])$$

Continuing our simplified task generation example, we let $\alpha_{TOD} = 0.14$ and $\alpha_{PVT} = 0.06$. Figure 2 shows the histograms of Figure 1 after one iteration of smoothing. Serving as a demonstration of the algorithm, the following is a derivation of A_1["4 PM", "work"]. Note that its value is shown as 1.96 in the figure.

First, we calculate the smoothing coefficient associated with the time of day factor. Because this factor is continuous, there are two neighbors ($n_{TOD} = 2$).

$$C_0[\text{“4 PM”}, \text{“work”}][\text{“TOD”}] = \frac{\alpha_{\text{TOD}}}{n_{\text{TOD}}} \cdot \frac{\alpha_{\text{TOD}} + \alpha_{\text{PVT}}}{\alpha_{\text{TOD}} + \alpha_{\text{PVT}} + \sqrt{A_0[\text{“4 PM”}, \text{“work”}]}} = \frac{0.14}{2} \cdot \frac{0.14 + 0.06}{0.14 + 0.06 + \sqrt{2}} = 0.0086730\ldots$$

The smoothing coefficient associated with the previous task is calculated in a similar fashion. In this case, because tasks are discrete and there are four in total, there are three neighbors ($n_{PVT} = 3$).

$$C_0[\text{“4 PM”}, \text{“work”}][\text{“PVT”}] = 0.0024780\ldots$$

PVT / TOD	D[...]["off"]				D[...]["work"]				D[...]["meet"]				D[...]["eat"]			
	off	work	meet	eat	off	work	meet	eat	off	work	meet	eat	off	work	meet	eat
7 AM	0.00	0.00	0.00	0.00	1.00	0.81	0.98	0.98	0.00	0.19	0.02	0.02	0.00	0.00	0.00	0.00
8 AM	0.00	0.00	0.00	0.00	1.00	0.62	0.96	0.95	0.00	0.38	0.04	0.05	0.00	0.00	0.00	0.00
9 AM	0.00	0.00	0.00	0.00	0.98	0.15	0.81	0.72	0.02	0.85	0.19	0.28	0.00	0.00	0.00	0.00
10 AM	0.00	0.00	0.00	0.00	0.65	0.01	0.77	0.27	0.35	0.99	0.23	0.72	0.00	0.00	0.00	0.01
11 AM	0.00	**0.00**	0.00	0.00	0.79	**0.17**	0.99	0.81	0.13	**0.55**	0.01	0.13	0.08	**0.28**	0.00	0.06
12 PM	0.00	0.00	0.00	0.00	0.45	0.02	0.86	0.87	0.02	0.02	0.00	0.00	0.54	0.95	0.14	0.13
1 PM	0.00	0.00	0.00	0.00	0.64	0.06	0.71	0.99	0.00	0.00	0.00	0.00	0.35	0.94	0.29	0.01
2 PM	0.04	0.08	0.02	0.00	0.62	0.08	0.76	0.97	0.04	0.08	0.02	0.00	0.30	0.77	0.21	0.02
3 PM	0.30	0.47	0.04	0.17	0.38	0.02	0.90	0.66	0.28	0.46	0.04	0.16	0.04	0.05	0.01	0.01
4 PM	0.37	0.51	0.04	0.36	0.31	0.01	0.93	0.33	0.32	0.48	0.03	0.31	0.00	0.00	0.00	0.00
5 PM	0.74	0.90	0.25	0.74	0.15	0.00	0.73	0.15	0.11	0.09	0.02	0.11	0.00	0.00	0.00	0.00
6 PM	0.85	0.96	0.39	0.84	0.10	0.01	0.59	0.10	0.05	0.04	0.01	0.05	0.00	0.00	0.00	0.00

Figure 3. Normalized arrays, with initially-populated bins shaded (values shown in bold are referenced in the text).

The c_0 coefficient is obtained as follows.

$$\begin{aligned} c_0[\text{“4 PM”},\text{“work”}] \; = \; & 1 \\ & - n_{\text{TOD}} \cdot C_0[\text{“4 PM”},\text{“work”}][\text{“TOD”}] \\ & - n_{\text{PVT}} \cdot C_0[\text{“4 PM”},\text{“work”}][\text{“PVT”}] \\ = \; & 0.97522\ldots \end{aligned}$$

Using A_1 in place of X_{i+1}, the convolution equation gives us the value shown in Figure 1.

$$\begin{aligned} & A_1[\text{“4 PM”},\text{“work”}] \\ & = c_0[\text{“4 PM”},\text{“work”}] \cdot A_0[\text{“4 PM”},\text{“work”}] \\ & \quad + C_0[\text{“4 PM”},\text{“work”}][\text{“TOD”}] \cdot A_0[\text{“3 PM”},\text{“work”}] \\ & \quad + C_0[\text{“4 PM”},\text{“work”}][\text{“TOD”}] \cdot A_0[\text{“5 PM”},\text{“work”}] \\ & \quad + C_0[\text{“4 PM”},\text{“work”}][\text{“PVT”}] \cdot A_0[\text{“4 PM”},\text{“off”}] \\ & \quad + C_0[\text{“4 PM”},\text{“work”}][\text{“PVT”}] \cdot A_0[\text{“4 PM”},\text{“meet”}] \\ & \quad + C_0[\text{“4 PM”},\text{“work”}][\text{“PVT”}] \cdot A_0[\text{“4 PM”},\text{“eat”}] \\ & = 1.9616\ldots \end{aligned}$$

We terminate the smoothing processes after a pre-defined number of iterations n. If empty bins still remain, one can use a greater number of smoothing iterations, select lower resolutions on the continuous factors, or supply more input data.

3.3 Normalization and Extraction

After smoothing the histograms, the feature values B_n are normalized to yield a set of arrays collectively named D. This is the last step in the calibration process. In our ongoing example, normalization means dividing the B_n values by corresponding A_n values.

$$D[\langle bin\rangle][\langle feature\rangle] \; = \; \frac{B_n[\langle bin\rangle][\langle feature\rangle]}{A_n[\langle bin\rangle]}$$

Figure 3 shows the normalized arrays of the example after 9 iterations of smoothing. This information may be used during a simulation to generate plausible tasks, a process called extraction. Suppose that the current simulated time is 11:30 AM, and a simulated occupant has just completed a “work” activity. To select the next task, the simulation looks up the normalized feature values associated with these TOD and PVT factors.

$$\begin{aligned} D[\text{“11 AM”},\text{“work”}][\text{“off”}] \; &= \; 0.00\ldots \\ D[\text{“11 AM”},\text{“work”}][\text{“work”}] \; &= \; 0.17\ldots \\ D[\text{“11 AM”},\text{“work”}][\text{“meet”}] \; &= \; 0.55\ldots \\ D[\text{“11 AM”},\text{“work”}][\text{“eat”}] \; &= \; 0.28\ldots \end{aligned}$$

In this case there is roughly a 17% chance that the occupant will continue working, a 55% chance he/she will meet with other occupants, a 28% chance of taking a break for food, and very little chance he/she will leave for the day. The simulation selects the next task at random according to these probabilities.

Suppose that the occupant completes the “work” activity at 11:45 AM instead of 11:30 AM. In this case, if we choose, we may interpolate probabilities. The center of the “11 AM to 12 PM” slot is 11:30 AM, a quarter of an hour before the simulated time, and the center of the “12 PM to 1 PM” slot is 12:30 PM, three quarters of an hour after the simulated time. Using a linear interpolation, the probability that the occupant begins eating is $(3/4) \cdot 0.28 + (1/4) \cdot 0.95$, or 44.75%. The other probabilities would be interpolated in a similar fashion.

4 PROTOTYPE MODEL

Different models of occupant behavior may be defined using the method of Section 3, but with alternative sets of attributes, factors, and features. In a prototype model that we implemented to test the method, we generate for each activity its task, number of participating occupants (NPO), and duration. The same smoothing algorithm is applied in each case; our implementation uses multi-dimensional arrays to support an arbitrary number of factors. The histogram population, normalization, and extraction procedures differ somewhat between the three generated attributes, as explained below.

4.1 Task Generation

In our prototype, we generated tasks using three factors. Two of these, the time of day (TOD) and the previous task (PVT), were used in the simplified example of Section 3. The third factor is the "task suspension interval" (TSI).

The rationale for using TOD as a factor is that the timing of arrivals, departures, and lunch breaks is highly dependent on the time of day. We discretize TOD into 15-minute intervals, giving us a higher resolution than the 60-minute intervals used in the previous section. We use $\alpha_{TOD} = 0.23$.

The rationale for the PVT factor is that certain transitions between tasks are more likely to occur than others. We in fact chose $\alpha_{PVT} = 0$, which prevents a transition from occurring in the simulated schedules if it is not found in the real schedules. With $\alpha_{PVT} = 0$, one would not expect two washroom breaks to be generated back-to-back.

The TSI measures the time elapsed since a task was last performed. Its inclusion as a factor helps reproduce intervals between tasks. This is useful to spread out washroom breaks, or coffee/lunch breaks, for example. We adopted a logarithmic discretization, with separate bins for task suspension intervals of 0 to 15 minutes, 15 to $15{\cdot}\sqrt{2}$ minutes, $15{\cdot}\sqrt{2}$ to 30 minutes, 30 to $30{\cdot}\sqrt{2}$ minutes, etc. We selected $\alpha_{TSI} = 0.02$.

Whereas the TOD and PVT factors are scalar values, the TSI factor requires a separate time interval for each possible task. This complicates the mathematical procedure demonstrated in Section 3. When populating a histogram with a single activity, or when extracting a set of probabilities, multiple TSI values must be used to reference histogram bins.

4.2 NPO Generation

To generate the numbers of participating occupants for each activity, we again use the TOD factor. Because we decided to generate an activity's NPO after generating its task, we are able to use the current task (TSK) as the other factor.

We believe TOD might well influence the NPO, as an occupant taking a break is more likely to have company around noon than in mid-afternoon. We again discretize the TOD at 15-minute intervals, but now use $\alpha_{TOD} = 0.25$.

It is obvious that the NPO should depend on the TSK, as a task representing meetings is more likely to be collaborative than, for example, a washroom break. A α_{TSK} value of 0 ensures that washroom breaks are always treated as individual activities.

When generating the task in Section 3, we required one feature per possible result. When generating the NPO, there are far more possible results. Here we use exactly two features, one called "sums" and another called "square_sums". For each activity in the real schedules, we first subtract $1/2$ from the number of participants npo. We add $(npo - 1/2)$ to the correct bin in the "sums" feature histogram, and $(npo - 1/2)^2$ to the same bin in the "square_sums" histogram. After n iterations of smoothing, we have for each bin $\langle bin \rangle$ the values $A_n[\langle bin \rangle]$, $B_n[\langle bin \rangle][\text{“}sums\text{”}]$, and $B_n[\langle bin \rangle][\text{“}square\_sums\text{”}]$. The normalization equations are as follows.

$$D[\langle bin \rangle][\text{“}mean\text{”}] = \frac{B_n[\langle bin \rangle][\text{“}sums\text{”}]}{A_n[\langle bin \rangle]}$$

$$D[\langle bin \rangle][\text{“}variance\text{”}] = \frac{B_n[\langle bin \rangle][\text{“}square\_sums\text{”}]}{A_n[\langle bin \rangle]} - D[\langle bin \rangle][\text{“}mean\text{”}]^2$$

During a simulation, with the time of day and the current task known, appropriate mean and variance values can be interpolated from D. We formulate a gamma distribution with these properties, randomly select a positive value from this distribution, add 1 to that value, then convert that real number to a positive integer by rounding down. The result is the NPO.

4.3 Duration Generation

We decided that the duration of an activity should depend on the time of day (TOD), the task (TSK), and the number of participating occupants (NPO).

The rationale for the TOD factor is that, for example, an occupant is less likely to work for several hours continuously if it is nearly time to leave the office. We again select 15-minute intervals, and use $\alpha_{TOD} = 0.23$.

It is obvious that an activity's duration is highly dependent on the task; in fact we insist that $\alpha_{TSK} = 0$. A positive α_{TSK} leads to absurdly long washroom breaks, as their durations become influenced by those of other types of activities.

The NPO factor is used because we imagine that an occupant may take a longer lunch break if he/she has company. We use separate bins for activities with 1 participant, 2 or 3 participants, 4 to 7, 8 to 15, 16 to 31, 32 to 63, and 64 or more participants. For smoothing, $\alpha_{NPO} = 0.02$.

Duration generation is similar to NPO generation. We again use two features, adding recorded durations to one feature and their squares to the other. This allows us to randomly generate durations using gamma distributions. As a duration is a continuous quantity, we do not need to offset or round off values as we did for the NPO.

5 RESULTS

Ideally, the fictional schedules generated by a our occupant behavior simulation would be indistinguishable from the real schedules used to calibrate the model. There is no single "best" metric to determine how well the two sets of schedules resemble one another. Here we present a brief qualitative analysis of our results, followed by a few statistics. In each case, the model was calibrated using the same 27 real schedules. These schedules were recorded manually during weekdays by an Autodesk Research employee, referred to in this section as "the real occupant". On some days, each activity was entered into a spreadsheet shortly after its completion. In other cases the occupant recorded all activities at the end of the day, aided by a webcam and motion detection software.

Observe the generated schedule in Table 2. The schedule appears plausible in several regards: the arrival and departure times; the fact that the occupant spends most of the day doing desk work; the 6-minute "desk meeting" in the morning; and the fact that the washroom breaks are spread out and range from 1 to 4 minutes in duration. The 13-person lunch break outside the building happens to be consistent with a few of the 27 real schedules.

Table 2. A generated occupant schedule.

Time	Task	NPO
9:57 AM	work@desk	1
10:19 AM	meet@desk	2
10:24 AM	work@desk	1
11:37 AM	break@washroom	1
11:41 AM	work@desk	1
12:01 PM	break@sharedroom	1
12:12 PM	break@washroom	1
12:13 PM	work@desk	1
12:28 PM	break@outside	13
1:31 PM	break@washroom	1
1:34 PM	work@desk	1
2:42 PM	break@washroom	1
2:44 PM	work@desk	1
4:08 PM	break@sharedroom	1
4:13 PM	work@desk	1
6:13 PM	off@outside	1

The real occupant never left the building and returned twice in a single day. This pattern of behavior appeared to be reflected in the generated schedules, presumably due to the TSI factor that measures the time elapsed since a task was last performed. But because breaks outside the building ("break@outside") were classified as separate tasks from breaks on site ("break@sharedroom"), the TSI factor did nothing to prevent the outside break in Table 2 from being preceded by the break at 12:01 PM. By combining "break@outside" and "break@sharedroom" into a single task, a modeler could discourage the generation of multiple breaks around noon. But on the other hand, one might then lose the tendency for a 13-person "team lunch" to take place at a restaurant, or the trend that 5-minute breaks around 4:08 PM tend to occur on site. The more general point is that different classification schemes will reproduce different behavioral patterns, but no single classification scheme will be ideal.

Figure 4 shows profiled probabilities that an occupant can be found working at their desk. The jagged line was produced by counting, for each minute of the day, the number of times the real occupant was recorded performing desk work, then dividing by the total number of recorded schedules (27). The smoother line was calculated in the same fashion, but using 10000 generated schedules. The real and simulated profiles follow roughly the same path, though evidently the simulation overestimates the probability of desk work around lunch and underestimates it before and after.

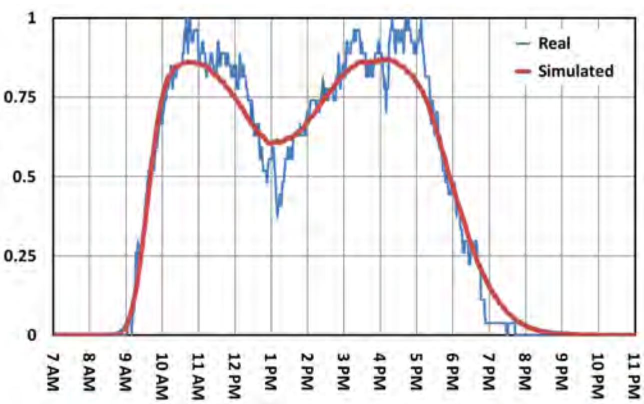


Figure 4. Desk work probability profiles.

Despite the differences between the profiles at certain hours of the day, the real and simulated time-averaged probabilities of desk work were in fairly close agreement. On average, the real occupant worked at their desk only 3.4 minutes longer per day than the simulated occupant (411.6 minutes total compared with 408.2).

While desk work is typically performed several times per day, the occurrence of certain other tasks is best measured over several days. The real occupant took a break outside on 14 of 27 days, and so the desired probability that the "break@outside" task occurs in a day is about 51.9%. From 10000 generated schedules, the result was roughly 53.6%. Given that we are trying to predict human behavior, the error of 1.7% seems tolerable.

6 FUTURE WORK

Recall that Tabak's USSU System [10] allowed simulated occupants to interact with one another, sharing activities such as meetings and presentations. Although we do quantify the number of participants for each activity, we have yet to present a means to utilize this attribute for occupant interaction. For example, suppose our method is used to simulate a building with 10 occupants over the course of a single day. If at some point a 5-person activity is generated for one of those occupants, then the same activity should occur at the same time in the schedules of 4 of the remaining occupants. So long as the 10 schedules are generated independently, however, this is extremely unlikely. Adding interactions between occupants, such that simultaneously-generated schedules are inter-dependent, remains important future work.

Discussing the schedule-calibrated method of [9], Page points out that software users will find it impractical to supply the large amounts of input data necessary to yield realistic simulated behavior. The problem applies to our own method as well, and is exacerbated by the possibility that users will want to populate their simulated buildings with different types of occupants. Designing a building for software company, for example, an architect may wish to perform simulations with different numbers of junior programmers, senior programmers, managers, and sales and support staff, each with their own behavioral patterns. We hope to combine our method with personas, descriptions of fictional individuals, to allow occupant behavior to be customized using only modest amounts of additional information.

Once we have developed a customizable model of the behavior of interacting occupants, we will need to combine it with models of other subsystems in an effort to predict building performance. At very least, the occupant behavior model should influence models of building equipment. If a simulated occupant begins performing desk work, for example, his/her simulated computer should respond and draw additional power. It is important to note that interactions can occur in the opposite direction as well, with other building subsystems influencing human behavior. If an HVAC system produces an intolerable temperature increase in a working area, an occupant might respond and move to a different location. Alternatively, he/she may open a window, impacting the HVAC system and potentially the actions of other occupants. In some cases it may be desirable to allow "exceptional behavior", like opening a window or vacating an uncomfortable area, to take precedence over the activities generated by our schedule-calibrated method.

7 CONCLUSION

A number of methods have previously been developed to simulate the behavior of building occupants based on the recorded schedules of real occupants. The schedule-calibrated method we have proposed and demonstrated is notable for its flexibility; one can determine the level of detail with which occupant behavior is modeled by selecting various activity attributes, and one can alter the behavioral patterns that get reproduced in a simulation by selecting different sets of factors. Tested with our own chosen attributes and factors, the method yielded plausible fictional schedules with acceptable statistical accuracy. Future work includes the modeling of interactions between occupants, the customization of occupant behavior using personas, and the integration of occupant models with those of other building subsystems in an effort to predict energy demand.

REFERENCES

[1] McKinsey Global Institute. *Curbing Global Energy Demand Growth: The Energy Productivity Opportunity*, 2007.

[2] Joe A. Clarke. *Energy Simulation in Building Design*. Butterworth-Heinemann, second edition, 2001.

[3] Jan Hensen, Ery Djunaedy, Marija Radošević, and Azzedine Yahiaoui. Building Performance Simulation for Better Design: Some Issues and Solutions. In *Proceedings of the 21st Conference on Passive and Low Energy Architecture (PLEA)*, pages 1185–1190, Eindhoven, Netherlands, 2004.

[4] Autodesk Research, CIMS Lab, and Faro Technologies. *Digital 210 King Project*. www.digital210king.org.

[5] Gerhard Zimmerman. Modeling and Simulation of Individual User Behavior for Building Performance Predictions. In *Proceedings of the Summer Computer Simulation Conference (SCSC)*, pages 913–920, San Diego, California, 2007.

[6] Pieter-Jan Hoes, Jan Henson, M.G.L.C. Loomans, B. de Vries, and D. Bourgeois. User behavior in whole building simulation. *Energy and Buildings*, 41(3):295–302, 2009.

[7] Danni Wang, Clifford C. Federspiel, and Francis Rubinstein. Modeling occupancy in single person offices. *Energy and Buildings*, 37(2):121–126, 2005.

[8] Yohei Yamaguchi, Yoshiyuki Shimoda, and Minoru Mizuno. Development of District Energy System Simulation Model Based on Detailed Energy Demand Model. In *Proceedings of the Eigth International IBPSA Conference*, pages 1443–1450, Eindhoven, Netherlands, 2003.

[9] Jessen Page. *Simulating Occupant Presence and Behaviour in Buildings*. PhD thesis, École Polytechnique Fédérale de Lausanne, Switzerland, 2007.

[10] Vincent Tabak. *User Simulation of Space Utilisation: System for Office Building Usage Simulation*. PhD thesis, Eindhoven University of Technology, Netherlands, 2008.

Finding Synergy in Simulation Modeling by Architects and Engineers in Conceptual Design

Alexander Hirsig
Harvard University Graduate School of Design
Gund Hall, Cambridge MA 02138, USA
ahirsig@post.harvard.edu

Keywords: survey, sustainability, conceptual design, barriers, architects, collaboration.

Abstract

The interest in having architects utilize building simulation software to better inform conceptual design decisions has been present in the design industry for more than a decade, yet the adoption of simulation by architects in practice has been slow. Past research into the barriers of adoption has often looked through the lens of particular software, with conclusions aimed at improving the tools' interface and applicability to the architects' process. While software has made great advancement, we hear more often today of the engineer stepping forward to apply simulation to the conceptual design process rather than the architect learning to use simulation. In a survey of ten North American firms, diverse in offering design and/or engineering services, individual practices for incorporating simulation into conceptual design were compared. The more common presence of the engineer in conceptual design modeling promises a new collaborative approach to the sustainable design process, overcoming a significant barrier to the architects own adoption while highlighting a need for greater efficiencies in communication between software and analysis time.

1. INTRODUCTION

ASHRAE describes conceptual design as "the first crucial element in producing a green building" (Gruman 2003). For many architects, designing a "green building" has an immediate association with the many passive and active design strategies devised from earlier periods of interest in sustainability. In representative literature such as D. Watson and K. Lab's architectural guide Climatic Design (1983) 50 sustainable design strategies are illustrated for consideration in the conceptual design phase, ranging in such issues as site selection, microclimate control, massing, and building articulation. These simplified design strategies help elucidate for architects important principles of complex physical interactions between the building and its environment, while providing "rules of thumb" that can be implemented in the brief timeframe of conceptual design without extensive engineering analysis or knowledge in physics. Building simulation tools are anticipated by many to be the next generation to such assumptions. Capable of efficiently resolving the large amounts of data describing the complex dynamic relationships between the building design, its active systems, and the environment, designers can evaluate the impact of various conceptual design decisions on building performance.

Leaders in the field of building simulation have long stressed the importance of getting the benefits of simulation analysis into the hands of the design decision maker (Clarke 2001) (California Energy Commission 1999). Conceptual design has been the sacred domain of the architect and performing simulation starts to redefine this boundary, blurring traditional responsibilities between the disciplines of architecture and engineering. Since their adoption in the mid 1970's, building simulation has been performed almost exclusively by the engineering profession, with their application historically regarded as an engineer's tool for assuring code compliance (Nall 1985). Until recently, the software was never written to be understood outside of the needs of the engineer and modeling in applications such as DOE 2 was done purely through text input. New applications present user friendly interfaces and 3d modeling environments for design visualization while allowing the architect to interpret the significance of their analysis as an engineer might. Still, prominent voices like ASHRAE publish it is "virtually impossible" to design a green building without major involvement by the engineer well before any programming and massing studies are completed (Gruman 2003). It is understandable why both architects and engineers appear vying for the role of simulator in conceptual design. By surveying both engineers and architects simulating in conceptual design, it is intended to better understanding the roles and emerging relationship of architect, engineer, and simulation around conceptual design.

2. SURVEY PARTICIPANTS

The simulation user survey was taken in spring of 2008 and targeted primarily larger offices within North America providing Architectural design, Engineering design, or combined services (A&E). The firms were identified primarily from trade publications, conference proceedings, and simulation software developers for their efforts to integrate simulation in design practice. The firm profiles' show they tend to work on projects of a similar magnitude and by location operate within a similar market for the potential to introduce simulation tools. With their multiple offices and strong presence in the industry these firms are anticipated to provide a fairly balanced reflection of the general perception and trends of simulations' adopted use in practice. Firms solicited were: Gensler, HOK, Perkins + Will, SOM, Burt Hill, Cannon Design, Stantec, Transsolar, Atelier Ten, and Buro Happold.

3. SURVEY RESULTS

Nine of the offices responded directly to the survey. Follow up phone interviews to better understand specific write-in responses and processes were made with five of these firms. Insight from one A&E firm is limited to information via phone and email conversation only and included only as applicable within the standard questions. Seven out of the ten offices offer architectural services, three of which offer only architectural services while the remaining four offer both A&E. Three of the ten offices offer only engineering services[1]. Six respondents have an educational background in architecture (two of which have additional education in engineering); four reported an educational background in engineering.

A distinct division was found between firm services and years of in-house use of simulation tools. Offices providing only architectural services had adopted the software within the past two years. All remaining offices offer engineering services and responded having used building simulation software for more than five years.

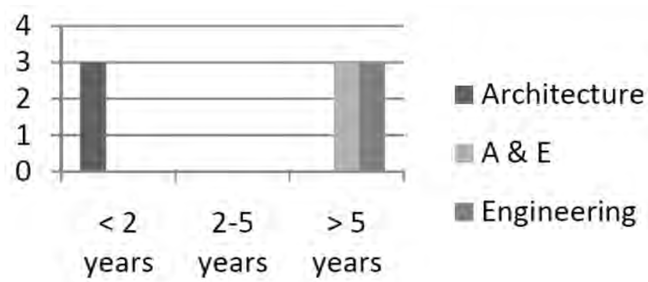


Figure 1. History of simulation software use by firms.

Respondents were asked to estimate the percentage of office projects in which building simulation was performed at some phase; and from these, the percentage which utilized simulation in conceptual design. The results in Figure 2 suggest increasing the use of simulation tools at any phase on projects results in a higher likelihood the tools will be used to aid in the conceptual design process.

Nine firms collectively listed seventeen different simulation software used during the conceptual design phase. Reasons for selecting particular software were multiple and varied by firm, however ease of user interface and speed of input were the most universally given. Engineers identified a much broader array of simulation software, typically capable of more complicated or specific analysis.

[1] Atelier Ten provides environmental consulting, for this survey their role appears similar to consulting engineers.

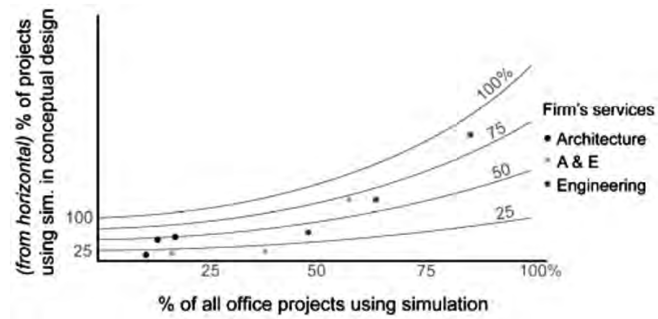


Figure 2. Both frequency of simulation on projects, and for conceptual design.

In all offices providing both A&E services the engineer was said to be typically responsible for performing any conceptual design simulations. In offices providing only architectural services a specific person(s) was responsible for simulation across multiple projects. The most common means for simulation was described as an in-house specialized team created to support sustainable building design and simulation, typically located in one office bridging to support the firm's other offices. Within consulting engineering offices it was common to hear that all employees should be familiar or capable of some simulation while specific persons had expertise in particular simulation types such as thermal modeling, CFD analysis, or lighting. In all but one office, the analysis results were said typically communicated second hand through a printed report or presentation at a design team meeting.

In follow up, users were asked if there were apparent advantages that the simulation was performed by the architect; or if the architect had done the simulation if they believed it would have been advantageous. Four engineers similarly noted that architects generally lack the technical knowledge to be able to interpret the results of much building simulation, so its success in use is limited by the complexity in analysis. A few of the architects noted that while they see an advantage in faster feedback time, the support of the engineers is essential for certain types of simulation. The suggested comfort boundaries of simulation analysis types for the architects became more evident through the module use patterns that emerged in the final question. All users agreed that the use of the simulation facilitates increased communication between architect and engineer.

The work required to prepare a data model for simulation has in the past been identified as a significant barrier to the tools adoption by designers (Grau and Wittchen 1999). The majority of users in this survey responded that designs are typically generated through "other tools" and then imported or reconstructed in the simulation software environment. One reason was it being easier to find persons to build conceptual models in software such as Rhino or 3ds Max, in addition to preferences for their graphic output. Another reason was keeping continuity with the data model in the greater design process; leading one user to create design variations through BIM software and export the multiple options in gbXML format for analysis, rather than creating variations directly in the simulation software. Multiple A&E firms reported utilizing the gbXML exchange format for their in-house transfer of the architect's design to the engineer's simulation.

No users identified being comfortable with any simulation software package in less than a week. Multiple users (particularly architects) noted feeling comfortable in the software Ecotect within two to four weeks. IES <VE> was consistently noted requiring more than 1 month, and respondents suggested a range of up to two years before being competent in some applications. In addition to the user interface of the software, it was noted by multiple users that one's education in building physics and engineering was important to the time it takes to learn the software. Hands-on 'trial and error' experience was identified by nearly all users as the most useful resource in learning the software. Second was learning directly from other experienced simulators. Vendor training sessions and tutorials were both equally recognized to a lesser degree. Confirmation with engineers on how the systems should appear behaving was noted by one architect modeling.

Offices were asked if they perceived limitations to simulation use because of a project's size or building typology. No respondents found building typology as a limitation. Limitations did arise from the inadequate support for importing or modeling a complex building design. Multiple firms noted a strong parallel between a projects budget (often proportional to its size) as a limitation to its use. A scenario was provided comparing a small project with a highly innovative design which may require simulation of an original or complex system for validation, and a large project with a more conventional design approach which may require simulation either for code compliance or for the magnitude of environmental impact small adjustments can have. The time and cost required for simulation of both projects can be identical but for the smaller building may represent one-fourth of the overall project budget available, making it unfeasible.

Only two out of nine offices believed the use of the software brings increased liability to the firm, in that the use of simulation early in the design provides a quantifiable basis for decisions by the client, from which the many assumptions required in modeling at this phase yield potentially serious consequences if later discovered unfeasible or mistaken. Based on reported higher frequency of use by these offices it does not appear that the sense of increased liability inhibits the software's routine use, rather cautions its use. One firm felt that the "add service fees" properly assumed taking on and managing the additional liability.

An additional fee for energy modeling was suggested during an Architectural Record interview as an industry means to realize its use (Fortmeyer 2006). Two of three offices providing only architecture services reported no additional service for simulation. Interestingly, the two offices with architectural services (one Architecture and one A&E) reporting additional fees also reported among the lowest for percentage of projects incorporating building simulation. Conversely, from the four offices reporting the highest use of building simulation among projects no conclusions could be drawn as two reported simulation as an additional service. When asked what prompted the use of simulation in conceptual design, it was most often reported as a means for the design team to confirm the project is on target with the goals set by the client.

In an open question of what these users saw as the "biggest limitations for the simulation software's regular use in conceptual design" a string of barriers were cited, however no more than two respondents ever cited the same limitation. Those limitations identified by two users include: inadequate time available in conceptual design to construct a meaningful model, lack of experienced simulation operators available in the industry, architects lack of knowledge to assess if simulation results are accurate, finding engineers interested in participating in the conceptual design phase, tradition and unwillingness of designers to change old processes, and the infancy of the software as limiting its adoption. Limitations listed by only one user include: lack of agreement among professionals on the validity of the software predictions (arising from different software providing different results), site constraints frequently precluding significant variation in massing and orientation therefore reducing benefit from simulation, and the discontinuity in use by architects. The limitation from discontinuity was elaborated to say architects will typically follow a project through construction which can result in a substantial time lag between the next conceptual design exercise and the software's use. This can be exacerbated by the relatively short duration of conceptual design, reducing ones period of experience and making it difficult to become familiar with the tool's interface, develop modeling strategies, and stay abreast of the software's updates.

In the final question, offices were asked to estimate the frequency in which simulation software is being used to explore the range of traditional sustainable design strategies suggested in Climatic Design in addition to others during conceptual design. The seventeen category matrix is grouped around seven main themes including Site Planning, Daylighting, Thermal, Programming, and Energy.

For architects modeling, the most frequently used simulation module was for daylighting studies, with the most frequent analysis being studying daylighting benefits from sun shading and reflecting devices. The study of passive thermal efficiencies in the building were lesser explored and quickly fell to 'rarely and never' in exploring natural ventilation. A few engineers in the survey linked this falloff to the expertise required in properly operating this and other sophisticated[2] building simulation analysis. The frequency with which the architects were simulating to study the effects of exterior wind exposure on the proposed design was also predominately 'rarely and never.' A few of the engineering offices offer CFD simulation in-house, which was said to increase its viability for their use on projects, but also noted an increased difficulty in CFD modeling for exterior wind studies, include having an operator experienced with CFD analysis, the complexity of defining boundary conditions, and invested time/cost to evaluate a single fixed wind direction when reality holds much greater unpredictability.

The tools currently available were shown to not yet fully support the broad interest of sustainable design strategies historically considered by architects at this phase. Important to near twenty percent of sustainable design strategies in D. Watson and K. Lab's Climatic Design, the surveyed revealed significant isolation between landscape and dynamic simulation. Frequency of studying impacts of landscaping on building performance was performed 'rarely' by all firms. Multiple engineers noted that the software simply isn't developed enough for this type of analysis, that building simulation software are typically designed to support up to the building envelope. Another limitation suggested that landscaping takes years to mature and caries unpredictability, requiring engineers to essentially design without landscaping in mind in order to meet building comfort requirements at initial occupancy.

Architect's reported on average 'rarely' using the simulation tools for building energy consumption analysis. From this figure and comments by users it appears the practicality of architects using these tools diminished at the introduction of mechanical systems into the analysis. One architecture office stated they simply "do not see themselves becoming energy modelers with these tools."

4. CONCLUSION

The climate design strategies architects relied upon decades ago are gradually being replaced by trust in computation. A renewed social priority in building performance coupled with increasing design complexity is also demanding greater accountability and specialization within the design industry, while the recent trend toward Integrated Practice is challenging the traditional structure of the conceptual design team and inviting the engineer to be an active voice in the early process. The majority of both architects and engineers in this survey shared the sentiment that the average architect's technical knowledge falls short of what is required for many of the analysis types offered by these tools, both in recognizing accurate model results and the ability to interpret their broader impact (such as the impact of more daylight on acceptable thermal gains/losses). Whether the architect or the engineer performed the simulation modeling did not diminish the frequency of simulation in conceptual design. If our goal is simply to see the benefits of simulation applied at conceptual design then a foreseeable future is one in which the engineer is engaged for design input and simulation modeling.

Additional evidence to suggest a collaborative approach may prevail is the ongoing effort necessary to efficiently and effectively use simulation tools. While engineers' often continue the use of building simulation through final design, the architect can find infrequent use beyond this phase. Architecture firms in this survey appeared able to overcome this potential time lag by having an individual bridge multiple projects and in some instances offices. A small to mid-size firm's means will likely influence how simulation is approached, and as an equal share of the design industry would be complementary to research.

In a collaborative arrangement, the ability of the architect to exchange the design model with the engineer in an efficient and effective manner becomes highly valued. The time required for constructing a model for meaningful analysis was seen as a significant limitation. Unfortunately there was little agreement on which specific limitations in the software are the most significant. Nearly twice as many software packages were used than the number of firms participating, complicated by individual process and type of simulation sought. To encourage simulation in collaborative design, efforts might further develop the methods and content of exchange with prominent design software.

[2] "Sophisticated" in this context is understood to refer to CFD, natural ventilation studies, and energy analysis.

References

California Energy Commission. "Building Design Advisor." Buildings End-Use Energy Efficiency, Public Interest Energy Research Program (PIER), 1999.

Clarke, J.A. Energy Simulation in Building Design. 2nd Edition. Oxford: Butterworth Heinemann of Reed Educational and Professional Publishing Ltd., 2001.

Fortmeyer, Russell. "When Less Powers More." Architectural Record, December 2006: 164-168.

Grau, Karl, and Kim B. Wittchen. "Building Design System and CAD Integration." IBPSA. Kyoto: IBPSA, 1999. 607-615.

Gruman, David L. ASHRAE Green Guide. Atlanta, GA.: American Society of Heating Refrigerating and Air-Conditioning Engineers, Inc., 2003.

Nall, Daniel H. "Building Energy Simulation and the Architect." Building Simulation 1985. Seattle: International Building Performance Simulation Association, 1985. 45-48.

Watson, Donald, and Kenneth Labs. Climatic Design: Energy-Efficient Building Principles and Practices. McGraw-Hill Inc., 1983.

The Power of Data

210 King Street: A Dataset for Integrated Performance Assessment

Ramtin Attar[1], Venk Prabhu[2], Michael Glueck[1], Azam Khan[1]

[1]Autodesk Research
210 King Street East,
Toronto, Ontario, Canada M5A 1J7
{firstname.lastname}@autodesk.com

[2]Carleton Immersive Media Studio
Carleton University
Ottawa, Ontario, Canada K1 5B6
vprabhu@connect.carleton.ca

Keywords: Dataset, Point cloud, Building Information Modeling, Performance analysis.

Abstract

This paper presents a Building Information Modeling (BIM) re-creation of a designated heritage building located in Toronto, Canada. By taking advantage of BIM as a centralized database, which describes both geometric and semantic aspects of a building, this model can be leveraged as a source of input for many forms of analysis. In addition to the BIM model, we present a comprehensive point cloud dataset gathered using terrestrial laser scanning technology. Based on an existing and a living building, this model is an ideal candidate for simulations that can be cross referenced with information gathered on-site.

1. INTRODUCTION

As the result of an increasing emphasis on energy efficiency and building performance, we are witnessing a growing global trend calling for further integration of sophisticated computational analysis and Building Energy Performance Simulation (BEPS) to support design decisions. Simulation is a powerful method for gaining insight into the behavioral characteristics of a system and it is becoming an integral part of the design evaluation process. Numerous projects are attempting to define the qualitative and quantitative measures of sustainability and "greenness" of a building. This is a complex task and many different kinds of data are needed to implement and asses these measures.

In the context of architectural design and urban planning, there is currently no common method for describing and representing data required for simulation analysis [Hand et al. 2005]. In order to integrate simulation as a core philosophy throughout the life-cycle of a building, we need to enable more effective mechanisms to create, distribute and cross validate information critical to simulation. In this paper, we contribute a comprehensive and high quality dataset of an existing structure to provide a basis for ongoing analysis within the simulation community. While noting that the overall performance of a building can be characterized as a trade off among many different attributes and aspects of design, our dataset can provide an opportunity for comprehensive cross validation among various methods of analysis performed on a shared model. Furthermore, we hope that this kind of data sharing will foster more collaboration between the scientific and design research communities.

2. DIGITAL 210 KING DATASET

This dataset is a digital re-creation of a designated heritage building situated at 210 King Street East in downtown Toronto, Ontario, Canada (see Figure 1). This eclectic workspace is the current home of Autodesk's Toronto office and it spans four historic Toronto warehouses, built between the 1930s and 1960s, with a total 145,000 square feet of office space. Toronto architecture firm Kuwabara Payne McKenna Blumberg was commissioned to carry out the integration and renovation of the warehouses, which was completed in November 1997. At the time, the company was called Alias|wavefront and was subsequently acquired by Autodesk in January 2006.

Figure 1. 210 King building combined with 204 and 214 King Street Warehouses.

Three of the warehouses, including 204 and 214 King Street East, are designated as heritage buildings, requiring that original features of the buildings be preserved. 210 King Street East, on the other hand, suffered damage from a fire and did not receive the same designation, making it an ideal candidate for the main entrance to the office. This space was renewed through the creation of an expansive two-storey lobby and striking steel staircase, the signature feature of the office. In 2009, growing out of an interest to foster awareness towards energy consumption and to promote more sustainable practices for existing buildings, the 210 King building was chosen as a living laboratory to analyze and investigate factors involved in building performance and energy efficiency. We started the process by digitally reconstructing (reverse engineering) the existing building into a BIM model using laser scanning, collected architectural blueprints and site inspection.

2.1. Point Cloud

Terrestrial laser scanning is a powerful and relatively fast survey method for collecting 3D information at a large scale. The first digital dataset we present in this paper consists of an accurate geometric documentation of the 210 King office building by collecting real-world spatial data as points in 3D space. We limited the scope of the scanning to the entire fifth floor interior and the rooftop terrace, as well as the lobby and the exterior of the building. In total, we carried out 53 individual scans, which were combined to produce a dataset containing over 1.3 billion data points (see Figure 2). These point sets were registered in terms of coordinate systems and elevation above the sea level by a survey team. This allowed for geo-referencing of the entire dataset by knowing the precise position of each point in terms of real-world coordinates.

Figure 2. The combination of 53 scans as one dataset.

2.2. Building Information Model (BIM)

In addition to the point cloud representation of our office building, our dataset includes a comprehensive Building Information Modeling reconstruction of the 210 King office building using Autodesk Revit. Due to the complexity of reverse engineering an existing structure, we treated the creation of our BIM model as an iterative modeling exercise. Our first attempt was a tremendous exercise in filtering through the mountain of information we collected from laser scans, hand drawn sketches, AutoCAD drawings, and on-site inspections.

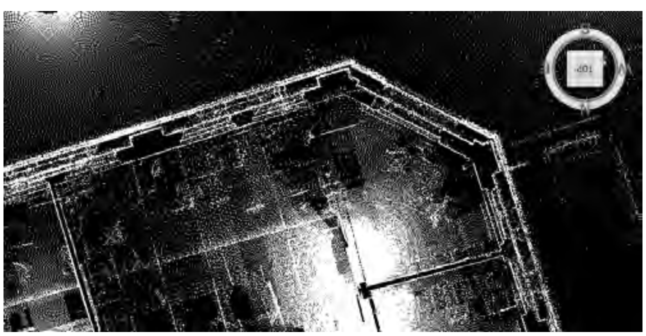


Figure 3. Plan of building envelope in Autodesk Navisworks after combing interior and exterior scans.

We used the AutoCAD drawings of the building from the time of renovation as a basis for our model and integrated measurements from the other sources to fill in missing information, such as wall thickness measurements and the composition of brick walls (see Figure 3). Through a layering approach, we built our first iteration of the model from bottom up by compiling all the CAD drawings in appropriate levels while cross referencing them with our laser scans and site inspections.

The first version of our model provided a great environment where we could compare and reconcile the complexities of the many existing datasets. However, we found that our approach to the problem of modeling an existing building needed to be reviewed. We approached the first model by considering the existing office building as a whole from the very start, which resulted in a lack of conceptual clarity among the four buildings. Since our digital reconstruction involved four different buildings built at different times, the interaction between the buildings themselves had to be understood beyond simple connecting blocks. Accurately capturing these interactions would not be easy, since the overall geometrical and structural configuration of the four buildings could not be simply rationalized into one method. Thus, in our second attempt, we thought of this task as if we were designing four new buildings that would later interact. We started the model by constructing the major party walls defining each of the four buildings, thus embedding a clear conceptual separation among the four existing structures. With each building still being considered separately, we added internal structural components: the pillars, the floors, and major interior walls. As a final step, we created openings in the walls where the buildings now connect to one another (see Figure 4).

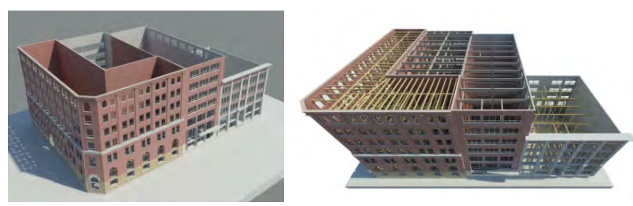


Figure 4. (left) Completed envelopes and (right) structural framing of the four buildings.

BIM provides a powerful mechanism for the assembly of components that can be repeated and placed in different locations. However, one of the key challenges in using BIM for an older building was how to best reconcile between traditional methods of construction and the family-based (component-based) approach of BIM. In the case of the 210 King project, we were faced with many interior elements that could not simply be reduced to repeatable component families since most of these elements had been thought through and designed on-site in order to fit a specific location. Therefore, most of these elements had to be treated as the assembly of much smaller components that were combined together based on their local configuration. By establishing the entire building's structural grid and each individual building's structural wall as a datum, we were able to create an overall anatomical setup for interior modeling. Additionally, we paid more attention to the visual fidelity of our model by creating drawings and rendering outputs throughout the process. In particular, the laser scan data proved invaluable for refining measurements of difficult to reach areas, such as the building envelope.

3. VALUE TO THE COMMUNITY

Despite significant advances, simulation programs still require a tremendous amount of user input about building components. Additionally, there is generally little information about how a dataset was created, how it is maintained and what purposes it is suited for [Hand et al. 2005]. Furthermore, full access to a comprehensive dataset that can be utilized for many different simulation tools is still rare to find. For this project, all the collected and authored materials are published for unrestricted use at www.210king.org.

Given the natural ability of BIM to maintain different kinds of information pertaining to the assembly of a building, spatial zoning, operational management and additional metadata, we can expect a more efficient and reusable dataset that could save a tremendous amount of manual input work. Based on a living building, BIM allows for semantic data, such as operational and contextual information, to be captured within an integrated information model shared by various parties and stakeholders.

While BIM acts as a powerful database of a building, no simulation tool supports the detailed level of attributes inherent in a BIM model. Therefore, there are many factors to be considered from the outset. In the context of the digital reconstruction of an existing building, a key factor is to determine a balance between the immediate function of an asset and its long term re-purposing in other contexts [Jemtrud 2006]. Therefore, we decided to approach our model as a comprehensive repository of data by choosing BIM as a scalable-integrated approach to life cycle information management and building performance analysis. BIM has already been utilized as the initial step in creating energy performance models and thermal performance management, for instance, is one of the most adopted use case scenarios of BIM models [Laine et al. 2007]. BIM has also been extended to different simulation scenarios, such as crowd simulation [Eastman et al. 2008]. As Eastman has noted, the use of BIM has long term advantages due to its capacity to store accurate 3D input and metadata required by simulation tools.

A key obstacle is that simulation is often performed on a completely different platform outside of the design environment. Another important goal of using BIM is to help the research community toward a better understanding of the mindset of a designer. There is a broad spectrum of models for measuring the energy performance of a building. On one side of this spectrum are specialist research users who develop a model for a very specific research task and on the other side are architects and engineers who make use of models as a design tool. The quality and efficiency of a model is domain-specific and it depends largely on practical knowledge and tolerances specific to a field. Thus, it is crucial to share the datasets as an open source assets in order to create an ongoing discussion among various disciplines. Furthermore, performing simulation on a common dataset can help to reveal certain aspects of each simulation product through cross validation while promoting a more holistic understanding of simulation in the context of design.

3.1. Heritage/Retrofit

The application of BIM as an integrated design method has brought significant advances to design and delivery of new construction projects. In this project we aim to extend the concept of BIM to an existing heritage complex where we can analyze the behavioral and performance characteristics of our building as a living system. In recent years, there have been a number of precedents [Penttila et al. 2007] that have explored the use of BIM in retrofit projects and we expect to witness a growing trend in integrated-performance assessment as our existing buildings are being scrutinized for their energy consumption and operational performance throughout their life-cycles.

3.2. To Support Validation

Measuring a building’s performance has many practical challenges and different kinds of data are required from both the actual condition of the building and results of a simulation model (performance model). While for some metrics it is relatively easy to obtain performance data from existing conditions, for others mathematical models and estimates are more practical. Our interest in creating a dataset based on an existing building was to take advantage of the analytical possibilities of a 3D digital model by setting the real building and its digital proxy into parallel analysis and cross validation.

4. HOW TO USE THE DATASET

Our BIM model acts as a repository of data that describes different aspects of the 210 King building. The main motivation behind using BIM is to centralize the building data while leveraging it for various analysis purposes. This would minimize the task of making input files for simulation since we can virtually incorporate any form of metadata into our BIM model. There are already a number of existing workflows among Autodesk Revit and some commercially available simulation software. For instance, using Autodesk Revit, one can export an analytical model that can be used either in conjunction with Autodesk Green Building Studio to evaluate the energy profiles and carbon footprints of the building, or it can be imported to Autodesk Ecotect to perform a full range of environmental analysis and simulation [Autodesk]. This workflow is facilitated through the gbXML extension, which focuses on a very specific set of definitions and data requirements for energy analysis. Alternatively, a similar workflow can be followed to interface with other simulation engines such as EnergyPlus.

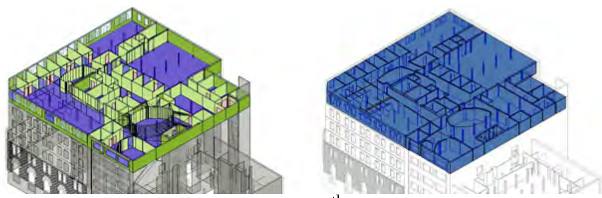


Figure 6. Analytical models of 5th Floor using gbXML.

It is important to note that our analysis is not limited to energy modeling and environmental analysis. All the elements in a Revit model are described with architectural semantics and we can assign additional data to each element or family of elements. BIM's geometry and semantic data model can also be extended by adding behavior patterns through its Application Programming Interface (API). This method allows external simulation platforms to perform more sophisticated simulation analysis concerning building usage and post-occupancy user analysis. For example, Yan and Liu [2007] combines a BIM model with a simulation engine built with Microsoft XNA Game Studio Express for path-finding simulation and visualization while integrating different kinds of simulation components such as fire and smoke. This would make the BIM model an ideal dataset for running other forms of analysis such as crowd simulation or fire evacuation. In general, our BIM model can be scaled or altered to fit a specific analysis.

4.1. File Structure

In general our dataset is divided into two groups: laser scan and BIM files. We have also converted our datasets into several common formats.

File Name	Use Case
210King-Pointcloud-Combined.xyz	This file has combined all 53 scans into one large dataset of points in ASCII text format.
210King-BIM.rvt	This file is created in Revit Architecture 2010 and it contains all the BIM data for our project.
210King-Sheets.PDF	This file is a set of architectural drawings of 210 King that can be used for quick inspection and review of the project.
210King-IFC.ifc	Industry Foundation Class export of Revit file.
210King-Thermal.xml	This file is a thermal model exported from Revit MEP that can be used for analysis in Ecotect or Green Building Studio. Alternatively you can use the Revit file to re-appropriate this file.
210King.obj	This is a basic Wavefront OBJ file that can be used in most 3D applications.

Table 1. Files included in the 210 King dataset.

5. CONCLUSION AND FUTURE WORK

In this paper we have presented a comprehensive dataset going beyond the geometrical and superficial description of an existing building by digitally reconstructing an accurate, high-resolution model. Within this context, a combination of old hand-drawn sketches and CAD drawings had to be studied and sorted in order to digitally reconstruct the building. An accurate geometrical documentation was performed using laser scanning technology and a comprehensive BIM model has been assembled to support various simulation analyses.

Our BIM model currently does not include some subsystems. As future work, we would like to complete our dataset by accurately modeling all the HVAC, ducts and electrical systems. By incorporating sensor networks, we are also planning to provide robust building performance data based on existing conditions that can be used for validation and calibration of simulation outputs.

The application of BIM for existing buildings, especially heritage buildings, is a new area of research. There is much work to be done in order to develop guidelines and efficient methods of how to model existing buildings that often lack sufficient architectural documentation. Our current process of modeling is not efficient, especially as far as taking advantage of powerful datasets such as laser scans. Therefore, we would like to further investigate more efficient and automated methods of working with laser scans and CAD models.

Acknowledgements

We thank FARO Technologies for their generous support and help in developing the 210 King point cloud data set. Also thanks to Julius Ndulue, Applications Engineer at FARO Technologies, Tobias Hathorn and Kyle Bernhardt, Autodesk AEC Division, for their technical support of this project.

References

Autodesk http://images.autodesk.com/adsk/files/building_performance_analysis_using_revit.pdf

Eastman, C., Teicholz, P., Sacks, R. and Liston, K. 2007. "*BIM Handbook: A Guide to Building Information Modeling for Owners, Managers, Designers, Engineers and Contractors*." John Wiley & Sons, New Jersey,108.

Hand, J.W., Crawley, D.B., Donn, M. and Lawrie, L.K. 2005. Improving the Data Available to Simulation Programs. IBPSA Building Simulation. 373-379.

Jemtrud, M. 2006."Technological Mediation and Digital Craft." Practices Vol. 7. 25-35.

Laine, T., Hanninen, R. and Karola, A. 2007. Benefits of BIM in the Thermal Performance Management. IBPSA Building Simulation. 1455-1461.

Penttila, H., Rajala, M. and Simo, F. 2007. Building Information Modelling of Modern Historic Buildings. eCAADe. 607-613.

Yan, W. and Liu, G. 2007. BIMGame: Integrating Building Information Modelling and Games to Enhance Sustainable Design and Education. eCAADe, Frankfort, Germany. 211-218

Intuitive Structures: Applications of Dynamic Simulations in Early Design Stages

Andrzej Zarzycki
New Jersey Institute of Technology
University Heights
Newark, NJ 07102-1982
andrzej.zarzycki@njit.edu

Keywords: dynamics simulations, building information modeling, BIM

Abstract

Developments in digital design have brought a new design freedom into architecture. Emerging tectonic trends, combined with research into new materials and fabrication technologies, make it possible to pursue imaginative designs with new expectations of space and form. However, these innovative designs often exist exclusively as visual propositions, deprived of the deeper structural, constructional, or functional logic necessary for well-developed designs. Similarly, the proliferation of analysis software tools has helped engineers to calculate sophisticated structural models, yet this ability seldom translates back into architecture. Consequently, these two parallel developments, while promising in their individual capabilities, fall short in terms of synergizing into successful designs. To bring these two distinct components together, this paper discusses the strategies for generative design validation using dynamics-based modeling tools that realistically portray physical processes. Through the use of dynamics-based software, a promising direction for generative architectural design emerges. An architectural form not only can be analyzed based on its structural performance, but also can be derived through the process of structural simulations.

1. INTRODUCTION

Emerging computational tools and theories expand the tectonic language of architecture and renew research in the nature of design creativity. While design outcomes often challenge established geometries and traditional architectural forms, they frequently converge on broader aesthetics of contemporary design, taking cues from other creative disciplines such as product design.

In this pursuit, digitally inspired thinking renegotiates conventional boundaries between architecture, landscape design, sculpture, and structures. These renegotiations respond to the broader contemporary cultural context of a unified or holistic view of the world. However, these movements toward integrated design often remain within a sphere of intentions rather than actions. To a large degree, this is due to an insufficient level of design research, experimentation, and practice dealing with the actual implementation strategies for these ideas.

While digitally inspired thinking allows for a broader reading of architecture, promoting innovative and unique designs as well as new expectations regarding its spatial, formal, and material characteristics, these emerging designs often exist exclusively as visual propositions, deprived of a deeper structural, constructional, or functional logic.

Structural analysis software has helped engineers in calculating sophisticated structural models and understanding the intricacies of complex structural strategies. However, the ability to model such structures is seldom utilized in the development of architectural forms, and in everyday practice it rarely informs the design process or design criticism. Consequently, these two parallel activities (advanced structural modeling and architectural form making), while promising in their individual capabilities, have not yet been synthesized. The use of building information modeling (BIM) software has begun to integrate these processes; however, for the most part, designers and engineers continue to operate within classical, architectural-versus-structural paradigms. This still un-reconciled gap between architectural and engineering modes of production calls for further research into means and methods for the unification of the design approach.

In an attempt to integrate these parallel developments, an emerging design approach uses computational building performance simulations to create a new relationship between building technology education and architectural design studio teaching. The renewed interest in building technology in general, and performance simulations in particular, sets new expectations for digitally based architectural education and practices. It sets an expectation for architecture to behave like a 21st-century structure, not merely be fashioned to look like one. Performance-based design is a particularly promising direction in regard to architectural generative processes [Aish 2005] in which a form can be not only evaluated based on the performance criteria, but also derived through the very process of simulation.

Performance-based simulation is emerging as a critical component of the contemporary design process [Kloft

2005][Oxman 2008], where it can function as a mechanism for the generative design validation. Performance-based simulations could facilitate human design by interactively responding to design parameters or function as semi-intelligent, self-optimizing agents that preselect promising generative scenarios and then channel them through a hierarchical portion of the design production, e.g. BIM software (see Figure 1). The genetic algorithm (GA) [Sasaki 2008] [Benoudjit and Coates 2008] and other evolutionary algorithms (EAs) [Burry 2004] are among the strategies that integrate structural analysis with architectural design [Schein and Tessmann 2008]. For example, Schein and Tessmann have developed a procedure for the space truss optimization based on a collision detection analysis. However, this and similar tools are still in the developmental stages and are harder to implement in a classroom context to test complex designs.

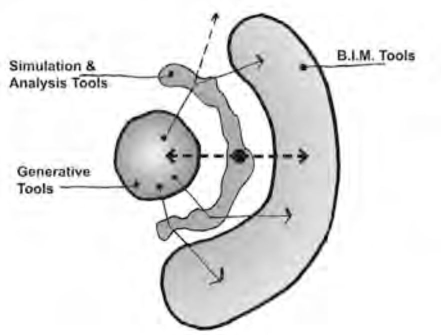


Figure 1. Simulation and analysis tools as a translation layer between generative design and BIM.

This paper focuses on the strategies for generative design validation with the use of digital simulations, particularly dynamics-based modeling tools. Specifically, tools that employ rigid/soft body dynamics such as cloth simulations, forward and inverse kinematics (FK/IK) as well as particle interactions. This approach was used in a classroom setting as an alternative, or perhaps a complement, to other methodologies such as Genetic Algorithm (GA). Our interest in this approach was dictated not only by relatively unexplored possibilities associated with this toolset, but also by its applicability as a teaching tool in an academic context.

2. DYNAMICS-BASED DESIGNS

The gap between generative design tools, which are often used to pursue exclusively formal gestures, and BIM tools is narrowing. Generative tools start considering a form's performance as well as material behaviors, while BIM tools define architecture as a parametric, spatially resolved object that can be freely manipulated and explored. This mutual convergence between generative and BIM tools is particularly effective in a scale of design components, where individual elements and properties can be parametrically interrelated. Both approaches also establish an active link between an object (component) and the entire system (whole) with an ability to manipulate individual design characteristics. While each software environment achieves this in a different way, the ability to interrelate a fragment with the entire design is common for both environments: generative dynamics and parametric BIM.

For example a rigged, IK bone system can demonstrate behavior similar to parametrically controlled composite beam-column (see Figures 2, 3). Both are defined by degrees of freedom as well as controlled by a set of constraints. While there is still a need to develop ways to effectively bridge these two digital design environments, the strategies for forming this connection emerge with parametric simulations and dynamics playing key roles. Consequently, dynamic based simulation not only create an opportunity for design validation, but also form a natural stepping stone towards parametrically defined architectural models (details) that could be utilized throughout the entire design process.

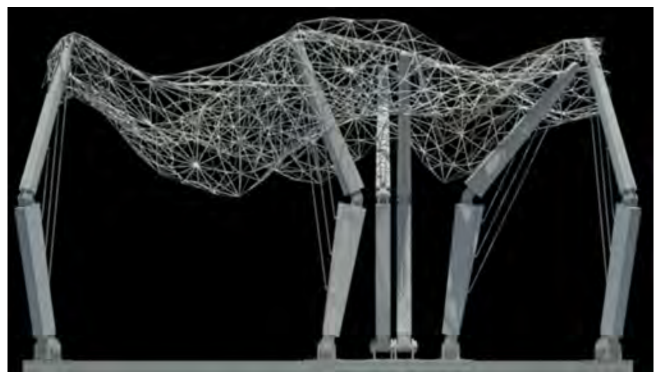


Figure 2. IK/Cloth hybrid structure after translation into BIM model with parametrically controlled columns.

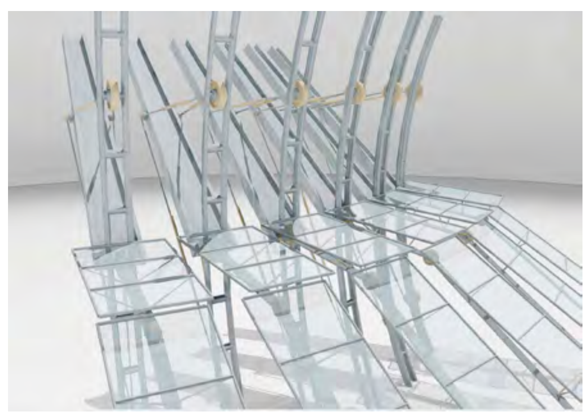


Figure 3. IK system after translation into BIM model with parametrically controlled components.

Recognizing this opportunity and testing design possibilities afforded by this approach became a central theme for a class taught by the author. Students in the class focused on traversing this 'continental divide' between generative and building modeling software with promising, yet hard earned, results. Students' work discussed later in this paper shows this convergence.

3. IN THE CLASSROOM

Special effects tools such as dynamics, cloth or inverse kinematics (IK) can facilitate form finding in a more intuitive and visually accurate way than traditional digital modeling tools. Further, this intuitive and visually accurate way is coupled with a usually instant feedback typical to dynamic simulation. This combination of increased accuracy and interactivity brings a new promise to digital design as well as to design education.

Dynamics tools such as cloth, particles or IK bring a combination of interesting characteristics together into design. On one hand, they are very suggestive, visually inspiring modeling tools that function well as generative tools. On the other hand, they start considering material and form behavior, and as such bring a component of real live performance into design. Both of these interactions are processed interactively, unlike more involved simulation tools such as Finite Element Analysis (FEA); see Figure 7.

In the class, we focused on design methodologies relating to the use dynamics-based tools. We looked at approaches that incorporated optimization and form generation mechanisms. Specifically, mechanisms that openly consider form, but also interact with simulations in a bi-directional manner. This bi-directionality becomes a vital component in the form generation feedback loop. It resembles 'the chicken and the egg' problem: one needs an idea of a form to run a simulation, and in turn, one uses simulations to derive a form. While the form finding could have been achieved in various software packages, an ability to animate transformations and interactively change design parameters was seen as crucial feature of an effective generative tool. Animation tools allow for scanning entire spectrum of possible solutions by analyzing a class of objects rather than an individual instance.

Furthermore, animating simulations puts a particular design scenario in a wider spectrum of design performance. This approach has broader design and educational benefits as discussed by Shea: "generating new forms while also having instantaneous feedback on their performance from different perspectives (space usage, structural, thermal, lighting, fabrication, etc.) would not only spark the imagination in terms of deriving new forms, but guide it towards forms that reflect rather than contradict real design constraints." [Shea 2004]

The class engaged these possibilities by employing dynamics simulation tools that are used in other industries, specifically, for the creation of special effects, gaming and character animation (see Figure 4). While this may seem as stepping outside a scientifically defined education, these tools were readily available and were well integrated within a small number of software packages. Since we had to rely on the set of software that students felt most comfortable with, as well as the need to cover a number of different simulations, we opted for the 3D Max/Maya approach with some data portability to other structural analysis software. This helped students to reduce the learning curve and optimize the software knowledge they currently held.

The following examples show specific applications of dynamics tools such as rigid/soft body dynamics, forward and inverse kinematics (FK/IK) and particle systems. While each of them represents a narrow aspect of design performance simulation, a combination of them quickly becomes a potent design tool.

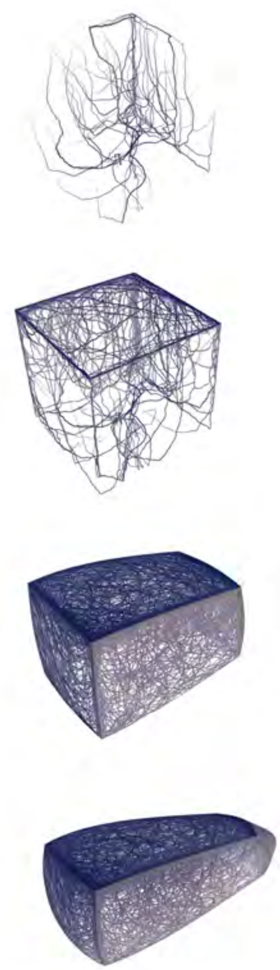


Figure 4. Generative form-finding. Figure above shows a semi-autonomous "vine" negotiating its growth in the relationship to continuously morphing form.

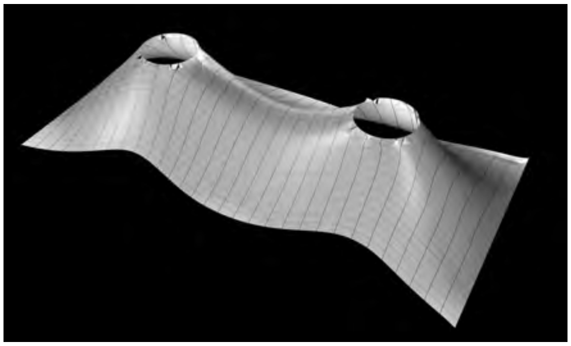


Figure 5. Cloth simulation example.

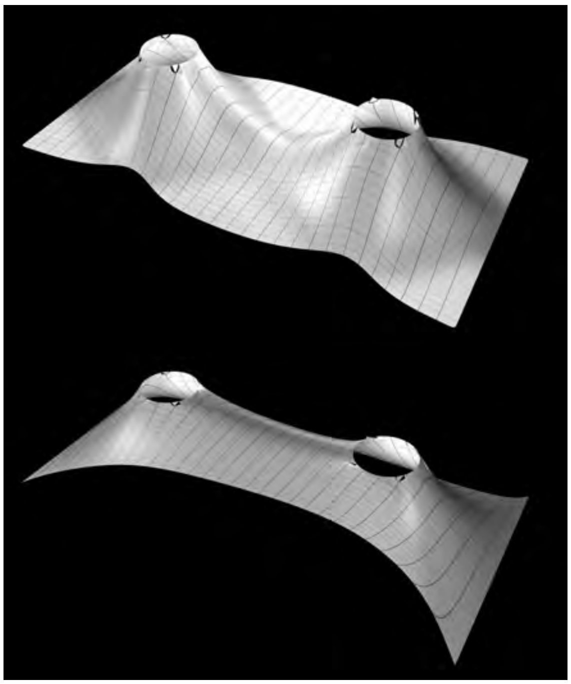


Figure 6. Variations in material properties result in different catenary shapes.

Cloth behavior exemplifies generative properties of performance-based simulations (see Figure 5). Cloth simulations, by the very nature of this material, follow the stress flow exactly and visualize the logic of a form (see Figures 7, 8) [Kilian 2004]. For these reasons, students were asked to develop a number of cloth simulations that would mimic a fabric-based architectural structure and purse material and geometric limits. Software packages provide a wide range of material properties such as weight, flexion, stiffness or friction (see Figure 6). They also consider physical forces including wind and gravity. In result, one not only can model a spatial configuration of the cloth object as a response to acting forces, but also include material properties allowing for tearing limits and fractures (see Figure 8). This interdependence between performance of a form and material parameters brings a certain level of reality into design discussion, even when particular units or physical values are not immediately understood by students.

Cloth dynamics-based simulations are analogous to rigid and soft body dynamics in its ability to incorporate physically driven behavior. An architecturally interesting extension of these capabilities is the ability to animate a cloth behavior with the use of colliders. Colliders in this application provide a skeleton for a canvas like membrane that has the ability to react dynamically to skeleton's reconfigurations. In such designed object, cloth becomes a dynamic skin that repositions itself based on the changed geometry of the collider framework. This can be achieved in the context of animated mesh or dynamics-based objects such as particles or bones.

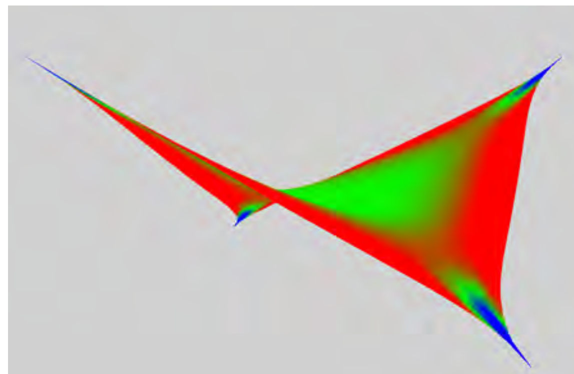


Figure 7. Cloth tension map; red color indicates fabric in tension and blue color indicates areas of compression.

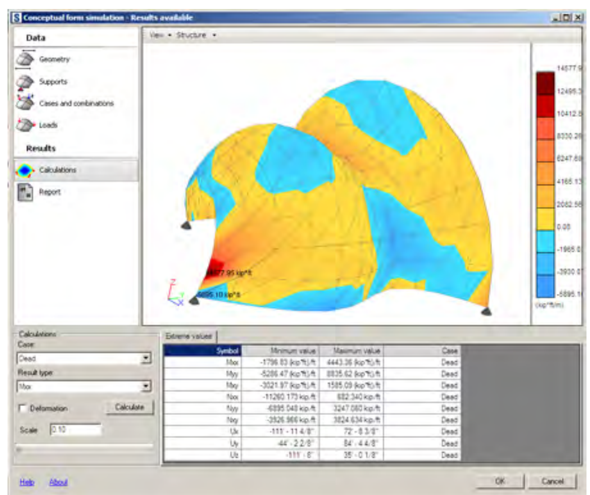


Figure 8. Similar results are usually achieved with advanced Finite Element Analysis (FEA) simulation software.

Inverse Kinematics techniques, adopted from character animation modules, were used investigate structural skeleton systems with integrated and interconnected framing members that mimicked sophisticated architectural structures (see Figures 9, 10). The ability to rig complex bone arrangements into hierarchical system with a small number of control points, allows for interactive and intuitive structural configuration. New skeletal shapes can be quickly derived from repositioning a small number of control points. After solving IK chain and hierarchical structure of the bone system, IK framework was connected with a cloth object. Resulting composite design integrated cloth with bone framework and could have been simulated dynamically as a single, morphing object.

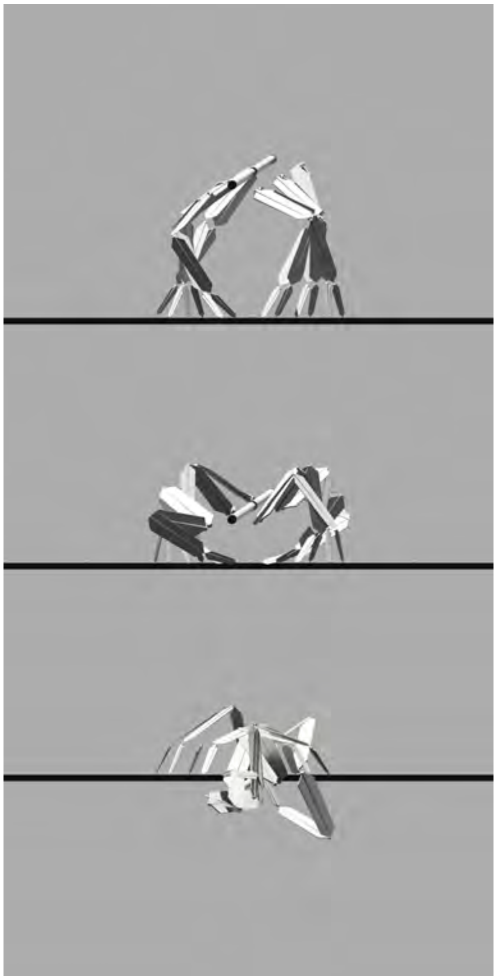


Figure 9. IK bone system helps to control structural frameworks.

While using IK in defining structural frameworks creates certain limitation in type of design solutions one is able to achieve, it also allowed students to pursue unusual and imaginary designs without need to resolve constraint requirements necessary in BIM system.

Particle systems bring yet another simulation opportunity into design. In our course, students used them to evaluate aerodynamic properties of an architectural form. This was a narrowly defined approach dictated by a wide range of various simulations they were expected to do. Other possibilities for particle system applications include aerodynamic simulations of urban spaces as well as smoke and fire spread in buildings.

The most interesting characteristics of a particle system are particles physically driven parameters. Particles can be designed to interact with other objects in a dynamic way, as well as to interact among each other. These inter-particle collisions not only allow modeling a particle system as a comprehensive force, such as wind, interacting with a building, but also within itself due to its volumetric properties. [Ophir 2008]

Figure 10. Interconnected frames mimicking adaptable structures.

4. BRIDGING THE GAP

After the initial development and simulations of generative designs students were asked to transfer them into BIM environment for further analysis. The path from generative to building modeling software was difficult and convoluted. Students often had to use other software packages to make transitions possible. This could have involved rebuilding a cloth surface in Rhino or recreating structural elements that behave like IK bones in Revit. While there are not direct and easy ways to go back-and-forth between various software packages, the process of ‘crossing the divide’ was educational and gave students better understanding of design possibilities afforded by various software packages. Additionally, by recreating IK chains in BIM software students became exposed to the logic of constraints and degrees of freedom.

Dynamic toolsets can define design in ways that would be difficult to arrive at with more traditional digital techniques such as NURB or solid modeling. This became particularly evident to students in the class who were attempting to recreate certain aspects of their IK models within BIM software. They quickly realized that using a constraint system of IK produced results faster than fully parameterized and initially less constrained BIM model/object.

Students learned from constraint and parametric models how to define parameters in a way that brings flexibility into a design system, but at the same time, define parametric flexibility that would not over-constrained their designs. Since each new parameter introduces a set of constrains (parameter range) a large number of parameters may result in increased constrains or inability to resolve them.

This parameter versus constraint relationship allows students to realize that creativity of solutions is achieved not by excessive "parameterization" of their design objects but rather by balancing parametric freedom and simplicity of an approach—structuring parameters for effective and creative use.

Dynamics–based generative models can become stepping stones for parametrically driven BIM models. This tendency can be seen in case of CS- FEM plug-in for Maya software, [Vollen et al. 2007] which is a further step towards integration of generative and validation tools within a single design environment.

5. DESIGN CASE STUDIES

In an academic environment, our primary concern is empowering students with tools and methodologies to generate creative propositions and to later evaluate them. From experience, whenever we had introduced aspects of simulations into studio design, we noticed a great deal of enthusiasm coming from students. This enthusiasm went beyond the collecting of an edgy portfolio piece and often addressed students' insecurity about the extent and applicability of their knowledge. Specifically, while doing light studies, students were often enthusiastic about visual outcomes and were ready to use them in their designs. However, these light studies often worked as yet another way to present or is some cases "sell" the project and did not necessarily help them to understand or make more informed decisions about designs.

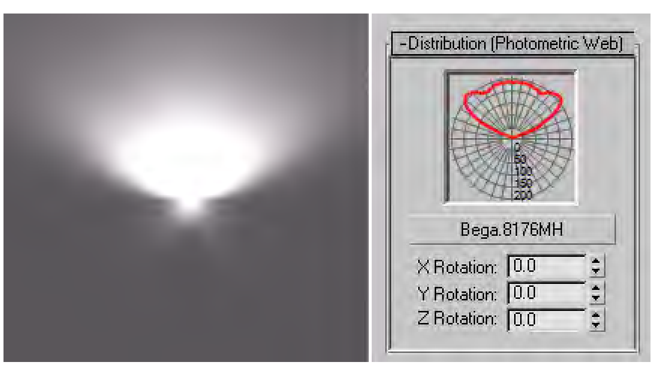


Figure 11. IES photometric light definition (right) and a rendered image showing an effect of light distribution on the wall surface.

This mindset became to change once students were introduced into physically based light analysis. Students were asked to choose lights based on the light manufacturer catalogs and the use provided by a light manufacture company photometric data (IES files) as light definitions. (see Figures 11-13). Furthermore, students were asked to render a number of views, including floor plan projection with tabulated illumination numbers in Lux or foot-candle units. This also became an opportunity to discuss various associated design issues such as levels of illumination. We went as far as discussing the color bleeding phenomenon and ways to account for it in design. Finally, students also felt empowered not only by broadening their conceptual design framework with concepts of building performance, but also, or perhaps primarily, by its scientific and tangible dimension underwritten by physically based values and behaviors.

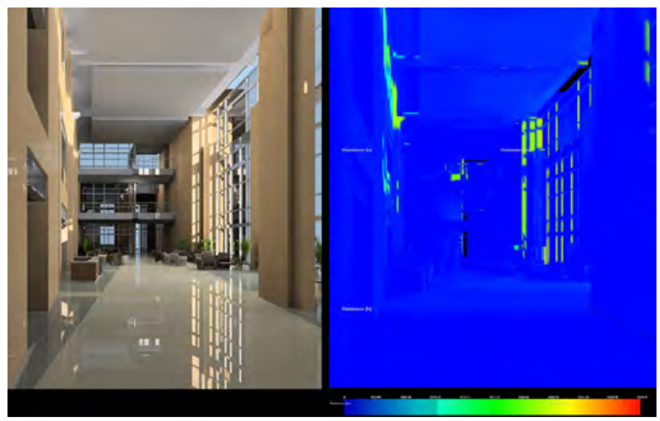


Figure 12. Sunlight illumination study.

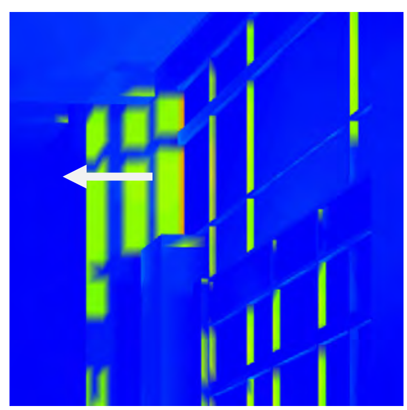


Figure 13. A detail of the illumination analysis showing changes in sunlight intensity when passing through the curtain wall glass.

This pedagogic approach builds on the notion postulated by Eduardo Torroja in 'Philosophy of Structures', where he emphasized the priority of qualitative over quantitative structural thinking. [Torroja 1958]

Computationally-based digital structural simulations address Torroja' postulate of qualitative structural thinking. They do it in a way that emphasizes a structural model with calculations being a critical determinant, but not primary visual communication component. Consequently, computer-based simulations can become a core element of structural design education by forming 'connections with ideas' [Torroja 1958] and creating opportunities for students' educational development.

Additionally, digital simulations allow students to look at more complex structural systems and to better understand their behavior. Specifically, educators can extend structural teaching models into interdependent systems that consider an entire structure. While calculations, in an architectural class context, usually stop with statically determinate structures, digital simulations can easily be extended into statically indeterminate systems such as continuous beams, at the minimum. This is an important distinction between traditional and computer assisted teaching methodologies. Traditional structural education would focus calculation-based learning on individual structural components such as a beam or a column. It would address integrated systems or complex framing in a descriptive, not computational way. Students would be told how a system would behave like, but would not be able to experience it by themselves.

Coincidentally, these complex systems need to be visualized most often because their behavior is less intuitive to students as compared to simpler models. Unlike the flexion of a beam or a column, of which a student might have had observed a similar phenomenon on his or her own in the past, complex and integrated systems typically lie beyond our immediate experience. As a result, we often calculate and experiment most with structural examples that are the easiest to experiment with, but also the least educational since they often are already intuitively understood by students. This realization is not proposing an elimination of simple model simulations, but rather argues for extending those simple models to understand them as components of a broader interdependent system.

Two critical benefits emerge from the use of digital simulations:

1) The development of an intuitive knowledge that may to some extent compensate for a lack of experience. In this meaning, intuitive (or primary process knowledge--we all have sensations before the verbalization or organization of a thought) knowledge is an unprocessed comprehension of an idea or a process that can be relied upon in preliminary decision making. This pedagogical approach responds to Michael Polanyi's "Theory of Personal Knowledge" where the author observes that knowing is an art form in which the knower understands significantly more than he or she can articulate. This comprehension of external facts without being aware of them specifically, called 'tacit knowledge', accounts for human ability to function in the world. "…tacit knowledge forms an indispensible part of all knowledge,' [Pol1983] and this is this part of knowledge, which allows us to process meaning and reach goals beyond our verbalized or processed thinking. Confidentially, what we often call experience is closely related to, such defined, tacit knowledge. This connection suggests that experience can be reinforced or partially substituted by other forms of learning. Simulations can be one of those.
2) An ability to ground a student in a physically based knowledge of architecture. In this sense, digitally based simulations relate to the teaching of materials and methods or building technology, since they bring physical properties and dimensionality to abstract designs.

6. CASE STUDY CONTRIBUTIONS

In recent years, we have witnessed a growing number of papers on the topics of generative and performance-based designs. These studies focused on theoretical underpinnings and/or relatively narrow applications that addressed particular functionalities. This study attempts to broaden this framework into multiple dynamics tools by interconnecting them into an integrated and comprehensive model. This is seen in an example that combines multiple dynamics tools, such as inverse kinematics (IK) and the cloth engine interoperability, into an architecturally relevant model.

Furthermore, this case study (student work) relates behavioral aspects of the dynamics-based tools with database models. It specifically maps individual capabilities and correspondences between both platforms and proposes a direction for further developments in the BIM platform. It shows the need for and opportunities associated with combining behavior-based and database characteristics into a single design model: broadening BIM not only as a database, but also as a behavior/performance model.

Finally, this case study allowed students to discuss an integrated design process, first by developing strategies for conceptual design and later by recreating conceptual designs within the BIM platform by mapping relationship between dynamics and BIM tools.

7. FINAL REMARKS

This paper discusses the integration of tectonic architectural studies with structural analyses and building performance simulations from an academic perspective, where the conceptual or visceral understanding of structures—a general and qualitative point of view—is more important than quantitative and numeric. The ability to visualize structural performance—stress, tension, shear—and interactively study the impact of these forces on an architectural form, can result in an integrated design process. Furthermore, these interactive simulations have an ability to translate into a visually inspired, virtual hands-on experience for students and young practitioners by helping them to develop an intuitive knowledge of architecture.

This simulation-based, interactive approach shifts the students' focus from the visualization of buildings or data to the visualization of physical processes and behaviors. The move is from static to more dynamic thinking. Consequently, through the use of dynamics-based software, a new and promising direction in generative architectural design emerges. An architectural form not only can be analyzed based on its structural performance, it can actually be derived from the process of generating structural simulations. This method of form generation brings the promise of greater design integrity within new creative horizons.

The fundamental question that needs to be further investigated is how we should relate a generative design system with performance simulations and analysis. The common-sense approach would be to combine both in the performance based generative system, not unlike discussed example of soft dynamics--cloth simulations with IK bones and their translations to parametrically-driven structural frameworks.

While it is highly desirable to create forms that immediately respond to design criteria, certain limitations may also be imposed through the introduction of too much reality, too quickly. After all, any generative system flourishes when it is not overly circumscribed by initial conditions. Success usually comes from a delicate balance of the random and the predetermined, where a small number of rules generate the greatest number of creative design solutions. Perhaps considering broader criteria in our evaluations of architectural designs would expand the range of possible design solutions.

References

Aish, R: 2005, From intuition to precision, The 23rd eCAADe conference, Lisbon, Portugal, pp. 10-14.

Benoudjit, Mohamed Amine and Coates, Paul S., Artificial networks for spatial analysis;, IJAC journal 2008

Burry, J, Felicetti, P, Tang, J, Burry, M and Xie, M: 2004, Dynamic structural modeling, The 22nd eCAADe Conference, Copenhagen, Denmark, pp. 312-317.

Kilian, A: 2004, Linking digital hanging chain models to fabrication, Proceedings of AIA/ACADIA, Cambridge and Toronto, Canada, pp.110-125.

Kloft, Harald; Non-standard Structural Design for Non-standard architecture; in Performative Architecture: beyond instrumentality; ed. Branko Kolarevic and Alim m. Malkawi; Spon Press, New York and London, 2005

Ophir, Y: 2008, Go with the flow: particle systems for program modelling in high-rise buildings, IJAC, Vol.06, pp. 172-196.

Oxman, R: 2008, Performance-based design: current practices and research issues, International journal of architectural computing, Vol. 06, pp. 1-17

Sasaki, Matsuto; Morphogenesis of Flux Structure, in From Control to Design: Parametric/Algorithmic Architecture; Verb Monograph, 2008 p.106; also available at www.isozaki.co.jp

Schein, Markus and Tessmann, Oliver 2008, Structural analysis as driver in surface-based design approaches in IJAC p.19

Shea, Kristina 2004, Directed Randomness, in N. Leach, D. Turnbull and C. Williams (eds), Digital Tectonics, London: Wiley-Academy, pp.89-101

Polanyi, Michael 1983, The Tacit Dimension; Peter Smith, Gloucester, MA, p.20

Torroja, E. Philosophy of Structures. Trans. J. J. Polivka and Milos Polivka. Berkeley: U of California P, 1958. 1-3.

Vollen, J, Clifford, D, Winn, K, and Gindlesparger, M, 2007 Digital Fabric: Generating Ceramic Catenary Networks." Expanding Bodies: Proceedings of the ACADIA 2007 Conference. Ed. Brian Lilley and Philip Beesley. Halifax, Canada: Riverside Architectural Press, 2007. 48.

Exploring Parametric BIM as a Conceptual Tool for Design and Building Technology Teaching

Andrzej Zarzycki
New Jersey Institute of Technology
University Heights, Newark, NJ 07102-1982
andrzej.zarzycki@njit.edu

Keywords: construction details, building information modeling, BIM, parametric models

Abstract

This paper discusses the adoption of BIM tools as an opportunity for design generation, validation, and implementation. It specifically focuses on parametric modeling in discussing construction details, assemblies, and generative explorations in the design context. The introduction of parametric thinking into architectural design allows for understanding the interdependencies between various elements of a building assembly and improves an architect's communication with consultants and builders. It also opens doors for "What if…" speculative exploration that allow for broader questioning of design intent. This second aspect of parametric thinking encourages students to bridge technical knowledge with creativity.

1. INTRODUCTION

Computer-based tools have radically changed the focus and modes of design thinking in architecture. While often criticized for its overemphasis on formal expressions and its pursuit of the spectacular, digital creativity has matured and begun to take into account a multiplicity of design factors that define architecture. These factors relate to performance simulation and analysis, fabrication, and building information modeling (BIM).

While usually associated with the back end of the design process (implementation), building information modeling could also redefine the way design ideas are generated by bridging formal creativity with design and technological innovation. This is achieved through a close integration of generative tools with parametric capabilities and intelligent database-enriched digital objects.

Presently, BIM-based tools lack significant generative design modules and thus become peripheral within the creative process [Penttila 2007]. At the same time, general-use, generative design software lacks the database dimension and material-based knowledge associated with its digital models. Architects may be able to develop interesting designs; however, it is impossible to verify whether these designs correspond to anything physically constructable, nor can they be associated with a particular scale or with particular material characteristics. This discontinuity in the creative process between generative and implementation design stages exemplifies a significant limitation of digital tools. To bridge this gap, this paper investigates generative qualities of the BIM platform through a relatively narrow but potent set of examples of parametrically controlled constructional details. It proposes extending BIM interoperability and parametric qualities into early, generative design phases, thus introducing two-directionality to a traditional process that follows a general-to-specific way of conceptualizing.

This paper discusses the integration of building science courses with the design studio, which offers lessons that can also be applied to everyday architectural professional practice. It proposes a design methodology that starts with a construction detail, and pursues designs that naturally emerge out of the assembly of discussed components. [Wallick and Zaretsky, 2009] While this is a long-practiced approach, this study broadens this method by considering a broader set of design solutions resulting from parametric alterations and alternations of original components. The final design project emerges through a series of explorations with fragments informing the entirety of the architectural design solution: fragments that are representative of the overall design. It is conceptually and metaphorically analogous to a fractal relationship, where a component implies an overall structure.

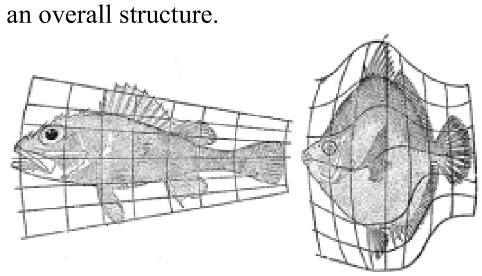


Figure 1. D'Arcy Thompson (parameter-based) transformations of biological forms [Thompson 1992]

However, this fractal-like quality is not a mathematically self-similar repetition of the same visual composition on a higher level of organization. Rather, it aligns itself with a biological analogy, where an individual component or design is altered and reused within the same organism or in its next evolutional version. Using this biological example, we can see that multiple versions of the same design, such as a leg or a skeleton, are parametric variations of the same original model. Studies by D'Arcy Thompson (Figure 1) immediately recall the phenotypical-visual transformability of distinct designs. When coupled with Dürer's proportional studies of humans, they track an analogous conceptual trajectory of phenotypical changes in an individual's development. While this progression is well demonstrated in evolutionary biology and explains the progression of species, an analogous conceptual framework is just beginning to be used in design and artistic practices. [Ambrose 2009]

2. DRIVES FOR ARCHITECTURAL IDEAS

"In conceptual art," according to Sol LeWitt, "the idea or concept is the most important aspect of the work. When an artist uses a conceptual form of art, it means that all of the planning and decisions are made beforehand and the execution is a perfunctory affair. The idea becomes a machine that makes the art." [LeWitt, 1967] Conceptual art, as defined by LeWitt, closely resembles contemporary modes of architectural practice, both digital and analog. In these modes of practice, the focus on the underlying concept often overrides all other considerations. Consequently, the design process becomes a linear implementation tool rather than an exploratory and idea-finding environment.

Conceptual art follows a didactic and hierarchical process based on design concept or idea, as opposed to an explorative and intuitive idea-building approach. The design approach inspired by conceptual art establishes the central idea as the ultimate criterion for design validation and values consistency with concept subordination over open-ended explorations with accidental discoveries.

The detail-based design approach for architecture, discussed in this paper, stands in contrast to past concept-centered design process as well as recent trends in which the weight of conceptual thinking, either in architecture or in the visual (fine) arts, has often taken precedence over tactile or material considerations. This has been evident both with traditional (analog) and with digital-based creativity. However, recent developments in fabrication, particularly in conjunction with the parametric BIM platform, create opportunities for balancing this emphasis on conceptual thinking by bringing material and assembly considerations to the forefront of architectural discourse. Architecture returns to the realm of making, rather than conceptualizing. Traditional or digital form making not only considers structural behaviors of particular geometries, as was the case with Antonio Gaudi's or Frei Otto's works [Otto 2001], but also starts considering material properties that could only be partially accounted for in Otto's soap-bubble models. Computational environments not only allow for re-addressing materiality that is often missing from the design process, but also allow for asking speculative "What if…" questions. Material properties can be parametrically investigated in similar ways to tectonics or building performance characteristics such as lighting or thermal behavior.

3. EDUCATIONAL APPROACH--METHODOLOGY

To connect generative creativity with professional practice and building technology education, the course uses BIM software. Presently, it is the only single platform that can successfully address constructability and design integration issues. However, working with BIM software has proven difficult for many designers because of the narrow range of designs that are possible with the applications. To overcome BIM's limitations as generative software, the course approach was to focus on selected software capabilities that allow for unencumbered creativity in the context of suitable design language. Thus, defining appropriate architectural precedence, in the form of case studies, was critical.

For appropriate precedence, we investigated contemporary designs representing high quality practices, which naturally translated into parametric and BIM platforms. Projects by Nicholas Grimshaw, Norman Foster, and Santiago Calatrava were just a few of the designs that fit well into the class methodology and were relatively easy to handle using digital tools. In selecting projects and particular assembly components or construction details, students were asked to study these precedences, model partial assemblies, and test them as a three-dimensional BIM models.

This assignment had two distinct purposes and phases. The first portion of the assignment—knowledge building—focused on research and modeling of an architecturally significant precedence. It let students get familiar with construction detail, assembly, and the interface between architectural and structural systems. The second portion—design formation—used the intrinsic ability of a parametric object (detail) to explore design scenarios that allowed for new design concept formation by transcending precedence into qualitatively new designs. When choosing examples for their explorations, students were asked to consider the open-endedness of their particular designs and their ability to develop meaningful variations.

Figure 2. Digital construction detail as an opportunity to research existing architectural precedence.

In this phase of the assignment, students learned about the spatial coordination of various elements and system components, their interconnectivity, and their interdependencies (Figure 2). Students were able to manipulate and experiment with parametric components and to follow interactively through the design alteration (Figure 3). Later, in the second part of the project, students explored the parametric possibilities of BIM models (Figure 4). Three-dimensional, parametrically resolved architectural details served as speculative, idea-generating devices for design. Students were expected to demonstrate the creative possibilities of their BIM models and to document their parametric explorations (Figure 5) through images, digital models, and a text narrative (final report).

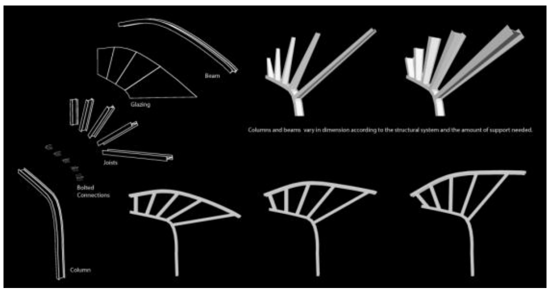


Figure 3. Digital construction detail as an opportunity to research existing architectural precedence.

The aim of this exercise was to help students to develop technical knowledge necessary for the pre-comprehensive and comprehensive studios. Specifically, it addressed the integration of building systems and their appropriateness to the design intent. Additionally, this assignment facilitated material, dimensional, and construction detail investigations in the context of contemporary architectural practice. The level of the applied constructional knowledge for this assignment matched that of the comprehensive studio work and of professional architectural practice.

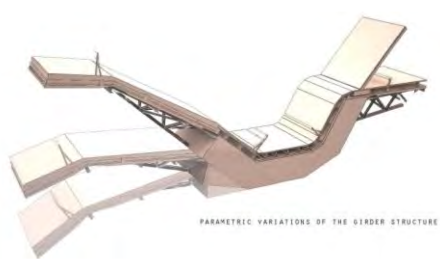


Figure 4. Parametric variations of the roof/deck structure.

4. BROADER DISCUSSION

With parametric analysis, students can immediately trace design changes and see how they impact other components in the assembly (Figure 6). Combining or nesting parametric components not only allows for an ease of modeling and a greater flexibility, but also allows understanding of how individual changes impact an overall design. Once a single parameter was changed in an overall, often complex, assembly of individual components, students were able to trace the propagation of changes throughout the database model and immediately evaluate the consequences of this particular change. Also, they could propose new designs through interactive manipulations of parameters and see changes propagated through the entire system assembly. This dual use of parametric digital models—for understanding of a significant architectural precedence (construction knowledge building) and for speculative explorations of possible design propositions—allows for greater integration between building science courses and the design studio. This is particularly applicable in the upper-level comprehensive studios where generative and implementation aspects of design need to be reconciled. In parametrically defined BIM environments, students can explore designs that are native to the world of construction—that do not have to be translated or reinvented as a result of the progression from a conceptual idea to a real product (Figure 6).

However, to be effective, this method has to approach design from a perspective characterized by inductive thinking, from particular to global, from precedence to a qualitatively new design. This reposition from didactic to inductive ways of thinking puts greater focus on design explorations and less focus on a hierarchical design process.

Figure 5. Analyzing parametrically-driven behaviors of element assemblies. Fully detailed truss at roof condition with parametric control of truss members and slab thickness.

Construction knowledge taught in architecture schools is often either irrelevant—discussing old, simplified architectural examples not related to students' current studio work—or else highly complex, representing contemporary design trends such as blobs and warps. In the latter case, it is often beyond students' ability to comprehend the information presented and apply it in their studio projects. Consequently, a certain built-in incomparability leaves students confused and less prepared for the professional life.

At the same time, the integration of building technology within upper design studios is critical. As a result of new digital tools and developments in professional practices, students increasingly develop designs that exceed their technological knowledge. This has the potential to further fragment expertise and weaken design practice by driving it toward paper-based architecture. It also has immediate implications for the education process and specifically for changes in technology teaching methods.

Parametric design follows an interesting paradox. A common argument for BIM, and for digital design in general, is that it allows for early decision making. Thus, BIM facilitates effective design progression from the conceptual to more concrete development and implementation stages. The other argument that is often put forward is that BIM allows for deferral of design decisions exactly because of its parametric properties.

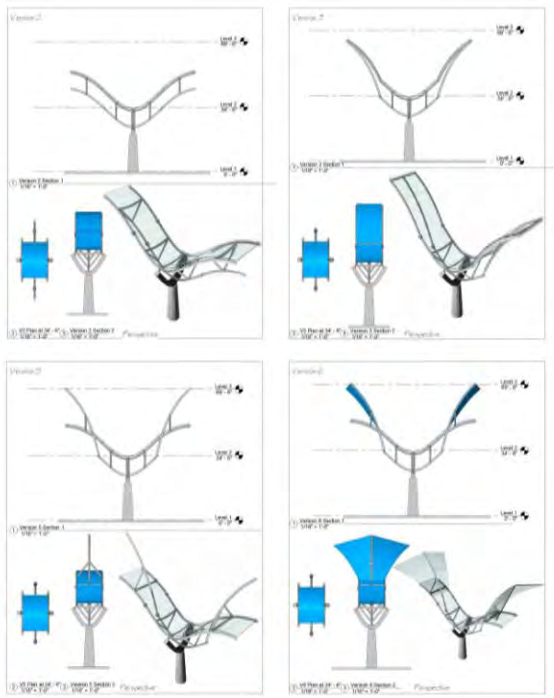


Figure 6. Parametric details allow for alternative design explorations and creating larger assemblies.

While both arguments are reasonable in their particular rationales, they also seem to exemplify both blessings and impediments to the design process. Depending on circumstances, early decision making may limit the procrastination and idle versioning common in architectural production, where a lack of direction or infinitesimal small variations in design alternatives effectively loop a designer into a closed design circle. Early decision making allows an experienced designer to validate his or her scenarios by introducing the constructability component into design.

At the same time, it is evident that the parametric capabilities of digital models allow for deferring specific design decisions while still considering a parametric component as an interdependent element of an overall system. In this application, parametric objects serve as intelligent placeholders for design. These placeholders can be changed if necessary, but, independent of the accuracy of their numeric values, they still function effectively as active elements of a larger interdependent system.

This property of parametric objects becomes a critical characteristic of BIM construction models, not only in understanding the models' assembly but also in applying them as explorative and generative tools for architectural design. This dual ability of BIM models—allowing designers to introduce constructional considerations in the early design stages, and later, due to the components' parametric definition, to develop variations and generate alternatives at the very end of the design process—reunites the act of conceptualizing with the act of making. It also renegotiates the boundary between design generation and design implementation. This renegotiated boundary will impact architectural practice and design team dynamics by increasing the requirement for each team member to contribute equally to the design and constructability of the project. Since design and implementation in BIM become more tightly intertwined, the separation into designer and detailers becomes meaningless. The next level of the design production integration removes architectural drafters from a design team structure.

5. CONCLUSIONS

The introduction of parametric thinking into building systems courses not only allows for understanding the interdependencies between various elements of a building assembly, but also opens doors for "What if...?" speculative exploration. This second aspect of parametric thinking encourages students to bridge technical knowledge with creativity.

While it is often convenient to discuss what a particular software application, tool, or methodology can or cannot accomplish, an equal burden should be placed on a user and his or her creative ability to apply these tools. The shift in tool development advocated in this paper is necessary, but an even more substantial change needs to occur in the way designers operate and conceptualize with these tools. This responsibility for growing up to match the capabilities of the tools we use is a much harder task, one that should be better handled both in professional practice and in academia. This paper naturally evolved as a response to research, teaching, and architectural practice pursuing design creativity in the context of digital, and specifically BIM, tools.

REFERENCES

Ambrose, Michael; Agent Provocateur — BIM and the Design Studio: Questioning Roles of Abstraction and Simulation in Design Education ACSA 2009 Annual Conference: The Value of Design; p.85

LeWitt, Sol "Paragraphs on Conceptual Art," Artforum, June 1967.

Otto, Frei, Rasch, Bodo; Finding Form; towards an architecture of the minimal; Edition Axel Menges, 2001

Penttila, Hannu; Dong, A, Vande Moere, A & JS Gero (eds), Early Architectural Design and BIM, CAADFutures'07, 291-302. Springer.

Wallick, Karl and Zaretsky, Michael, Fragmentation and Interrogation as an Approach to Integration; ACSA Proceedings, Value of Design, Annual Conference March 2009

Augmented Reality

DeskCube: using Physical Zones to Select and Control Combinations of 3D Navigation Operations

Michael Glueck, Sean Anderson, Azam Khan

Autodesk Research
210 King Street East,
Toronto, ON, Canada M5A 1J7
{firstname.lastname}@autodesk.com

Keywords: Augmented reality, Input device, 3D navigation.

Abstract

We present the DeskCube, a new passive input device, together with a space-division scheme using physical above-the-surface zones to select and control the desired 3D navigation operations that gives users simple scene-in-hand control over the virtual 3D world.

1. INTRODUCTION

Navigating around objects of interest in 3D environments can be difficult for novice users. In addition, most 3D software packages require two-handed input to control virtual cameras in three dimensions. The non-dominant hand uses keyboard hotkeys to switch between standard interaction modes, such as pan, zoom, and orbit. The dominant hand uses the mouse to control the selected operation. Combining multiple degrees-of-freedom (DOF) into a single input device could reduce user effort to control the viewpoint of a 3D model, but it has been found [Hinckley 1994] that users generally have difficulty in controlling so many dimensions simultaneously. To address this problem, our design includes a high DOF input device that uses physical above-the-surface zones to select and control the desired operations. We present the DeskCube, a new passive input device, together with a space-division scheme that allows camera controls to be manipulated to give users scene-in-hand control over the virtual 3D world.

2. RELATED WORK

Augmented reality has long been used to integrate the physical and digital realms. ARTag [Fiala 2005] is a library that uses computer vision to translate real-world camera coordinates to virtual camera coordinates using fiduciary markers, essentially creating a mapping from the real world to the digital world.

Several multiple degree-of-freedom devices have been proposed to help users interact with 3D scenes. Recently, Woods et al. [2003] used the ARToolkit [Kato 2003], an earlier system similar to ARTag, to create a low-cost six-degree-of-freedom "mouse" using the information extracted from the toolkit to position a camera in 3D space. This gave users a camera-in hand interaction metaphor.

Figure 1. DeskCube configuration at a demonstration booth.

3. DESIGN CONSIDERATIONS

The design of the DeskCube was inspired as a tangible analogue to the ViewCube [Khan 2008], an interface widget to aid a user in orientation awareness in a 3D scene. Each side of the DeskCube would represent one of the six canonical scene views: top, bottom, left, right, front, and back. The DeskCube is intended to rest on the desk, in front of a web camera, while not in use (see Figure 1).

When designing the DeskCube, several considerations needed to be met. The size of the DeskCube needed to be the right size to be comfortably grasped. We also discovered that the weight of the tangible device was a factor in how users manipulated early prototypes. We noticed early paper prototypes, which were lighter, would be manipulated very quickly and irregularly, and that additional weight from heavier materials, such as plastic, encouraged users to interact with the DeskCube more smoothly and consistently. Since the ARTag library produces better results when motion is less disruptive, this helped improve interaction between the DeskCube and the 3D software.

Since the DeskCube was expected to be handled by a user directly, we needed to make sure that at least one fiduciary marker was visible at all times. There is a tradeoff between the number of markers one can present on each face and the size of the markers. While Fiala [2005] recommends arrays of markers for improved performance, the smaller size of the markers also decreases correct recognition of targets, especially when they are blurred by motion. We found that four markers per face sufficiently prevented errors due to markers being occluded by a user's fingers.

4. 3D NAVIGATION SCHEMES

Three 3D navigation schemes were investigated, allowing users to access different combinations of panning, zooming, and orbiting operations. To support these operations, we partitioned the space around the DeskCube into *action zones*, allowing users to access different navigation operations by moving the DeskCube. In all schemes, a central zone was designated as a *static zone*, where no navigation operations were activated.

Two different methods of transferring user input to navigation operations were used: position control and rate control. Position control directly maps user input to the movement in the 3D scene, such as moving the mouse left to pan the scene left. Rate control, on the other hand, is a mapping where the position of the input maps to the velocity of the movement in the 3D scene, for example when pushing a joystick left to pan the scene left [Zhai 1995].

The DeskCube uses position control for orientation changing operations, such as orbit and look. As the DeskCube is rotated, the change in degrees about its axis is mapped directly to the degree change in the scene camera. Rate control is used for operations such as panning and zooming. Leaving a static zone and moving the DeskCube into an active zone acts as a clutch to engage the operation associated with that zone. The farther into the zone the DeskCube is moved the more quickly the operation is applied. The operation can be stopped by moving the DeskCube back into the static zone.

4.1. Orbiting with Panning and Zooming

In our first navigation scheme, we wanted to support all three standard 3D navigation methods: panning, zooming, and orbiting. We used the scene-in-hand [Ware and Osborne 1990] interaction metaphor to guide our design of the DeskCube navigation, meaning that if the user was viewing the DeskCube from the top-left, then they should also be viewing the 3D scene from the top-left on the monitor.

We divided the space into three primary zones, which would control zooming behavior, and subsequently divided the middle zone into nine secondary zones to control panning direction. The middle area of the primary and secondary zones was a static zone where neither zooming nor panning occurred. Rotating the DeskCube in any of these regions accessed the orbiting operation simultaneously to either zooming or panning (see Figure 2).

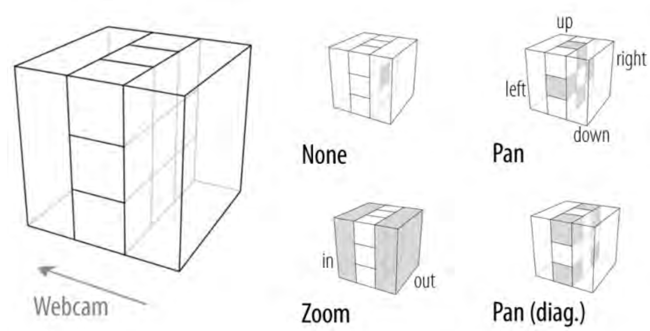


Figure 2. Schematic of modal operations for first scheme. Note that orbiting was always available in this scheme.

The main drawback to supporting all of these interaction techniques was that the DeskCube would need to be held in the static zone if the user did not want zooming or panning to occur. In order to allow the user to put the DeskCube down without effecting the camera position, we needed to provide a method of modally turning on and off the panning functionality. This was achieved through a keyboard hotkey. Unfortunately, this moved away from our original DeskCube and mouse interaction paradigm. However, we found that without some kind of modal switch, we would only be able to support a subset of these operations.

4.2. Looking with Zooming and Limited Panning / Modal Orbiting

In our second navigation scheme, we instead approached the interaction technique in terms of egocentric and exocentric usage. We defined the table top area as an exploratory zone and the area above the table as an orientation zone. The table top zone was divided into nine regions that allowed users to zoom in and out, and pan left and right, as well as a combination of zooming and panning operations. Changing the orientation of the DeskCube in any of these zones would result in a looking operation either to the left or right. If the DeskCube was lifted, only then would orbiting be available, while panning and zooming would be disabled altogether. This scheme allows users to explore a space, such as the interior of a room, by simply moving the DeskCube around the surface of their table top, and reorient themselves globally by lifting the DeskCube from the table top (see Figure 3).

While we felt this was an interesting interaction technique for the exploration of the interior of spaces, it removed the DeskCube from the same conceptual space as the ViewCube, which was designed purely for the exocentric orientation around objects from the outside.

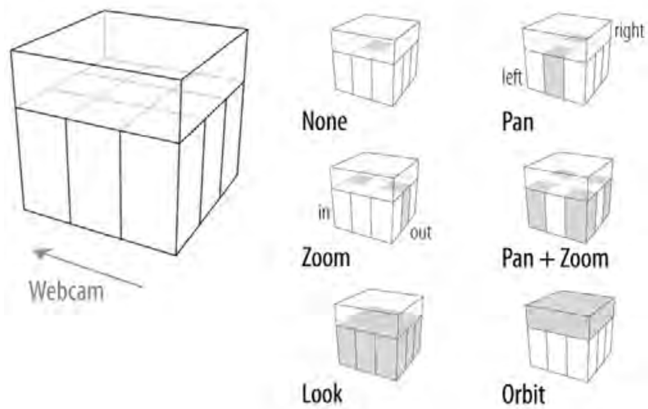


Figure 3. Schematic of modal operations for second scheme.

4.3. Orbiting with Zooming

In our third navigation scheme, we simplified the interaction technique by removing panning altogether. We also returned to the scene-in-hand metaphor. In addition, we drew inspiration from another feature of the ViewCube: *selection sensitivity*. In this case, if an object is selected, the ViewCube would represent the orientation of that object, which could be different from the global scene. This actually extends the metaphor to an object-in-hand technique. Selection and deselection of objects was performed by mouse clicks and used to navigate between selected objects. For example, in the city scene (see Figure 1), clicking on a building would reposition the camera with respect to that object. If no objects were selected, then by default all scene geometry would be selected.

When a new object was selected, the camera position would be animated toward the selected object and the camera would be turned according to the current DeskCube orientation. The camera would also be placed a fixed distance away from the object initially, as a factor of the bounding radius of the object. The user could then zoom in or zoom out from the object to achieve different views (see Figure 4).

This scheme provided the user with constrained interaction with objects found in the scene. Conceptually, it can be thought of as moving along the surface of a sphere centered on the selected object (turntable), roll, and changing the size of the sphere interactively; offering users four degrees-of-freedom.

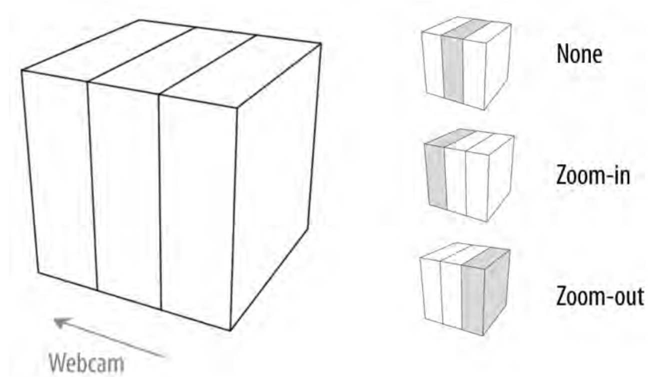


Figure 4. Schematic of modal operations for third scheme. Note that orbiting was always available in this scheme.

5. IMPLEMENTATION

The DeskCube system was implemented using the third interaction scheme. The ARTag library was used to map the DeskCube position and orientation from real-world to digital coordinated, and written as a plug-in to an existing 3D graphics engine. We used a Logitech QuickCam Pro 9000 to record the DeskCube. To ensure consistent readings, we disabled most of the automatic features (focus, white balance, exposure), in favor of a static manual calibration. The final DeskCube prototype measured 7.5cm along each side and was a hollow plastic cube with ARTag fiducial markers attached via stickers to each face.

The DeskCube and web camera system was set up on the left-hand side of the computer desk so that a right-handed user would be able to use the mouse and DeskCube simultaneously: controlling orientation and zoom with the left hand, while being able to select objects for inspection using the mouse with the right hand. To provide some stability to the zooming feature, we developed a zoned control system, whereby stickers attached to the desk would indicate tolerance zones where zooming in and zooming out would begin (see Figure 5). The zoom in region was placed 15cm away from the web camera, and the zoom out region was placed 30cm away from the web camera.

In the area between these two lines, no zooming would occur. Moving the DeskCube into the zoom-in region and zoom-out regions would perform those operations, respectively. The zooming speed was determined by the proximity of the DeskCube to the camera, increasing exponentially as it approached the webcam. Zooming was limited to the bounding radius of the object to prevent a user from entering the selected object on the near side, and a factor ten times the bounding radius on the far side.

The 3D graphics application displayed a model of downtown Toronto, Canada. Individual buildings were selectable (see Figure 1).

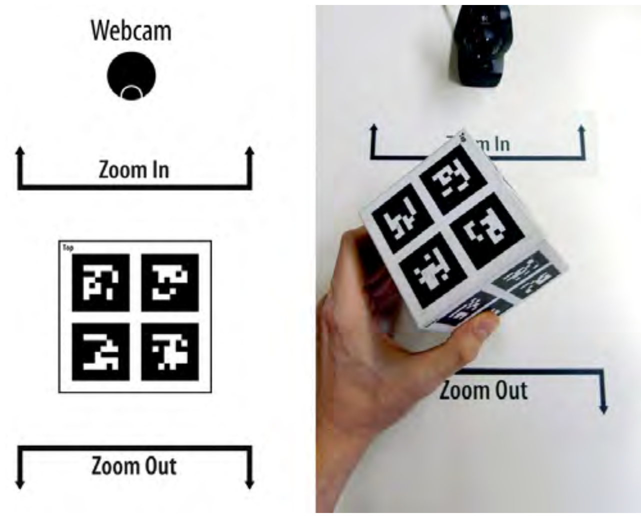


Figure 5. Layout of DeskCube regions on desk and position of webcam: schematic (left) and actual (right).

We used two methods to smooth out the DeskCube interaction: (a) averaging across samples from the ARTag library and (b) detecting when a DeskCube face was oriented orthogonally to the web camera.

Samples were read from the web camera at 30 frames per second and were averaged across the last 5 samples recorded. This was sufficient to smooth out the interaction and produced minimal lag when using the DeskCube.

An angular threshold was set for when the cube approached an orthogonal orientation. Once the cube was within 15 degrees of a face towards the web camera, the orientation of the user in the viewing application would start to pull towards the orthogonal orientation, like a magnet. This snapping behaviour allowed for more precise interaction by users to view canonical view directions, and also removed any jitter from perpendicular views of the cube, where small angular differences in samples from one web camera frame to the next could result in oscillating orientation changes in the scene.

6. DISCUSSION

A number of considerations for the design space of the DeskCube have been shown. Each one has strengths in some aspects but weaknesses in others. While we have implemented these scenarios and developed preferences for certain combinations, external testing is still needed to discover user performance and preference. To this end, we have conducted an informal pilot study to gather initial user responses and to observe the levels of difficulty in learning and using the DeskCube (see Figure 1). We installed the system at a public event held for users of 3D design software. At times we actively demonstrated the intent and use of the system while at other times we observed passers-by attempting to use the system. Approximately 200 people used the DeskCube system over a period of three days. Generally, users quickly understood the operation of the DeskCube and were able to carefully control the virtual camera. Furthermore, users were excited about this form of 3D control and several suggested that the system would be well suited to the presentation of 3D content for their clients. We were told that the use case of the DeskCube as a navigation aid while giving a presentation was appealing because it was easy enough to operate that the clients themselves could “take control.” As integrated webcams in laptops become ubiquitous, the DeskCube would be readily usable without additional equipment set-up. An architect could present a concept model of a building to a client, and use the DeskCube to offer different exterior views of the structure, without using the mouse or keyboard. This could facilitate a more natural presentation style and allow the architect to engage the client directly, without the distraction of interacting directly with the computer.

7. FUTURE WORK & CONCLUSION

Future research will focus on the design of a formal experiment to evaluate the design space of the DeskCube. We expect that further insights will be discovered leading to both improvements and an expanded design space. Also, a pilot could be undertaken to explore the DeskCube as a 3D presentation controller. In summary, we have shown the DeskCube, a new interaction device and space-division scheme to give users simple control over an urban 3D scene.

References

Fiala, M. 2005. ARTag, a fiducial marker system using digital techniques. In Proc. CVPR’05, Vol. 2, 590-596.

Hinckley, K., Pausch, R., Goble, J.C., Kassell, N.F. 1994. A survey of design issues in spatial input. ACM UIST, 213- 222.

Kato, H., Billinghurst, M., Poupyrev, I., Imamoto, K., Tachibana, K. 2000. Virtual object manipulation on a table-top AR environment. In Proc. ISAR’00, 111-119.

Khan, A., Mordatch, I., Fitzmaurice, G., Matejka, J., and Kurtenbach, G. 2008. ViewCube: a 3D orientation indicator and controller. In Proc. I3D’08, 17-25.

Ware, C., and Osbourne, S. 1990. Exploration and virtual camera control in virtual three dimensional environments. In Proc. I3D’90. 175-183.

Woods, E., Mason, P., and Billinghurst, M. 2003. MagicMouse: an inexpensive 6-degree-of-freedom mouse. In Proc.. GRAPHITE '03, 285-286.

Zhai, S. 1995. Human performance in six degree of freedom input control. Ph.D. Thesis, University of Toronto.

Input Devices for Interactive Architectural Visualization

Ultan Byrne
John. H. Daniels Faculty of Architecture, Landscape, & Design
University of Toronto
ultan.byrne@utoronto.ca

Tom Bessai
John. H. Daniels Faculty of Architecture, Landscape, & Design
University of Toronto
tom.bessai@daniels.utoronto.ca

Abstract

Opportunities for the visualization of architectural proposals are emerging at the interface of digital and physical modelling. The course of this research so far has resulted in three explorations of these new possibilities. The three are linked by the general theme of updating digital visualizations based on a user's interactions with a physical model. Each involves a different device for the translation of the user's input into a responsive visualization system. First, the repurposing of an existing device – a Microscribe Scanning Arm – for the coordination of a digital drawing set; second, the implementation of a sliding device for the interactive digital sectioning of an architectural space; and third, the implementation of a computer vision system for interactive digital perspectives. In each of the three cases, visualizations of a digital model respond to the user's interactions with a physical correlate. This paper will briefly introduce the three approaches, explain how they function, and demonstrate their application by means of Autodesk Research's Digital 210 King Dataset [1].

1. COORDINATING A DRAWING SET

At the outset of this research it was anticipated that the key challenge would be reliably calculating the real-world position and orientation of a user manipulated device in order to update the simulation accordingly. This challenge has already been treated successfully in terms of 3D input devices capable of six degrees of freedom (6 DOF). Among such devices – a variety of which are reviewed by Shumin Zhai [2] – the most readily applicable to architectural visualisation are contact based mechanical armature devices such as Immersion's Microscribe line of digitizers [3]. While desktop based "joystick" devices would eliminate the sense of direct interaction with a physical model, free moving devices often involve costly and complex tracking techniques. Furthermore, since Frank Gehry's well publicized deployment of contact based digitizers in the 1990's, the devices are likely to be generally familiar to users in the field of architecture [4]. Microscribe's G2X is just such a contact based mechanical digitizer: as the user manipulates the stylus, positional information is derived in terms of the angles of rotation at a series of joints linked by known manufacturing constants such as the length of each 'arm' segment. This positional information is calculated relative to an origin at the base of the device. In standard use, this coordinate system is linked to the coordinate system of 3D modelling software, providing a translation of the stylus' movement in physical space (for example along the surface of some object) into a point set for the construction of a digital model.

Immersion provides MicroScribe Utility Software in order to facilitate the use of the G2X with existing CAD

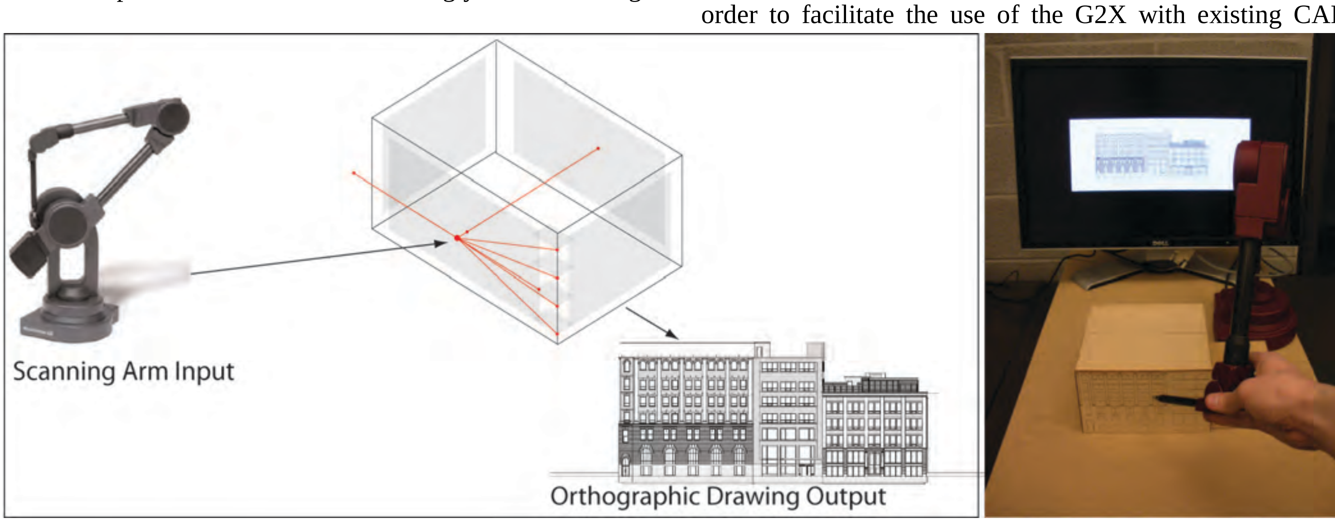


Figure 1. Coordinating an Orthographic Drawing Set with a Scanning Arm

software. The utility software communicates positional information from the G2X to an active program as a series of key presses. It is limited however, in that it only communicates the point position of the stylus tip (ie. three degrees of freedom) rather than the point position, direction vector, and rotation (six degrees of freedom). This limitation could be overcome by writing software to communicate directly with the device rather than communicating through the intermediary of the utility program.

The position information of the stylus tip was nonetheless sufficient to develop a first approach to digital/physical visualization: the coordination of a digital drawing set with a physical model. The Microscribe device was affixed to the base of an architectural model and used to take measurements of key points on the physical model. The key points included corners of the model's faces and a point at each floor of the building. Using the point measurements, software was written to calculate the distance between the position of the G2X stylus and each face of the model. The software then checks if the stylus is sufficiently close to the plane of a face or to the point indicating a floor plate. By linking a conventional drawing set of elevations and plans to this software, it is possible to coordinate an architectural drawing set which the user controls by physically interacting with a model of the building (see Figure 1). By touching the stylus to one of the model's faces or to a point at a floor plate, the user can examine that part of the building with the relevant orthographic drawing.

2. INTERACTIVE SECTIONAL CUTTING

The previous project took the continuous input of a physical device and used it to transfer between a discrete set of architectural representations. However, the formal complexities of some architectural projects resist clear representation by such means. This is particularly true in the case of contemporary architecture that eschews ordered repetition in favour of continuous variation. For example, in their documentation of the Yokohama Port Terminal, Foreign Office Architects (FOA) relied on a large sequence of sectional drawings to represent the continuous sectional variation between rooftop plaza, circulation, and programmed spaces [5]. This approach established a kind of cinematic succession of sections at equal intervals which successfully communicated the project's ambitions.

While FOA's approach makes the similarities and differences in section apparent, their representation as a series of static drawings on a page does not entirely communicate the three dimensional spatial implications of varying sections. Picking up on the approach of FOA, representing the sectional variation of an architectural space has been explored by calibrating a slider on a physical model to a sequence of digital sectional representations. Rather than a series of sections presented together, the sections become a user controlled sequence, which can be explored according to the user's interest in particular regions of the space.

Similarly, the authors of "SlideBar" describe the benefits of a specialized input device for the common class of task involving scrolling through text or sliding through menus [6]. While their objective is to develop a more specialized input device than those more familiar to computer users such as the mouse and keyboard, their SlideBar remains general in that it is meant to be used in any case of sliding or scrolling. The sectional slider device is more specialized still than SlideBar – just like the digitizing arm, it benefits most from a direct mapping between interaction and output, and hence the slider is designed according to the size of the physical model so that positioning the knob in a particular plane will activate that same plane as the sectional cut on the screen. By interacting with the slider on the physical model, the user can get a sense of the continuities and variations in the section of the space (see Figure 2).

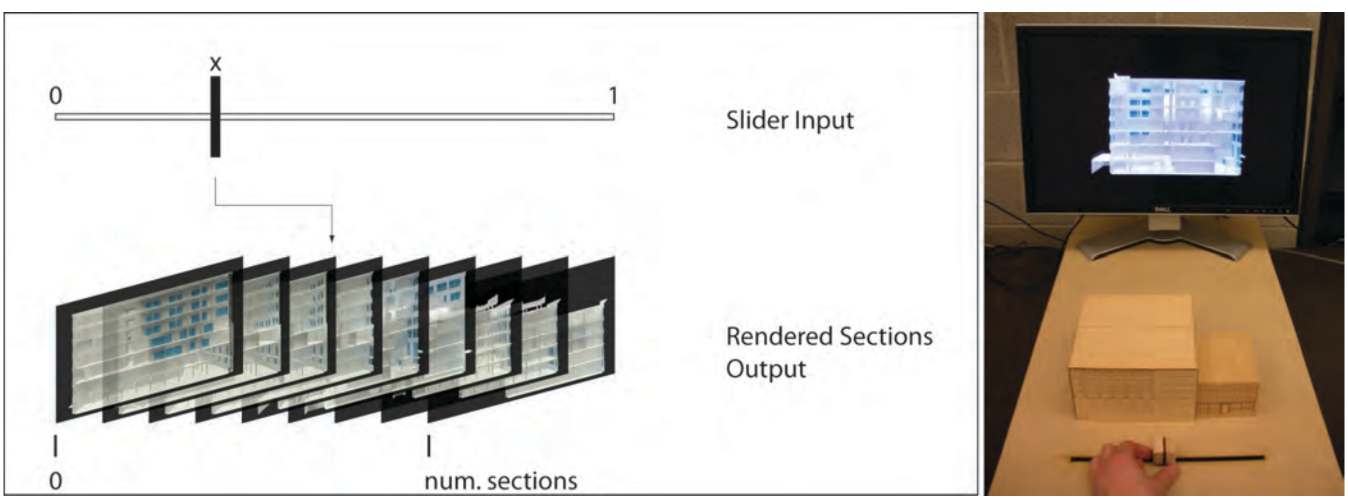


Figure 2. Controlling a Section Cutting Plane with a slider

To achieve the desired effect, a camera is first choreographed within animation software to fly through the space, perpendicular to the direction that sections will be cut. Next, images are rendered from this camera with a cutting plane activated, establishing the sequence of sectional representations. Software is then written to take the one-dimensional position input from the slider and update the screen with the appropriate section. This program requires an initial calibration for each different model or section sequence that is used. Calibration consists of partitioning the length of the slider in terms of the number of sections in the rendered sequence.

Although this second approach provides more continuous interactivity than the first, it remains highly constrained by the fact that the digital representations are themselves pre-rendered static images.

3. DYNAMIC PERSPECTIVES

The previous two projects both coordinated a pre-rendered set of representations according to the state of an input device. This approach limits the interactivity to discrete steps based on the size of the set. As the active degrees of freedom in the input device increase, the approach of discrete steps becomes increasingly inappropriate. For example, to effectively provide interactive perspectives on a space requires a device with 6 DOF. This implies either a pre-rendered set of unreasonable size or a step size between representations which is too large to convincingly provide interaction.

For the task of providing interactive perspectives on an architectural space, an alternative to the discrete pre-rendered approach is necessary. A more effective means of treating the problem is to render perspectives to the screen in real-time, providing a continuous rather than discrete domain of representations, one that is constrained in step size only by the precision of the input device.

Part of the objective of this work was to develop interactive visualisation systems which could be used without relying on expensive or complex interface devices. Hence, despite the Microscribe scanning arm's potential as a 6 DOF device and its demonstrated adaptability in the first project, it was not used as the input device. Another approach to translating real-world spatial positions into digital coordinates involves computer vision based systems. In particular, research in augmented reality has successfully treated the problem of translating spatial coordinates between virtual and real space [7, 8]. With the ARToolKit, it is possible to derive the relative positions and orientations of a webcam and a fiducial marker. In augmented reality applications, these measurements are used to provide a digital overlay for the camera's image that is calibrated to the position of the marker. Projects such as the "Magic Mouse" invert the typical arrangement of augmented reality applications using the position of the marker relative to the camera, rather than that of the camera relative to the marker. The "Magic Mouse" uses the positional data of the marker (mounted onto a glove worn by the user) to provide an inexpensive and easy to use 6 DOF input device [9].

Building on the approach of the "Magic Mouse", in this project the camera's position is fixed above the physical model and directed downwards. As the user moves the marker around within the camera's range, the ARToolKit software keeps track of the marker's position and orientation. This position and orientation are then communicated to the rendering system, which updates a view relative to a digital correlate of the camera's position. In short, the view displayed on screen at any time is from the position and direction of the marker-device held by the user (see Figure 3).

A limitation that has been identified so far with this approach involves the restrictions on the complexity of digital models that can be effectively visualised in real-time:

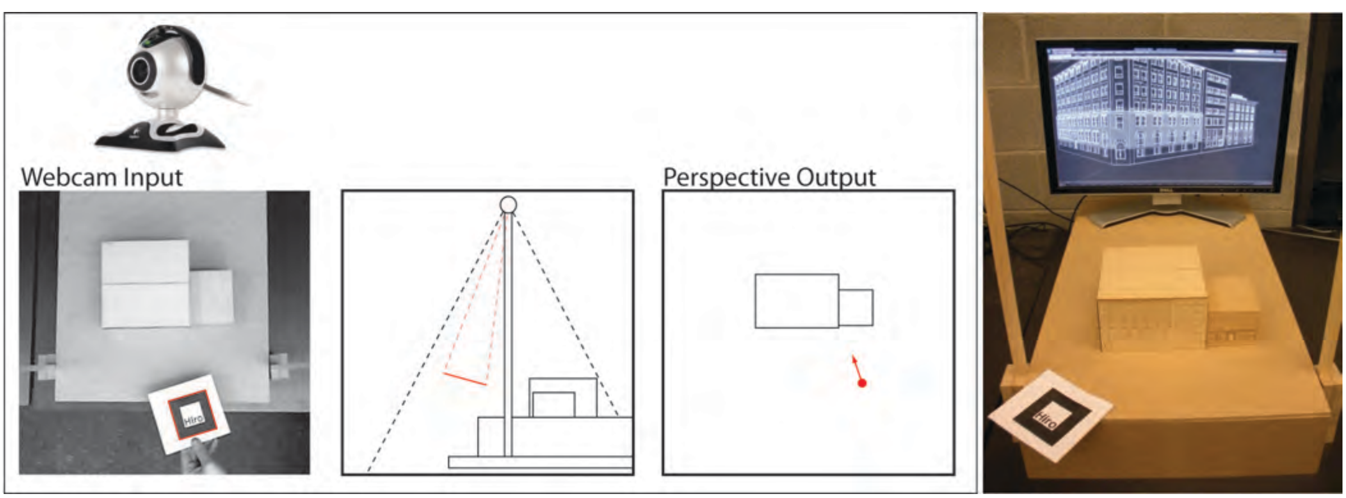


Figure 3. Real-Time Perspectives with Fiducial Marker Tracking

the high degree of detail of the Digital 210 King model makes it a very poor candidate for demonstrating the application of this approach. Using the built-in viewport rendering system of animation software, it is not possible to achieve reasonable frame-rates with such a complex model.

4. FUTURE WORK

The potential for this last research project to be extended is quite rich: a robust real-time rendering system adapted from another domain - for example a computer game's rendering engine – combined with a reasonably detailed digital architectural model could provide high quality visualisations in real-time, controllable through the user's manipulation of a fiducial marker. Because this system uses a real-time approach, it could be further extended by introducing time based variation into the visualisations. For example, variations in weather, time of day, or even entire systems of simulated users occupying the space would be possible.

Joining the three projects introduced above: real- time perspectives on a building's exterior, sectional views of its interior, and an architectural drawing set, into a bimanual interactive physical model would result in a very useful tool for communicating the intentions and specificities of an architectural proposal. It is likely that such a tool would prove most useful as a means of communicating with clients, who may be less adept at reconstructing a design from a static model and conventional drawing panels.

On a more ambitious scale, interactive methods of architectural representation could be integrated into the emerging approach of "virtual building", described by researchers at Ove Arup as the integration of 3D Models from all design consultants to coordinate, overlay, and compare information prior to actual construction [10]. Calibrating the layers of digital representations to a physical model would facilitate communication between architects, engineers, contractors, and clients, by providing a general and intuitive means of interfacing with Building Information Models.

5. CONCLUSIONS

As the fidelity of digital models to built projects improves, new opportunities are emerging for communicating architectural proposals. Reified in conventional drawings, perspectives, and fly-through animations, much of the communicative potential of digital models is lost. The real potential embedded in digital models is to provide an interactive experience determined by the user. Yet, the assumption cannot be made that the most desirable user with whom to communicate – a potential client – is able to interactively experience digital models within existing design software. Furthermore, existing input devices and software privilege individual rather than collaborative interactions. Hence, this paper introduces three projects, each of which embeds aspects of digital models into a physical model, and provides means for interaction through an intuitive input device.

A common platitude within architectural schools relates that when a physical model is presented alongside panels of 2D representations, the model will always draw far more attention and inevitably be the representation through which the project is understood. Strategies for linking digital representations to physical models are therefore likely to become ever more relevant as the sophistication of the former continues to increase, while the legibility of the latter remains primary.

References

[1] Digital 210 King Project Homepage: http://www.digital210king.org/

[2] "User Performance in Relation to 3D Input Device Design" Shumin Zhai. *Computer Graphics 32(4)* November 1998. pp50-54.

[3] Microscribe Digitizers have been discontinued by Immersion, who's homepage is: http://www.immersion.com/

[4] "Evolution of The Digital Design Process" Jim Glymph. *Architecture in the Digital Age. Ed.* Branko Kolarevic. Spon Press: 2003.

[5] *The Yokohama Project: Foreign Office Architects* ed. Albert Ferre et al. Actar: 2003.

[6] "Slidbar: Analysis of a Linear Input Device" Chipman, Bederson, Goldbeck *Behaviour & Information Technology 23.* January 2004. pp. 1-9.

[7] ARToolkit homepage: http://www.hitl.washington.edu/artoolkit/

[8] "Virtual Object Manipulation on a Table-Top AR Environment" Kato, Billinghurst et al. *Proceedings of the International Symposium on Augmented Reality*, pp.111-119. Munich, Germany. 2000.

[9] "Magic Mouse: An Inexpensive 6-Degree of-Freedom Mouse" Woods, Mason, Billinghurst. *Proceedings of Graphite* February 2003.

[10] "The Virtual Building" Bailey, Brodkin, Hainsworth et al. *The Arup Journal 2/2008.* pp.15-25.

Augmented Reality Framework supporting Conceptual Urban Planning and enhancing the Awareness for Environmental Impact

Holger Graf[1], Pedro Santos[1], André Stork[2]

Fraunhofer Institute for Computer Graphics[1]
Fraunhoferstrasse 5
D-64283 Darmstadt
{holger.graf|pedro.santos}@igd.fraunhofer.de

TU-Darmstadt[2]
Fraunhoferstrasse 5, D-64283 Darmstadt
(andre.stork@igd.fraunhofer.de)

ABSTRACT

This paper introduces a new augmented reality framework based on a multi-source urban planning backbone aiming at interactively investigating fast "what-if" analysis of urban planning simulations and creating awareness of possible environmental impact. The process of advanced urban planning nowadays includes the simulation of physical phenomena, its analysis, visualization and interpretation in order to evaluate the impact on the layout of the planning. For example, noise and air pollution, annoyances due to both nearby transportation infrastructure and urban traffic have become a serious concern for citizens. In order to provide major aid to the involved stakeholders, especially city managers, new techniques for the preparation, representation and interpretation of the typically large amount of resulting simulation data are required. These must be designed in order to enhance the perceptual and cognitive processes of users to facilitate faster interpretation and decision making. Hence, we introduce a new augmented reality framework which not only allows users to visualize but also to analyze physical phenomena in fast "what-if"-scenarios. By changing boundary conditions, parameters and re-simulating at interactive rates results can be augmented into real world planning layouts. Augmenting reality, urban planning and layouts with resulting simulation data through real-time visualizing tools provides a new and efficient interactive post processing unit for the exploration and analysis of the environmental impact due to changing conditions.

Keywords: Decision Support, Augmented Reality Framework, Multisource nD-modeling

1 INTRODUCTION

Numerical simulations have become crucial during the product development process for predicting different properties of new products as well as the simulation of various kinds of natural phenomena. Many applications of such simulations are found in different natural sciences, engineering domain and increasingly within the AEC domain. By using simulations in advance, sustainable validations and optimizations of a product's behavior or a planning site and its characteristics are possible. Nevertheless, simulation tools produce a vast amount of data taking up to several days to process, depending on the size of the mock ups, and therefore need dedicated tools to interpret analysis simulation results. Having a multi-resolution, three-dimensional model of cities into which vast quantities of data can be embedded is for sure one key of the problem. Being able to visualize this three-dimensional model, interact with the embedded data and use it efficiently to support decision making is definitely another. Some VR-based tools for interactive data exploration have matured resulting in faster decision making processes and enhancing the analysis of spatiotemporal relationships. The major aim is to enable engineers steering scientific discovery processes and investigate the impact of conceptual changes to the underlying domain or its near-by environment. These interactive visualization tools claim to enhance perceptual and cognitive processes of the user, leading to faster interpretation of results. One main problem, in interpreting such volumetric data, however, is the loss of relationship to the real environment, for which the data sets were originally computed. Here some innovative technologies based on augmented reality (AR) have been proposed to support this and over the last few years, several efforts have been spent on leveraging AR technologies into industrial working environments.

This paper presents an innovative AR framework for the interpretation of numerical simulation results within a conceptual urban planning context. It allows showing a direct impact of the physical phenomena on the physical urban planning objects and its environment. A tight coupling to a nD-Model database allows an enrichment of the real world planning through additional simulated information of interest. It is based on a new service oriented framework providing several resources needed for such analysis. Our framework has been designed such that:

- it provides a new simulation service offering the needed functionality to change boundary conditions or redefine simulation parameters

- a new AR environment to conceptually explore parameter changes within a real planning environment comprising tracking, visualization and simulation service
- an integration into a new backbone offering a diversity of services for several stakeholders involved in planning decisions thus offering potential extensions as well as the exploitation of nD-models

2 RELATED WORK

Currently most research in AR focuses on the challenges of high quality rendering, using advanced scene graph technology in combination with fast graphics accelerators (e.g. for occlusion calculations of real/virtual objects) and tracking technology for mobile applications. The registration of position and orientation of real objects, especially the user's head is under development making use of several tracking approaches (e.g. [1] or [24]). Markerless methods did achieve impressive results (e.g. [2]). Hence, it is obvious that major application areas of augmented reality can be found in the area of medicine, architecture, edutainment or cultural heritage. An overview of potential applications has been investigated in [4], [5], [6], [7]. Within industrial applications AR is traditionally placed in the late stages of the product development process, e.g. training or maintenance, e.g. [8]. Here, prepared user instructions (textual or graphical information) are displayed into the field of view of the user. Different tracking mechanisms are used to ensure an exact positioning of the displayed information. See-through head mounted display or video mixing techniques are used for the combination of real and virtual objects. Examples of the successful deployment of augmented reality in later stages of the product development process have been shown within [9]. For the use of AR during earlier stages of the product development, [10] presented a concept for the efficient presentation of product design, called Fata Morgana. In the field of architecture, engineering and construction (AEC) a growing number of technical feasibility studies of AR have shown the potential of AR during planning and construction phases. [11] establish an animated AR prototype designed to simulate activities in outdoor locations within construction operations. [12] presented an outdoor AR system that augments virtual objects of subsurface utility systems, such as buried pipes and cables, onto the real outdoor environment. [13] created an AR prototype for architectural assembly that provides users with AR guidance for assembling a space-frame structure. [14] identified AR as a technology with a high potential for coordination, interpretation and communication within certain construction, building and inspection tasks. [15] used AR for rapid assessment of earthquake-induced building damage. Their study, nevertheless, questions the lack of validation and its proof of suitability for the AEC sector. A good overview if existing AR approaches in the construction area can be found in [16]. For conceptual planning studies, several work has been published, i.e. in late 90's the Urp system [25], an AR workbench for urban planning [26] or recent efforts for AR support in early phases of design processes [27], [28]. A few approaches exist in order to integrate scientific visualization results at interactive rates into an AR environment, e.g. [20] makes use of Augmented Reality to lead 'on-site' inspections of simulation results within an airplane cabin. [21] uses a client server architecture in order to preprocess and render scientific data sets.

3 AR FRAMEWORK FOR CONCEPTUAL URBAN PLANNING AND SIMULATION

Conceptual simulation is a typical mantra within the scientific discovery process. Scientists are likely to use such thought experiments or "what if" reasoning, when it is either impossible or impractical to conduct a physical experiment. In addition, "what if" reasoning offers several advantages over quantitative reasoning strategies. These kinds of mental simulations do not require any numerical precision. This may be useful when precise quantitative information is not available or when a scientist attempts to develop a general, or high-level, understanding of a system [23]. Within a typical urban planning framework several applications and their corresponding services exist, such as applications for land use management, job services and consultation, etc. In this paper, however, we focus on the visualization of physical phenomena through simulation services evaluating the effect of the environmental behavior triggered by some pollution sources which can be placed in different locations and analyzed in view of different wind directions causing different pollutant diffusion. In this section we present the general service architecture - the Integrated Open City Platform (IOSCP)[1] (see Fig 1) as a middleware providing distributed services to client applications such as our AR enriched visualization tools. Within this platform services can be Industry Foundation Classes (IFC) databases or knowledge- and ontology-based extensions of the IFC model, such as nD-Model driven assessment services providing not only raw building data but also a variety of analytic capabilities. But services may also use other low-level services providing basic functionalities, e.g. a visualization service may connect to an interaction service.

3.1 IOSCP Service infrastructure

The IOSCP is a distributed communication and session management middleware based on CORBA providing interoperability over different hardware platforms. City

[1] http://www.c-s.fr/

services are distributed over the network. As a matter of fact, end users can access IOSCP applications through User Interfaces running on web browsers, Wi-Fi enabled PDAs, Smart-card terminals, etc. Also, the services themselves can make use of distribution (e.g. parallel super computers). The key principle used by IOSCP is the proxy concept. Once an application requests a service, it addresses an IOSCP naming server which instantiates a connection to a proxy object of the requested service. By doing so, we ensure monitored connections between client applications and services (see Fig.2).

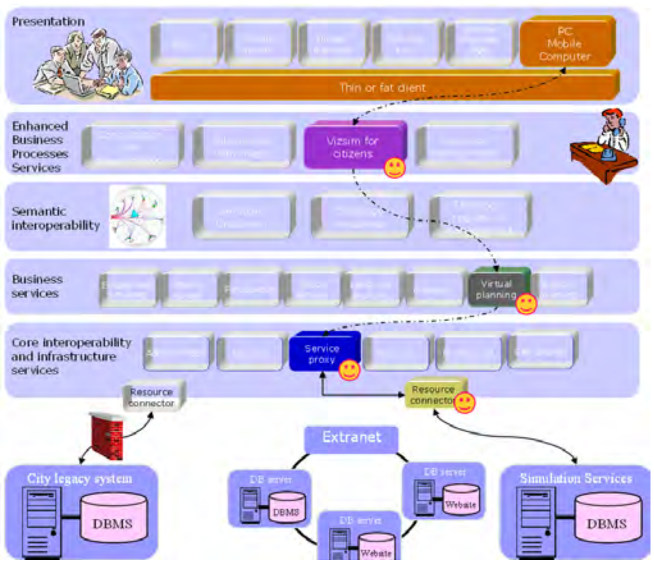


Figure 1 Urban Planning Platform (IOSCP)

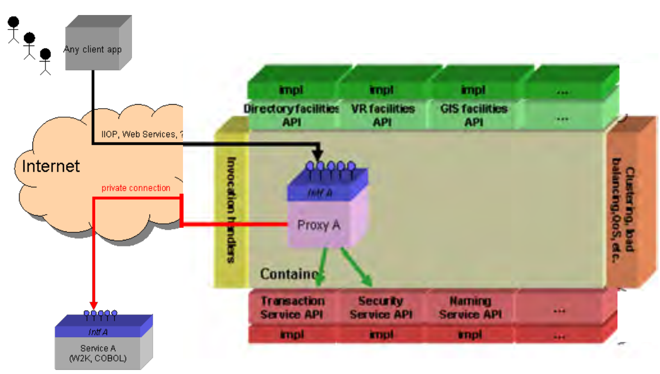


Figure 2 IOSCP Proxy Container Architecture

In addition this approach allows for scalability and extensibility of the urban planning framework as well as physical distribution of processing load.

3.2 nD-Modeling Service

An nD-model is an extension of the building information model (BIM) by incorporating all the design information required at each stage of the lifecycle of a building facility [17]. The nD-modeling approach involves the development of a holistic multi-dimensional computer model and tools using IFCs, to improve the decision-making process and project performance by enabling true 'what-if' analysis to be undertaken and to simulate and visualize the whole-life of the project. Decision making processes could then be improved in order to predict and plan construction processes, determine cost options, maximize sustainability, investigate energy requirements, examine people's accessibility, determine maintenance needs, incorporate crime deterrent features, examine the building's impact on the environment. The conceptual layout for the nD-model is given in figure 3:

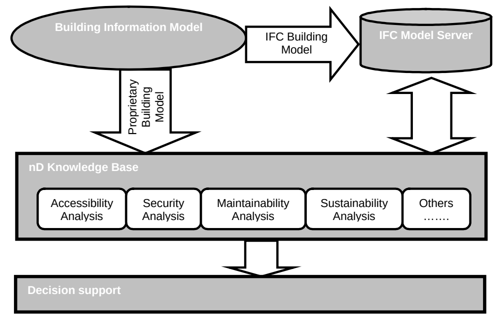


Figure 3 nD-Modelling Approach [18]

3.3 Middleware based Visualization Pipeline

Our visualization tool has been developed for interactively exploring data sets in virtual and augmented reality environments. Similar to known data-flow systems as AVS and Iris Explorer our tool is based on a modular structure enabling the creation and execution of visualization pipelines as described in [19]. The principal architecture of our system and its components is shown in Figure 5. For achieving interactive update rates the system has been designed in order to separate preprocessing and rendering. To obtain renderable images from raw IFC/simulation data, it has to be prepared accordingly in order to reduce the complexity of the data sets. This process is usually done within a pipeline of preparation and transformation steps, which transform raw data (containing 2D and 3D information) into 2D renderable images [22]. Before a mapping of 3D geometrical data into 2D pixels can be done, the data has to be filtered, classified, reduced and optimized according to the requirements of the rendering engine. For interactive exploration, those steps might have to be processed several times and impose a critical bottleneck for highly interactive environments aiming at an image refresh rate of 10-30 fps. Hence, we separated the preparation of several data sources (IFC/simulation data) and their rendering into different services. Our architecture is inspired by the observation that compared to the size of the complete flow field only a relative small portion is usually displayed. The calculated geometric- and topological data and their attributes, such as interpolated scalar- and vector values (pressure, density, velocity) are transferred through the IOSCP backbone to the AR Service (see Fig 4). Tasks of this service are the creation of graphical objects (points, lines, faces) and mapping of attributes to RGB-colors.

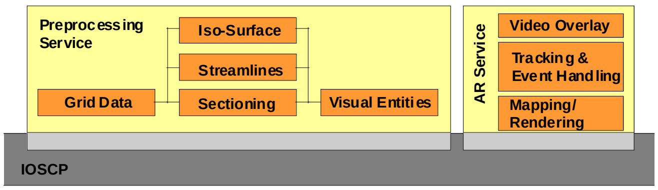


Figure 4 Middleware based integration of preprocessing and AR post processing services

Furthermore the AR service is processing user events and the incoming signals from the tracking environment as well as the capturing of the input from the video cameras.

3.4 System Architecture

To have a better understanding of how GIS/IFC/3D model data will be used to pre-calculate simulations and allow the user to then interact with the simulation results, we present the visualization pipeline implemented in more detail. The interactive visualization framework is an application composed of different IOSCP connected services (see Fig 5):

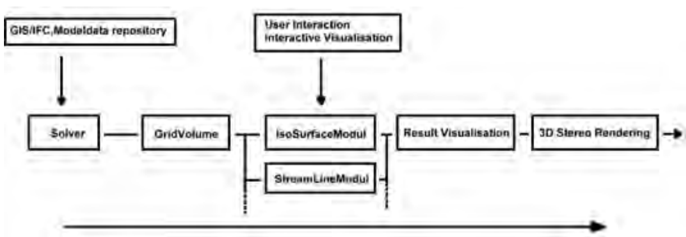


Figure 5 Visualization pipeline

Data Service

All data is stored in GIS/IFC/3D model repositories. Providing data to simulators (solvers) will be the task of an nD driven database service.

Solver Service

The solver is a module that calculates an acoustic or air flow simulation on an existing 3D city model producing a Grid Volume, where each node represents a simulation result for the respective 3D position in the volume. There may be different kinds of solvers depending on the kind of simulation to be run.

Visualization Service

Depending on the user's interaction a form of results representation is chosen. One of the many possible representations may be the display of iso-surfaces or cutting planes in a volume. Alternatively the user may interact with the simulation volume exploring streamlines of airflow simulations. The visualization service offers virtual reality stereo output on a back-projection wall as well as augmented reality stereo output seamlessly super-imposed on a real scene, visualized through head-mounted displays.

Tracking Service

To accurately compute the user's pose in any kind of environment, the tracking service connected to the IOSCP provides access to different trackers, for example optical infrared stereo tracking systems for indoor tracking or marker-less optical trackers for outdoor tracking. Available services can be written in many different programming languages as long as they are able to communicate with the IOSCP via CORBA or WebServices.

3.5 System Set-Up

A possible physical indoor system setup may consist of a video see-through HMD and an infra-red optical stereo tracking camera system with tracked interaction artifacts to analyze the pollution distribution on a model of a city square (see Fig.6).

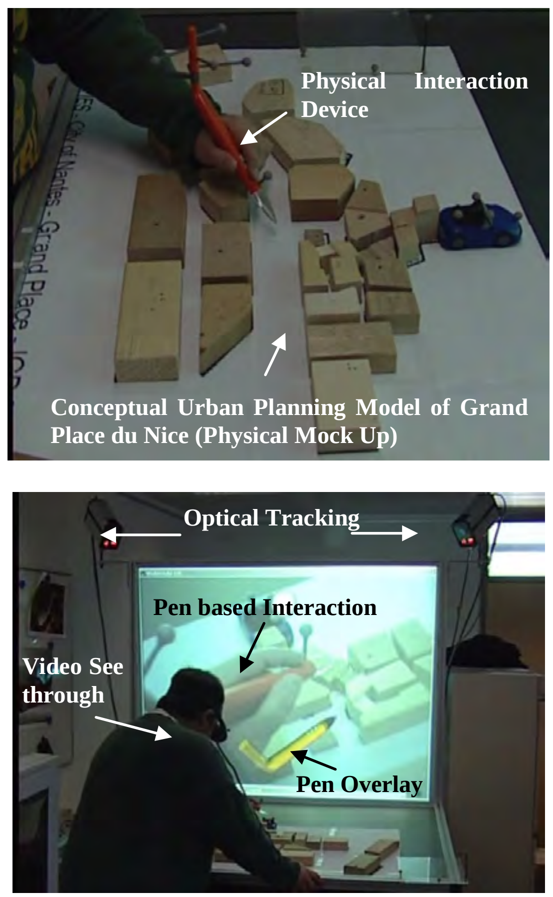


Figure 6 AR framework setup and interaction

The AR/VR interface is based on OpenInventor[2] and supports two-handed interaction by using a tracked personal interaction panel (PIP) in the one hand and a tracked pen in the other. The interaction panel is a transparent palette onto which virtual menus and controls are displayed[3]. The pen is used to point to objects and for manipulating menus and objects.

4 CONCEPTUAL FLOW FIELD EXPLORATION IN AR

The presented system allows us to review air pollution simulations and evaluate its impact on certain urban areas under changing conditions. Different pollutant distributions can be interactively placed and evaluated according to changing wind directions or locations. The quantities of interest are:

- The meteorological conditions (wind direction);
- The topography of the terrain;
- The influence of buildings.

The simulation engine which simulates the air pollution does also provide results for a more detailed dispersion of pollutants. Those can afterwards be analyzed by emitting puffs near the assigned pollutant sources. The pollutant sources in the simulation engine being used can be either vehicles or fixed sources (e.g. factories). The puff equivalent model of the sources and their propagation take into account:

- The nature of traffic (cars, trucks, speed, gasoline or diesel);
- The nature of pollutant;
- The building diffusion effects (wake, vertical surrounding).

The management process of the puff also includes the diffusion effect due to distance, the canyon effect due to narrow streets and an automatic puff partition when it is not possible to follow only one direction. Some changing boundary conditions such as traffic characteristics or wind in strength and direction can be evaluated “quasi”-interactively. Hence, the underlying simulation engines used to calculate the physical phenomena did focus on the solution of the Navier-Stokes equations (including a turbulence model) accounting for different wind directions and strengths as well as the solution of a diffusion convection problem based on the wind field for simulating the dispersion of the pollutant emission. For our AR exploration tool, we implemented different methods for visualizing the characteristics of a flow field in augmented reality environments (see color plates). As entrance point we re-designed the interaction panel (Fig. 7) and provide the users adequate access possibilities, which are used for selecting desired visualization methods, options and parameters (e.g. iso-values). Furthermore a coupling through the IOSCP provides a real-time feedback for the display. Here a change of boundary conditions such as wind direction resulted in an updated visualization and provided a new basis for the solution of the convection diffusion problem (pollution dispersion).

Figure 7 Menu for selecting different visualization methods and options as well as defining wind direction

The pollution sources for the local investigation of emitting sources at a certain location within the simulated weather condition can be set by activating the boundary condition type *Emission.* The system allows to re-define a point location and a line segment of a pollution source, at any location in the model. By moving the physical interaction device to the point of interest, a point source pollutant can be defined. Dedicated line segments provide the possibility to simulate pollution on a selected lined area (such as streets) by defining two point sources and their connections. If the user wants to explore the calculated wind field, the streamline and visualization services are triggered. Using a pen like interaction device, a new seed point is transmitted to the pre-processing service as soon as the tip of the pen is placed into the flow field and a pen button is pressed at the same time. At interactive rates the new computed streamlines are transferred and displayed. Moving the pen in the flow field and keeping the pen button pressed generates sequences of streamlines (Fig. 9). Further rendering options for the streamline representations can be chosen such as line, tubes, faces, arrows. A similar concept is used for animating particle flows. After activating the animation a new particle flow is started from the current pen position whenever the user places the pen in the flow field and presses the pen button. In the same way the interactive placement of cross sections within the flow field is accomplished by deriving the required plane normal and position from the current pen position and orientation. Figure 8c shows the exploration the resulting flow field with respect to the pollution dispersion using a cross section. Finally, our application also supports the generation and visualization of iso-surfaces. Different values can be interactively defined and analyzed (Fig. 8b).

---

[2] OpenInventor (2010), http://www.vsg3d.com/

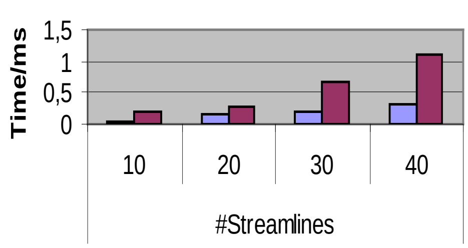
[3] Studierstube (2010), http://studierstube.icg.tu-graz.ac.at/

Several techniques presented in Figure 8 can be used to augment real physical models (Fig. 9).

5 RESULTS

The augmented reality setup we used for exploring the flow field within a conceptual urban planning set up is as follows: The user is wearing a head mounted display with two integrated micro cameras. An A.R.T. Dtrack system [24] is used for tracking the user's head-, tablet-, pen- and model positions with high quality. This system uses retro-reflective markers, which are tracked by two self-flashing infrared cameras. The video see-thru mode is accomplished by placing a video texture onto the backplane of the virtual environment. In our distributed architecture, the rendering client was running on a Core 2 Duo Windows PC with 4GB main memory, a NVIDIA PCI Express GeForce 8800 GTS Graphics board. The pre-processing service was running on a 2.33 GHz PC with 2GB main memory. For the flow computations we meshed the models using the FLUENT[4] pre-processor. The integrated fluid solver was used to create the air flow velocity field in the city simulating different weather conditions. Given this velocity field we have been able to study pollution distribution and dispersion through the solution of a convection diffusion problem. The realized solver did allow us to distribute the simulation load onto several processors providing an efficient parallelization (table 2). A specific parallelized convection diffusion solver provided direct access to the different pollution sources and its locations. Several result files have been stored into a simulation DBM which could be accessed through a simulation service interface via the IOSCP. Two models, a smaller containing 108604 nodes and 577748 tetrahedral elements ("2houses") and a larger urban model, the re-modeled 'Grand Place du Nice' ("Urba"), which had been generated using an unstructured grid containing 347819 nodes and 1632685 tetrahedral cells suited as input. The time needed to calculate different numbers of streamlines is shown below. It indicates that for fewer streamlines real-time feedback can be achieved, whereas a larger number of streamlines might significantly reduce the reactive response of the preprocessing service.

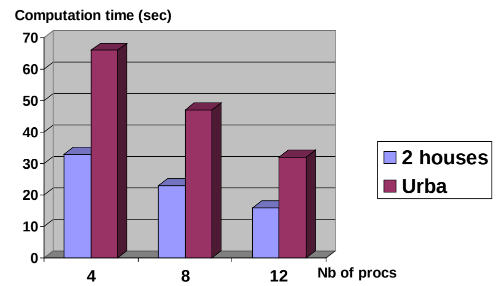
Table 1 Performance of interactive streamline calculation



[4] Since 2006 ANSYS Inc.

Table 2 shows the simulation time needed to calculate the diffusion convection problem dependent on the number of processors for two example sets. A pollutant diffusion step in the larger model could thus be simulated within 32 sec on 12 processors (It takes respectively 1 hour and 5 min using only 1 processor).

Table 2 Performance of parallel convection diffusion simulation



6 CONCLUSION & OUTLOOK

Within this paper we have presented an efficient solution for data preparation and processing of simulation results in a highly interactive 3D AR environment for conceptual investigations in urban planning. Based on an integrating planning platform with access to a nD-modeler, we could provide an efficient prototype for leading conceptual studies within urban planning set ups. The results we achieved indicate that fast "what-if" analysis might be feasible and that the framework we elaborated provides a good playground for testing alternatives. Though for even larger models it might become unfeasible, there is still a possibility to reduce the data sets by sub-modeling procedures bringing us again back to fast "what-if" interaction loops. The concept and layout of our solution in an augmented reality setup opens further possibilities to integrate ongoing research topics of physically based simulations, like interactive mesh manipulations, interactive flow field distortion and adaptation of the simulation results as well as mutual occlusion calculations. A real mixed reality setup is envisaged in which physical objects might influence turbulences, change of distribution of flow and interact with the simulation engine. As within conceptual studies the scientific precision is not needed, we envisage the inclusion of hardware accelerated physical simulation engines as e.g. used within games in order to provide a good and realistic overview of the concepts to be tested. Other research work should then address adaptive refinements in critical areas with a zoom into details with a possibility adjusting the scientific precision by intelligently changing the mathematical model.

7 REFERENCES

[1] Billinghurst, M., Poupyrev, I., Kato, H., and May, R.: 'Mixed Realities in Shared Space: An Augmented Reality Interface for Collaborative Computing.' Proceedings of ICME 2000, 2000, IEEE. pp. 1641-1644

[2] Reitmayr, G.; Drummond, T.W.: Initialisation for Visual Tracking in Urban Environments. In: Mixed and Augmented Reality, 2007. ISMAR 2007. 6th IEEE and ACM International Symposium on, 2007, S. 161 172 2.1.3

[3] Klinker, G., Creighton, O., Dutoit, A., Kobylinski, R. and C. Vilsmeier: 'Augmented maintenance of powerplants: A prototyping case study of a mobile AR system'. In IEEE and ACM International Symposium on Augmented Reality, Oct. 2001.

[4] Stricker, D., Klinker G. and Reiners D.: 'Augmented reality for exterior construction applications.' In W. Barfield and T. Caudell, editors, Fundamentals of Wearable Computers and Augmented Reality. Lawrence Erlbaum Press, 2001.

[5] Christou, I., et al.: 'Design and Development Issues for ARCHEOGUIDE: An Augmented Reality based Cultural Heritage On-site Guide', Proceedings of ICAV3D '01.

[6] Azuma, R.: 'A Survey of Augmented Reality. Presence: Teleoperators and Virtual Environments', 1997. 6(4): pp. 355-385.

[7] Azuma, R., Baillot, Y., Behringer, R., Feiner, S., Julier, S., MacIntyre B.: 'Recent Advances in Augmented Reality'. IEEE Computer Graphics and Applications, November 2001

[8] Santos, P., Gierlinger, T., Stork, A., McIntyre, D. 2007. Display and rendering technologies for virtual and mixed reality design review. In 7th International Conference on Construction Applications of Virtual Reality, CONVR 2007. Penn State University, University Park, PA, USA.

[9] Gierlinger T., Danch, D. Stork, A. 2009. Rendering techniques for mixed reality. In Journal of Real-Time Image Processing, Special Issue, Springer Berlin / Heidelberg. DOI 10.1007/s11554-009-0137-x

[10] Klinker, G., Dutoit, A., Bauer, M.: 'Fata Morgana – A Presentation System for Product Design', Proceedings of the IEEE and ACM International Symposium on Mixed and Augmented Reality, Darmstadt, 2002. pp 76 – 85.

[11] Behzadan, A., Kamat, V.R.: 'Visualization of construction graphics in outdoor augmented reality', Proc. The 2005 Winter Simulation Conference, Orlando, Florida, U.S.A, 2005, pp. 1914–1920.

[12] Webster, A., Feiner, S., MacIntyre, B., Massie, W., Krueger T., 'Augmented reality in architectural construction, inspection, and renovation', Proc. ASCE Third Congress on Computing in Civil Engineering, Anaheim, CA, 1996, pp. 17–19.

[13] Roberts, G., Evans, A., Dodson, A., Denby, B., Cooper, S., Hollands, R., 'The use of augmented reality, GPS and INS for subsurface data visualisation.' Proc. FIG XXII International Congress, TS5.13 Integration of Techniques, Washington, D.C., U.S.A, 2002.

[14] Shin, D., Dunston, P.S.: 'Identification of application areas for augmented reality in industrial construction based on the suitability', Automation in Construction 17 (7): 882–894, (2008)

[15] El-Tawil, S., Kamat, V.R., 'Rapid reconnaissance of post-disaster building damage using augmented situational visualization', Proc. 17th Analysis and Computation Specialty Conference, St. Louis, Missouri, U.S.A, 2006, pp. 1–10.

[16] Dunston, P.S., Shin D.H., 'Key areas and issues for augmented reality applications on construction sites', Purdue University, USA in Mixed Reality in Architecture, Design, and Construction, Xiangyu Wang, Marc Aurel Schnabel, 2008, Springer Verlag

[17] Lee, A., Marshall-Ponting, A.J., Aouad, G., Wu, S., Koh, I., Fu, C., Cooper, R., Betts, M., Kagioglou, M. and Fischer, M., 'Developing a vision of nD-enabled construction', Construct IT Report, ISBN: 1-900491-92-3

[18] Fu C., Aouad, G., Lee, A., Mashall-Ponting, A., Wu, S., 'IFC model viewer to support nD model application', Automation in Construction (15):178– 185, 2006, Elsevier

[19] Schroeder, W., Martin, K., Lorensen, B., 'The Visualization Toolkit', 3rd Edition, Prentice-Hall Inc, 2004. ISBN 0-13-954694-4.

[20] Regenbrecht, H., Jacobsen, S.: 'Augmentation of Volumetric Data in an Airplane Cabin', Demonstration at IEEE and ACM International Symposium on Mixed and Augmented Reality, Darmstadt, 2002.

[21] Benölken, P., Graf, H., Stork, A., 'Texture-Based Flow Visualization in Augmented and Virtual Reality Environments', Proc. of International Conference in Central Europe on Computer Graphics, Visualization and Computer Vision (WSCG), pp. 21-24, 2004

[22] Frühauf, T., "Graphisch-Interaktive Strömungs-visualisierung", Springer-Verlag, 1997

[23] Bell Trickett, S., Trafton, J.G., '"What if. . .": The Use of Conceptual Simulations in Scientific Reasoning', In Cognitive Science (31):843–875, 2007

[24] A.R.T., Advanced Realtime Tracking GmbH, Germany, http://www.ar-tracking.de.

[25] Underkoffler, J., Ishii, H., 'Urp: a luminous-tangible workbench for urban planning and design', In: Proceedings of the SIGCHI conference on Human factors in computing systems, 1999

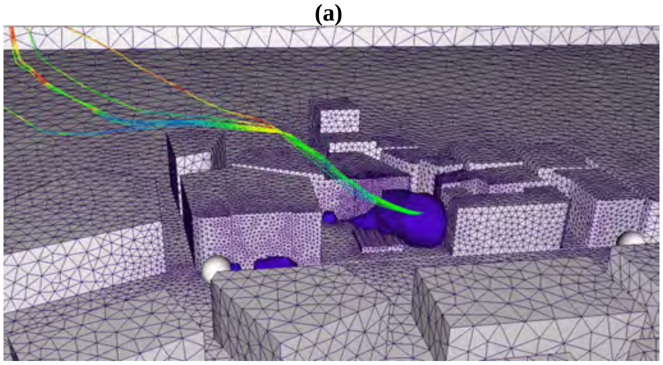
[26] Ishii, H., Underkoffler, J., Chak, D., Piper, B., 'Augmented Urban Planning Workbench: Overlaying Drawings, Physical Models and Digital Simulation', In: IEEE & ACM ISMAR 2002

[27] Aliakseyeu, D., Martens, J-B., Rauterberg, M., 'A computer support tool for the early stages of architectural design', In: Interacting with Computers, Elsevier B.V, 18:528–555, 2006

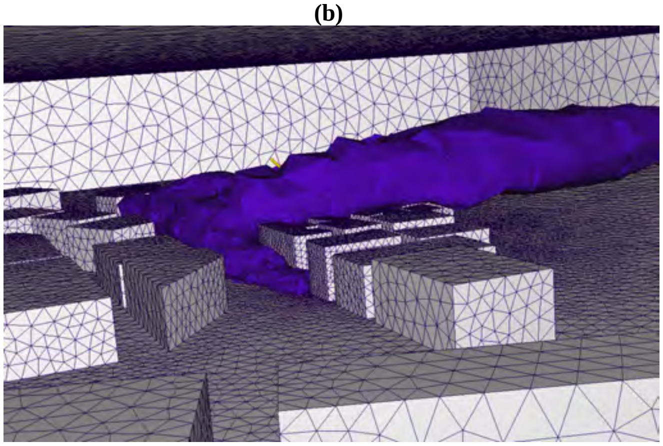
[28] Maquil, V., Sareika, M., Schmalstieg, D., Wagner, I., 'MR Tent: A Place for Co-Constructing Mixed Realities in Urban Planning' In: Proc. of the Graphics Interface Conference, BC, Canada, 2009

8 COLOR PLATES

(a)



(b)



(c)

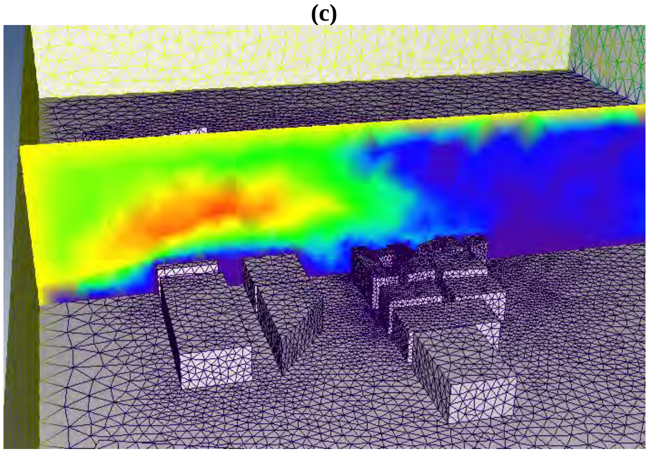
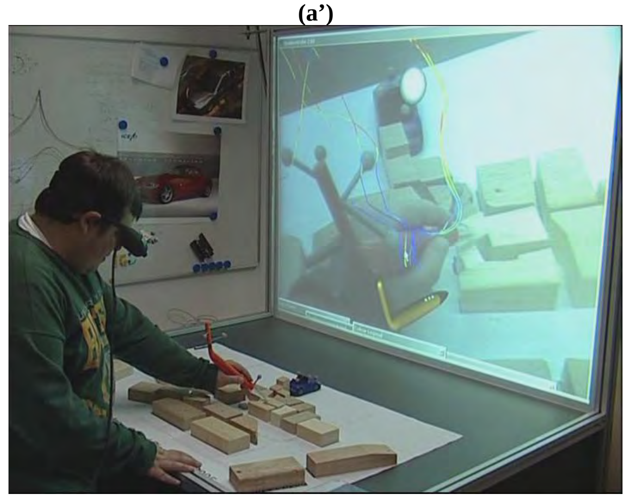


Figure 8 Scientific visualization and interactive post processing of pollution (sources iconised through white balls, streamline and iso-surface representation; conceptual analysis by changing icons and wind direction (a)), diffusion of pollution at given wind direction (iso-surface presentation over time; (b)); pollution dispersion and concentration (cross section mapping and representation, (c))

(a’)



(b’)

Figure 9 Augmented Reality setup for conceptual urban planning. Physical mock-up building blocks augmented by scientific simulation results (a’), (b’). Several post processing techniques such as mentioned in figure 9 could be augmented into the field of view of the operator, here streamline renderings.

Transportation

Supporting Outdoor Mixed Reality Applications for Architecture and Cultural Heritage

Pedro Santos, Dominik Acri, Thomas Gierlinger, Hendrik Schmedt
Fraunhofer Institute for Computer Graphics
Fraunhoferstrasse 5, D-64283 Darmstadt
({pedro.santos|dominik.acri|thomas.gierlinger|hendrik.schmedt}@igd.fraunhofer.de)

André Stork
TU-Darmstadt
Fraunhoferstrasse 5, D-64283 Darmstadt
(andre.stork@igd.fraunhofer.de)

ABSTRACT

This paper introduces new approaches to enable collaborative outdoor mixed reality design review in the architectural domain as well as outdoor mixed reality experiences in the cultural heritage domain.
For this purpose we present the results of three closely related European projects, IMPROVE and CINeSPACE, which are currently succeeded by MAXIMUS, continuing the development of technologies relevant to the two domains.
The paper focuses on the base technologies needed to develop usable outdoor mixed reality applications, such as marker-less optical tracking combined with sensor fusion for accurate pose estimation outdoors. Furthermore the paper presents a new visualization system developed within the project, one of the few daylight blocking head.-mounted displays available. In addition to pose estimation and visualization devices, the paper presents a VR framework adapted for the architecture and the cultural heritage domain featuring high dynamic range image acquisition combined with pre-computed radiance transfer rendering to be able to photo-realistically render urban content. For the cultural heritage domain this framework has been extended to allow accurate display and generation of multi-media content super-imposed on a real environment as well as city navigation.

Keywords: Outdoor Mixed Reality, Design Review, Display Technologies, Pose Estimation, Rendering, Human Computer Interaction, Cultural Heritage

1 INTRODUCTION

In the coming years we will see the rise of mixed reality applications. Mobile processing platforms are becoming faster and less power consuming. Display technologies are becoming more integrated and less intrusive. Rendering capabilities of mobile platforms are increasing. Advances in optical tracking enable accurate, real-time pose estimation on mobile devices with the precision needed to seamlessly super-impose virtual content.
Over a number of previous and ongoing projects we have developed supporting software for mixed reality environments that cover many areas of interest for the architecture and automotive domain. In this paper we present rendering components making use of High Dynamic Range image acquisition for accurate and photo-realistic rendering, new near to the eye display types which will work in outdoor environments even when pointed against the sun light and pose estimation approaches vital to any seamless super-position of virtual reality content on top of a real scene.

2 RENDERING

Realistic rendering in the mixed reality context is concerned with the seamless integration of virtual models into real-world scenes in real-time. To achieve this, the rendering system has to perform a lighting simulation to capture global effects like shadowing (i.e. shadows of virtual models must be cast into the same direction as those of real-world objects). Furthermore, the colour of virtual models due to incident lighting has to be consistent with the real-world scene. The consistent appearance of the virtual models in a real-world environment requires a physically based description of the model's materials as well as a description of the incident lighting which resembles the real-world lighting situation as closely as possible.

In our case we induce two additional requirements with respect to the rendering, namely real-time frame rates have to be achieved on a mobile computer (with rather limited resources as compared to desktop PCs) and the shifts in intensity and color due to the display's alpha channel have to be taken into account. The rendering work aims at delivering higher quality than standard rasterization-based systems which inherently use a local lighting model and cannot take global illumination effects into account. In the following we give a short overview of global illumination algorithms, followed by the state of the art in mixed reality rendering.

2.1 Mixed reality rendering systems

Traditionally, the requirement of interactive frame rates for mixed reality rendering systems has restricted such systems to the use of rasterization-based approaches which do not take global illumination into account (see e.g. Klinker et al. [Klinker et al 2002]). Advances in processor and graphics technology have later led to the adoption of global illumination algorithms. The ARIS project [ARIS 2004] calculates approximate soft-shadows by combining shadow maps of the most important emitters of a Radiosity mesh. However, good shadow quality requires a large number of shadow maps to be generated. Pomi et al. [Pomi et al 2003] utilize their distributed Ray Tracing system to render mixed reality content. They achieve high quality results, but require multiple computers to do so which prevents this approach from being used in mobile scenarios. Santos et al. [Santos et al 2007] apply PRT in the mixed reality design review system of the IMPROVE project to perform high quality rendering with soft-shadows in dynamic environmental lighting (captured by High Dynamic Range images). Franke et al. follow a similar approach in their X3D based rendering system. Current work by Gierlinger et al. [Gierlinger et al 2009] extend the IMPROVE rendering system into a hybrid renderer which fuses PRT and Ray Tracing in order to support accurate reflections and refractions while at the same time accounting for physically-based soft-shadows. This work is being performed in the context of the MAXIMUS project.

2.2 Current rendering solution

The rendering work in our projects aimed at developing a rendering engine that provides high quality real-time rendering in mobile mixed reality scenarios with support for the alpha channel of the proposed display. Due to the limited resources of mobile computers, the focus is not on basic research in the area of rendering algorithms, but on adapting existing algorithms in order to deliver interactive speed on the proposed mobile system.

2.2.1 HDR image acquisition

As discussed before, HDRI enables the efficient acquisition of real-world environmental lighting. To capture an HDRI using a standard digital camera it is necessary to take a series of photographs at different exposure times. These images are then combined to form an HDRI (see [Debevec et al 1998] for details), which can be done using HDRShop [HDRShop 2009]. The acquisition of the photo series can be done manually, but this usually leads to incorrectly registered images due to small movements of the camera when pushing the release. It is possible to register the images in a post-process (e.g. by using PFSTools [PFSTools 2009]), but an easier way to get correctly registered images is to remote-control the camera. AHDRIA [AHDRIA 2009] is a software package specifically developed for this purpose. To capture a full environment image in one shot, it is possible to utilize a mirror sphere located at the place where the virtual object will be placed. However, this approach results in rather low quality environment images due to the limited resolution provided by off-the-shelf cameras. For this reason we use a SpheroCam HDR [SPHERON 2009], which is a special high resolution HDR environment camera. The resulting images have a resolution of 11000x5500 pixels. From these images we generate a low resolution light probe image (128x128 to 512x512 pixels) which we actually use for the lighting calculation in the renderer and a high resolution Cube Map which we use for specular reflections.

2.2.2 Pre-computed radiance transfer

PRT is a real-time rendering algorithm proposed by Sloan et al.[Sloan et al 2002] to render low-frequency global illumination effects under dynamic environmental lighting. The algorithm consists of two passes: a pre-process and the actual run-time calculations. The pre-process calculates so called transfer functions at points on a model (either at the vertices or on a per-texel basis). These transfer functions encode global illumination effects like shadowing and indirect lighting and they are projected into the function space of the Spherical Harmonic (SH) functions.

Figure 1 Changing Illumination conditions on-the-fly[1]

[1] Model courtesy of Page and Park architects, Glasgow, Scotland

The projection process results in a set of projection coefficients which are stored together with the model and reused during the run-time calculation. We utilize Shirley's Galileo ray tracer [Shirley et al 2003] to evaluate the visibility function. During run-time we project a lighting environment provided as an HDR light probe image into the SH basis and upload the projection coefficients as uniform variables to the graphics board. Figure 1 shows two renderings of a building with different environmental lighting. Note the soft-shadows below the building. The lighting environment can be changed interactively.

The composition of the shadows cast by the model with the background image works as follows: first of all we have a white plane below the model that acts as the shadow receiver. For this plane we calculate the colour due to the environment lighting (without the geometry of the building casting shadows). This un-shadowed colour is calculated once per frame on the CPU since it is the same colour for all vertices of the plane (it is only dependent on the plane normal and the actual lighting environment). We then compute the shadowed colour for the plane on the GPU and derive the change in illumination as (un-shadowedColor – shadowedColor)/un-shadowedColor. The result is then multiplied with colour value of the background image.

The rendering engine is tailored to the display hardware supporting the alpha channel of the display (which will also be used to cast virtual shadows onto real-world objects). Finally, we have worked on adaptation algorithms to account for intensity / colour shifts induced by the display's alpha channel.

2.2.3 Benchmarks

In this section we present benchmarks for a typical scene rendered using our system. The model is courtesy of Page\Park Architects, Glasgow, Scotland.

Scene	Building
Number of triangles	1.5 million
Number of vertices	3.3 million
Pre-processing time preview quality	5 minutes
Pre-processing time medium quality	28 minutes
Pre-processing time high quality	106 minutes
Rendering speed	25 fps

Table 1 Benchmarks for the Building scene

The timings for pre-processing and rendering are shown in Tables 1-2. The computer used for pre-processing and rendering was an Intel Core2Duo 6850 (3 GHz) with 2GB of RAM and a Geforce GTX 280 running Windows Vista x64. The rendering was done using 9 coefficients (3 SH bands) for the representation of the environment light and the transfer functions. The transfer coefficients were calculated in the pre-process using 100 visibility samples per vertex for preview quality, 625 visibility samples for medium quality and 2500 samples for high quality.

3 DISPLAY TECHNOLOGIES

Manufacturing appropriate displays for outdoor environments is usually a big challenge. Video-see through displays work well, but reality is viewed through a camera not accurately capturing a scene. Optical see-through displays in turn allow users to visualize reality through the user's own eyes, however in sunny environments traditional LCD or OLED displays are not strong enough to super-impose virtual content and retina displays are much too expensive. Together with Trivisio [Trivisio 2009], we have developed an innovative daylight blocking display that overcomes those limitations while representing a high-quality, cost-effective solution.

3.1 Daylight-blocking device

To the knowledge of the authors, the only comparable works previously done are some first prototypes by Kiyokawa et al. [Kiyokawa et al 2003] and applications for workbenches from Mulder et al. [Mulder et al 2005] that also explicitly use the light-blocking technique. In contrast to real daylight blocking most technologies only augment the brightness of the virtual content reaching the eye as to provide a perception of occluded reality. One of the most effective ways is found for example in Microvision Retina Scanners.

In contrast to the previous approaches to daylight blocking our solution is lightweight and portable allowing for mobile mixed reality applications, such as on-site architectural design review or indoor automotive design review. The most striking difference to conventional displays is the fact, that only by using light-blocking it is possible to create virtual shadows on real objects, which is not possible with conventional display technologies. The innovative contribution of this work (see Fig. 2) is the application of two LCD panels per eye with two distinct functionalities [Santos et al 2008a][Santos et al 2008b]. The first panel (the daylight-blocker) is located at the beginning of the optical path where real world light enters the system. It is able to pixel-wise block real light with an 8 Bit alpha-channel. The second panel is used to create the virtual image which is mixed with the light passing the daylight-blocker. The current display resolution is SVGA (800x600) in true-color, while the resolution of the daylight-blocker is XGA (1024x768) with 8 Bit grayscale.

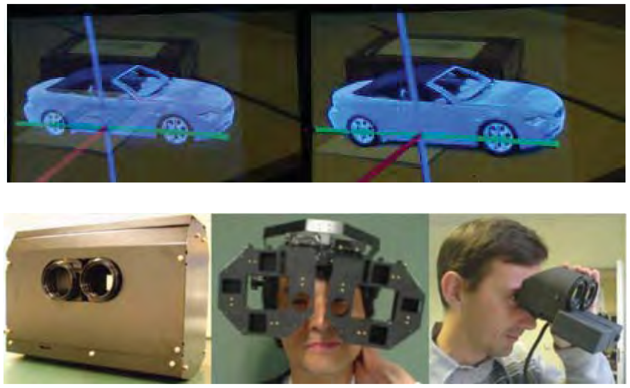


Figure 2. Kiyokawa 2001, Kiyokawa 2003, Our HMD 2009

The final decision on the LCDs used was influenced by the availability of LCD panels, the maximum flexibility and the best brightness. Therefore we selected the transmissive LCD for pixel based real light blocking inside binoculars from Sony in XGA resolution.

For the micro-display generating the virtual image of binoculars (info channel), because of the planned optical scheme only self luminous or back-lighted displays were suitable (reflective types were not practical). From the technical point of view the specifications, in terms of resolution, pixel size, diagonal and aspect of ratio, should match the selected micro-display of light blocker. On the other hand also availability and price where influencing the decision. The chosen display is a backlight illuminated transmissive LCD from Kopin in SVGA resolution for generating the virtual image of binoculars.

The final device consists of two DVI inputs, each receiving the virtual image information and the alpha channel information for the respective eye. Integrated into the device is a small VGA camera "VRmC-3plus" from VRmagic to be used for marker-less tracking in outdoor mixed reality environments.

3.2 Driver

From the rendering point of view the daylight blocking prototype consists of four displays that need to be fed. For each eye there is one color display which has a resolution of 800x600 pixels and a light blocking TFT with a resolution of 1024x768 pixels. Each pair of color and alpha display is accessed via a single DVI input and expects an image of a resolution of 2048x768 pixels. We want to drive all displays from a single PC which means that we have to provide a resolution of 2048x768 on each DVI output of the graphics board. We do this by defining a custom display resolution of 4096x768 (which is possible in current Nvidia display drivers) and configure the desktop of a Windows PC to run in “horizontal span” mode. This setup allows us to create a full-screen window of the needed size of 4096x768 pixels.

In this high resolution window we create four viewports. The layout of the viewports is shown in Fig.3. From left to right the viewports of the window are: left eye color (800x600), left eye alpha (1024x768), right eye color (800x600) and right eye alpha.

The rendering is performed using frame buffer objects (FBO). First the scene is rendered into an FBO. Afterwards the FBO texture is applied to a view-aligned quad that fits exactly into the color viewport of the left eye (the cameras used for the view-aligned viewports are orthographic). The texture coordinate of the view-aligned quad are scaled such that the rendered image fits into the top-left 800x600 pixels of the viewport. Next the FBO texture is applied to a view-aligned quad of the alpha-channel viewport. We render the texture using a shader that calculates alpha values from the original rendering output and writes them to the RGB frame-buffer of the alpha-viewport. This procedure is repeated for the right eye. To calculate the alpha values we define shadow receivers in the virtual scene which are used to blend virtual shadows over real-world objects.

Figure 3. Daylight Blocking Display Driver

In the example shown in Figure 3 there is a shadow receiving plane below the car model. The material of the shadow receivers is diffuse white. The alpha value for each fragment is calculated by performing the lighting calculation twice: First the receiver is rendered without shadows and the resulting color is stored. Second the receiver is rendered with PRT shadowing turned on. The ratio of shadowed color to un-shadowed colour is used as alpha value.

4 TRACKING

A robust and precise pose estimation of the user's display is essential for outdoor mobile mixed reality applications. It is needed so virtual content is seamlessly super-imposed on top of a real scene without flickering. From the beginnings of pose estimation for mixed reality applications when Ivan Sutherland [Sutherland 1968] used Charles Seitz and Stylianos Pezaris ultrasonic technology to track his first HMD, several technologies have been taken into consideration. However not all of the technologies meet the requirements of outdoor mobile mixed reality scenarios:

- **Accuracy** so virtual content is positioned correctly in the real scene.

- **Latency** so no lag is perceived between the pose change of virtual and real parts of a scene.
- **Update rate** so pose changes of the virtual content in the real scene are smooth.
- **Freedom of movement** so the tracking solution is un-tethered, has a small form factor and is light-weight
- **Robustness** so tracking is stable against interferences to its sensors.
- **Large tracking range** covering a broad selection of reference objects on which to super-impose information
- **Low computational load** so the main mixed reality application will run smoothly
- **Low energy consumption** to increase battery life time of the overall mobile system

The latest trend in outdoor pose estimation is hybrid tracking, in which vision sensors are combined and complemented with other pose sensing sensors such as accelerometers, gyros and compasses. In particular the demands of mobile mixed reality applications which need to run on mobile platforms with lower processing power lead to a number of research efforts in that field. In addition the use of several sensing devices improves robustness and stability of the tracking and can effectively be used to compensate for outliers or occlusions in vision-only based approaches. Because inertial sensors tend to drift due to noise accumulation, acceleration and rotational speed data, in particular from accelerometers, can only be used for a short period of time. However it is enough until new absolute data is available from the vision based tracking algorithms.

4.1 State-of-the-Art in Hybrid Tracking

Klein and Drummond [Klein et al 2007] presented a hybrid visual tracking system, based on edge detection and matching against a CAD model. It uses rate gyroscopes to track the rapid camera rotations. Gyroscopes have the benefit of returning and absolute orientation when combined with a 3D compass. Bleser et al. [Bleser et al 2006] combined SFM (structure from motion), SLAM (Simultaneous Localization and Mapping) and model-based tracking. A CAD model was first used for initialization and then 3D structure was recovered during the tracking allowing the camera to move away and explore unknown parts of the scene. DiVerdi et al. [DiVerdi et al 2007] have developed their so called ground cam which extends the precision of an inertial differential tracker by using optical flow on camera pictures taken from the ground the user is walking on. Combined with GPS they are able to calculate an absolute position in space. Reitinger et al. [Reitinger et al 2007] propose an interactive 3D reconstruction system for urban scenes consisting of a scout equipped with a UMPC, a USB camera and a GPS. By iteratively adding poses, a 3D model of the target object is created and refined on the fly.

Running a loop of rendering a simplified textured 3D model of the environment to predict the displacement of the corresponding feature locations in a live camera view is shown in Bleser et al. [Bleser et al 2008]. A camera combined with inertial sensors is used for tracking. In Lee et al. [Lee et al 2008] Distinctive image features of the scene are detected and tracked frame- to-frame by computing optical flow. The user's bare hand is used to define the world coordinate system. Hybrid tracking here is viewed as the combination of SIFT and optical flow. To decrease the gap between GPS precision and vision-based pose estimation approaches Reitmayr et al. [Reitmayr et al 2007] proposed a combination of both. The 2D GPS position together with average user height is used as an initial estimate for the visual tracking.

Many times certain types of pose estimation sensors are better suited for indoor or outdoor usage and it would make sense to have a way for the user to connect to the best possible combination of pose estimation sensors at any given point in time. Huber et al. [Huber et al 2007] and Pustka et al. [Pustka and Klinker 2008] suggest a tracking middleware to facilitate development of stationary and mobile applications by providing a simple interface and encapsulating the details of sensing, calibration and sensor fusion. Kotake et al. [Kotake et al 2007] have proposed a hybrid camera pose estimation method combining the input of an inclination sensor and correspondence line segments. In this method, possible azimuths of the camera pose are hypothesized by a voting method under an inclination constraint. Fong et al. [Fong et al 2008] calculate the carrier phase differences between two local GPS receivers so they are able to increase precision from 10m in average when using one receiver to 10cm when using two GPS receivers which is extremely interesting for outdoor mixed reality applications.

4.2 Current sensor fusion solution

In this paper we introduce a novel sensor fusion pose estimation approach which we combine with the first compact daylight blocking optical stereo see-through display for mixed reality we presented a year ago [Santos et al. 2008a] [Santos et al. 2008b]. Through two new feature matching algorithms and appropriate sequencing of tracking algorithms using different sensors we attempt to achieve time-constant tracking update rates while keeping efforts for pose estimation at a fixed share of the overall available computing performance. By doing so we are able to guarantee the main mixed reality application a fixed share of the remaining computing performance on the mobile platform while preserving high tracking stability.

As opposed to previous implementations which use model based marker-less tracking in their sensor fusion approaches, the current system combines image based marker-less tracking with optical flow and an inertial sensor. In this first stage our pose estimation already covers an almost complete 180 degrees hemisphere without inertial sensor support.

4.2.1 Calibration

Zhang implemented a flexible camera calibration technique [Zhang 1999] that requires multiple views of a planar pattern shown at a few different orientations. Either the camera or the pattern can be freely moved, without previously knowledge of this movement.

4.2.2 Marker-less tracking

For image-based marker-less tracking we have developed a multithreaded SIFT method with two novel feature point matching algorithms which identify non-matching feature point pairs of reference and current images better than previous solutions (KD-tree, BBF search). By doing so, they help increasing tracking stability significantly. Initially both our algorithms receive the list of potential feature pairs from SIFT.

The first algorithm generates a list of largest possible triangles composed of a three-tuple of feature pairs and computes the appropriate afine transformations from reference image to current image. By selecting a single feature point in the reference image and applying all generated transformations to it, the resulting accumulation field shows outliers which are identified leading to removal of their corresponding feature pairs.

The second algorithm is based on epipolar geometry (see Fig.4) which is usually applied for optical stereo tracking. In our case however, instead of using a left and right eye live image to triangulate the position of feature points, we have used the marker-less reference image and the current live image as a basis for identifying non-matching corresponding feature pairs. The fundamental matrix F [Pedram et al. 2007] describes the relation between both images taken from different perspectives of one and the same scene. For two corresponding image points m_1 and m_2, one can set up the following equation:

$$m_2^T * F * m_1 = 0$$

Figure 4. Above: Matching epipolar lines from different perspectives; Below: Identified wrong correspondences in red, correct ones in blue

We now easily identify non-matching feature pairs, because the equation must be non-zero for non-matching correspondences.

In a preliminary test comparing a live camera screenshot with the reference image of the same scene, both containing around 500 feature points, we managed to identify 30% more wrong correspondences with our algorithms than before, thus making our marker-less optical tracking more stable.

4.2.3 Optical flow

To further reduce processing load, we have implemented optical flow similar to the well-known Lucas/Kanada algorithm [Lucas et al 1981]. Optical flow detects how much each image pixel moves between adjacent images, in our case from one frame to the next of the live video stream. Due to the nature of this algorithm, it will work best when there are no occlusions and no changes to the live scene over time, else corresponding features are lost. However, optical flow complements image-based marker-less optical tracking very well in that it can reduce processing load by being used intermittently. In addition, in our implementation image-based marker-less optical tracking will immediately take over when optical flow results deteriorate too fast by a sudden occlusion or dynamic object moving in the scene, as to always providing accurate tracking (see Fig.5).

Figure 5. Sensor fusion tracking sequence

4.2.4 GPS and Inertial sensor

To further complement our outdoor tracking module, we use an XSens MTi-G sensor which includes accelerometers, gyroscopes (3d compass) and a GPS receiver. Using the coarse position estimate of the GPS receiver together with the 3D compass, limits the number of reference images we need to compare the live scene with at a specific geo-location, from a certain vantage point, to a single reference image. Data from gyroscopes helps compensating optical flow estimates for camera rotation to detect when optical flow is deteriorating and a new full image-based, marker-less feature and pose calculation is needed. In addition to gyroscopes for rotation which give reliable absolute orientation data, we are starting to use the sensor's accelerometers for short periods of time when the reference image is no longer in sight to estimate translation. Currently however, results are not yet very stable and we still have to investigate further how to reduce possible electro-magnetic interferences of the different devices and sensors.

5 CONCLUSION

We presented supporting technologies for outdoor mixed reality environments in the areas of rendering, near to the eye displays and pose estimation, particularly suited to the areas of architecture, urban planning and cultural heritage.

Figure 6. Daylight Blocker; Original; Historic; Live view

Our rendering approach allows for photo-realistic rendering results based on HDR image acquisition of the surroundings. To our knowledge our daylight blocking device is unique in the sense that it is an economic, commercially available choice when compared to other display systems, portable, lightweight and ready to be used in mobile outdoor mixed reality scenarios, making it possible for the first time to seamlessly mix real and virtual content in good quality, even in bright sun-light outdoor environments. Our hybrid tracking extension complements the two technologies allowing for accurate pose estimation in outdoor environments. We achieve around 400fps on a Intel Q6600 CPU and 70fps on a portable Sony Vaio VGN-UX1XN UMPC while tracking remains highly stable.

The combined technologies allow architects to visualize their projects directly on location and discuss changes with their customers. They allow the public to better perceive and evaluate proposed alterations of the city landscape by urban planners. Finally, they help preserve and disseminate cultural heritage in a city by offering tourists a new immersive experience such as time travel on location (see Fig. 6).

6 ACKNOWLEDGEMENTS

Our work has been founded by the following European project grants: IMPROVE FP6-IST-2003-004785; CINeSPACE – FP6-IST-2005-034990; MAXIMUS – FP7-ICT-2007-214993;

7 REFERENCES

[Sloan et al 2002]: Sloan, P., Kautz, J., and Snyder, J. 2002. Precomputed radiance transfer for real-time rendering in dynamic, low-frequency lighting environments. In Proceedings of the 29th Annual Conference on Computer Graphics and interactive Techniques (San Antonio, Texas, July 23 - 26, 2002). SIGGRAPH '02. ACM, New York, NY, 527-536.

[Klinker et al 2002]: Klinker, G., Dutoit, A. H., Bauer, M., Bayer, J., Novak, V., and Matzke, D. 2002. Fata Morgana " A Presentation System for Product Design. In Proceedings of the 1st international Symposium on Mixed and Augmented Reality (September 30 - October 01, 2002). Symposium on Mixed and Augmented Reality. IEEE Computer Society, Washington, DC, 76.

[ARIS 2004]: ARIS project (2004) European research project aris, augmented reality image synthesis through illumination reconstruction and its integration in interactive and shared mobile ar-systems for e (motion)-commerce applications. http://aris-ist.intranet.gr

[Pomi et al 2003]: Pomi, A., Marmitt, G., Wald, I., Slusallek, P. 2003. Streaming Video Textures for Mixed Reality Applications in Interactive Ray Tracing Environments. In Virtual Reality, Modelling and Visualization 2003 (VMV), November 19-21, Munich, Germany. 2003

[Santos et al 2007]: Santos, P., Gierlinger, T., Stork, A., McIntyre, D. 2007. Display and rendering technologies for virtual and mixed reality design review. In 7th International Conference on Construction Applications of Virtual Reality, CONVR 2007. Penn State University, University Park, PA, USA.

[Gierlinger et al 2009]: Gierlinger T., Danch, D. Stork, A. 2009. Rendering techniques for mixed reality. In Journal of Real-Time Image Processing, Special Issue, Springer Berlin / Heidelberg. DOI 10.1007/s11554-009-0137-x

[Debevec et al 1998]: Debevec, P. Rendering synthetic objects into real scenes: bridging traditional and image-based graphics with global illumination and high dynamic range photography. In SIGGRAPH '98: Proceedings of the 25th annual conference on Computer graphics and interactive techniques (New York, NY, USA, 1998), ACM, pp. 189-198.

[HDRShop 2009]: Website: http://www.hdrshop.com/; Last visited: 01.12.2009

[PFSTools 2009]: Website: http://pfstools.sourceforge.net/; Last visited: 01.12.2009

[AHDRIA 2009]: Website: http://www2.cs.uh.edu/ ~somalley/ hdri_images.html; Last visited: 01.12.2009

[SPHERON 2009]: Website: http://www.spheron.com/; Last visited: 01.12.2009

[Shirley et al 2003]: Shirley, P., and Morley, R. K. Realistic Ray Tracing. A. K. Peters, Ltd., Natick, MA, USA, 2003.

[Trivisio 2009]: Website: http://www.trivisio.com/; Last visited: 01.12.2009

[Kiyokawa et al 2003]:Kiyokawa K., Billinghurst M., Campbell B., Woods E., An Occlusion-Capable Optical See-through Head Mounted Display for Supporting Co-located Collaboration, Proceedings of the 2nd IEEE and ACM International Symposium on Mixed and Augmented Reality, p.133, October 07-10, 2003

[Mulder et al 2005]Mulder, J.D.Realistic occlusion effects in mirror-based co-located augmented reality systems. In Virtual Reality 2005. Proceedings of VR2005, IEEE 12-16 March 2005, p.203-208

[Santos et al 2008a]: Santos, P., Gierlinger, T., Machui, O., Stork, A.2008. The daylight blocking optical stereo see-through hmd. In IPT/EDT 2008. Proceedings, 4.

[Santos et al 2008b]: Santos, P., Gierlinger, T., Machui, O., Stork, A.2008. An innovative daylight blocking optical stereo see-through hmd. In Siggraph 2008. Full Conference DVD-ROM, 1.

[Sutherland 1968]: Sutherland, Ivan E.: A head-mounted three dimensional display. In: AFIPS'68 (Fall, part I): Proceedings of the December 9-11, 1968, fall joint computer conference, part I. New York, NY, USA : ACM, 1968, S. 757 764 2, 2.1

[Klein et al 2007]: Klein, G.; Murray, D.: Parallel Tracking and Mapping for Small AR Workspaces. In: ISMAR '07: Proceedings of the 2007 6th IEEE and ACM International Symposium on Mixed and Augmented Reality. Washington, DC, USA : IEEE Computer Society, 2007. ISBN 978 1 4244 1749 0, S. 1 10 2.1.3

[Bleser et al 2006]: Bleser, G.; Wuest, H. ; Stricker, D.: Online camera pose estimation in partially known and dynamic scenes. In: ISMAR '06: Proceedings of the 2006 Fifth IEEE and ACM International Symposium on Mixed and Augmented Reality (ISMAR'06). Washington, DC, USA : IEEE Computer Society, 2006. ISBN 1 4244 0650 1, S. 56 65 2.1.3

[DiVerdi et al 2007]: DiVerdi, S. ; Hollerer, T.: GroundCam: A Tracking Modality for Mobile Mixed Reality. In: Virtual Reality Conference, 2007. VR '07. IEEE, 2007, S. 75 82 2.1.3

[Reitinger et al 2007]: Reitinger, B.; Zach, C.; Schmalstieg, D.: Augmented Reality Scouting for Interactive 3D Reconstruction. In: Virtual Reality Conference, 2007. VR '07. IEEE, 2007, S. 219 222 2.1.3

[Bleser et al 2008]: Bleser, G.; Stricker, D.: Advanced Tracking through Efficient Image Processing and Visual-inertial Sensor Fusion. In: IEEE Virtual Reality 2008. Proceedings, 2008, 137-144 2.1.3

[Lee et al 2008]: Lee, T.; Hollerer, T.: Hybrid Feature Tracking and User Interaction for Markerless Augmented Reality. In: Virtual Reality Conference, 2008. VR '08. IEEE, 2008, S. 145 152 2.1.3

[Reitmayr et al 2007]: Reitmayr, G.; Drummond, T.W.: Initialisation for Visual Tracking in Urban Environments. In: Mixed and Augmented Reality, 2007. ISMAR 2007. 6th IEEE and ACM International Symposium on, 2007, S. 161 172 2.1.3

[Huber et al 2007]: Huber, M.; Pustka, D. ; Keitler, P. ; Echtler, F. ; Klinker, G.: A System Architecture for Ubiquitous Tracking Environments. In: Mixed and Augmented Reality, 2007. ISMAR 2007. 6th IEEE and ACM International Symposium on,2007, S. 211 214 2.1.3

[Pustka and Klinker 2008]: Pustka, D.; Klinker, G.: Dynamic gyroscope fusion in Ubiquitous Tracking environments. In: Mixed and Augmented Reality, 2008. ISMAR 2008. 7th IEEE/ACM International Symposium on, 2008, S. 13 20 2.1.3

[Kotake et al 2007]: Kotake, D.; Satoh, K. ; Uchiyama, S. ; Yamamoto, H.: A Fast Initialization Method for Edge-based Registration Using an Inclination Constraint. In: Mixed and Augmented Reality, 2007. ISMAR 2007. 6th IEEE and ACM International Symposium on, 2007, S. 239 248 2.1.3

[Fong et al 2008]: Fong, W.T.; Ong, S.K. ; Nee, A.: A di erential GPS carrier phase technique for precision outdoor AR tracking. In: Mixed and Augmented Reality, 2008. ISMAR 2008. 7th IEEE/ACM International Symposium on, 2008, S. 25 28 2.1.3

[Zhang 1999]: Zhang, Z.: Flexible Camera Calibration By Viewing a Plane From Unknown Orientations. Version: 1999. IEEE 1999

[Pedram et al. 2007] Pedram Azad, Rüdiger D. Tilo Gockel G. Tilo Gockel: Computer Vision - DasPraxisbuch. 1. Auflage. Elektor Verlag, 2007. – ISBN 978–3–89576–165 3.2,3.2,3, 4.5.2.1

[Lucas et al 1981] B.D. Lucas and T. Kanade, "An Iterative Image Registration Technique with an Application to Stereo Vision",DARPA Image Understanding Workshop, 1981, pp121–130 (see also IJCAI'81, pp674–679).

Conversion of One- to Two-Way Streets in Birmingham Downtown: A Feasibility Study

Virginia P. Sisiopiku, Ph.D.
The University of Alabama at Birmingham
1530 3rd Ave S, Hoehn 311
Birmingham, AL 35205-4440
vsisiopi@uab.edu

Jugnu Chemmannur, MSCE
TRIA, Inc.
Birmingham, AL 35244
jugnu.c@gmail.com

James Brown, PE
Gonzalez-Strength and Associates, Inc.
2176 Parkway Lake Drive
Birmingham, AL 35244
jbrown@gonzalez-strength.com

Key Words: Signal optimization, one-way operations, two-way operations, traffic impacts analysis.

ABSTRACT

Many urban areas are looking for ways to attract new businesses and residents in city centers and transform declining downtown areas into vibrant elements of the urban landscape. As a principal revival option, studies are conducted nationwide to understand the upshots and relevance of conversion of one-way streets to two-way operations. This paper summarizes the approach, findings and evaluation methodology for a study that investigated the feasibility and potential impacts of street conversion options on traffic operations in downtown Birmingham, Alabama. Of particular interest were issues related to conversion of grid systems from one-way to two-way operations. Recommendations on the most promising design and control configurations for implementation are also provided.

1. INTRODUCTION

In the 1950s and 60s, one-way streets were implemented in cities across the US in an attempt to rid downtowns of traffic congestion while avoiding widening streets or constructing new facilities. Although this was done during the pre-freeway era in order to improve vehicle access to downtowns, most one-way street couplets were not converted back to two-way operations even after freeways came into existence [2].

During the 1990s, a nationwide trend emerged to revert some one-way couplet streets back to two-way streets. The intent of this conversion was to help increase the livability of the surrounding neighborhoods and to make the streets more compatible with pedestrian and bicycle traffic [2]. The prevailing wisdom is that two-way streets enhance a neighborhood's environment, reduce speeds to levels that are more compatible with pedestrian presence, and create a healthy business environment [2].

Businesses tend to locate in areas in which they will be able to conveniently serve their customers. One-way streets generally have high traveling speeds, which may reduce a prospective customer's ability to discern individual store fronts. Slower driving speeds and a higher propensity to walk can enhance location awareness. In addition, two-way operations may increase storefront visibility because there is less "eclipsing" of storefront exposure [3].

Lately, with the gas prices and living costs soaring high, suburban life is not as inexpensive as it used to be. People's priorities have hence started to change, and more and more people have started to exhibit an inclination to stay around city centers for ease of travel and increased accessibility to facilities. In response to this trend, many city planning agencies are considering one- to two-way street conversions as a means of downtown rehabilitation for revival of city centers.

2. LITERATURE REVIEW

The focus of the literature review was on the operational impacts of one- to two-way conversions; changes in crash frequencies and/or patterns; motorist response to changes; cost of implementation; economic vitality; and impacts of conversions on pedestrian, bicycle, and transit usage. The review and synthesis of technical literature helped with identifying the issues and challenges associated with street conversions as well as with determining best practices.

2.1 Travel Patterns and Accessibility

The literature review identified a number of case studies that successfully implemented one- to two-way conversions. For instance, a portion of the downtown Milwaukee street system that historically operated as one-way street system was converted to a two-way system in 2006. The two-way street system was shown to exhibit improved business accessibility, create a less confusing circulation system for downtown visitors and business customers, and permit transit passengers to board and exit city buses at the same intersection and in closer proximity to their destinations [6].

In response to the Vision 2020 Plan, researchers conducted a study on restoring two-way traffic to North Front Street in the City of Wilmington, NC. The study findings revealed that this conversion would improve access to existing businesses and provide better traffic circulation with the redevelopment of the northern downtown area [4].

Literature reports show that 40% of businesses in Vine Street, Cincinnati, were closed after a street was converted to one-way from two-way [3]. Also, a one-way system usually yields approximately 120 to 160% more turning movements when compared to a two-way system and displays 20% to 50% greater travel distances between origin and destination [3]. Similarly, in a microscopic simulation study, Meng and Thu found that, when a one-way only network was compared to a two-way only network, the total vehicle miles traveled (VMT) of the one-way network was about 20% more [5]. These findings showcase the great potential of using two-way operations instead of one-way.

2.2 Traffic Speed and Safety

Although several studies were identified that evaluated safety related data, no studies were found that considered both crash frequency and crash severity pre- and post-conversion. However, there are some conflicting data concerning crash frequency.

Improved safety was reported in several locations. For example, in the city of Edmonton, Alberta post two-way conversion vehicle crashes decreased by 4% [7]. Portland, OR also displayed 51% fewer accidents at intersections and 37% fewer between intersections [8]. Moreover, police departments in both Louisville, KY (pop: 1,005,000), and Hamilton, Ontario (pop: 662,000), reported that patrolling on the converted two-way streets was easier than before, and shop visibility has increased [7].

However, a number of studies reported increase in crash rates following the two-way conversion. Three years after the 1993 conversion of a major route to two-way operation in Indianapolis, IN, accidents on that route had increased by 33%. In 1996, Lubbock, TX, converted several one-way streets to two-way. Two years later, monitoring found a 12% decrease in traffic on those routes but 25% more accidents causing 34% more property damage [8]. According to Cunneen and O'Toole, a measurement of street safety by Denver planners before and after conversion found that conversion of a one-way street to two-way led to a 37% increase in accidents [8]. Furthermore, the two-way streets reportedly produced 163% more pedestrian accidents in Sacramento and 100% more pedestrian accidents in Portland, OR; Hollywood, FL; and Raleigh, NC [8]. However, increased pedestrian activity was also observed post-conversion.

It should be noted that two-way streets may experience a larger number of accidents but that those accidents tend be less severe because average speeds tend to be lower. Because the risk of death and injury is an exponential function of vehicle speed, a 20% reduction in speed would result in a greater than a 20% reduction in the chances of death and injury [7].

2.3 Pedestrian Considerations

Converting streets to two-way operations allows the introduction of medians, which can be used as pedestrian islands. In this way, a pedestrian needs only to cross one direction of traffic at a time and thus to cross fewer lanes at a time. The decrease in vehicle speeds on two-way corridors is also an enhancement to pedestrian safety because the risk of serious injury and death is an exponential function of speed [7]. Even if the accidents are less severe, two-way street operations may indeed increase the risk for pedestrian crashes, as pedestrian/vehicle conflicts typically increase.

2.4 Traffic Operations

The survey of available studies indicates that, when one-way couplets were converted to two-way operation, some cities (e.g., Edmonton, Alberta (pop: 900,000);

San Francisco, California (pop: 739,426); Hamilton, Ontario (pop: 662,000)) experienced displacement of traffic onto parallel corridors. This displacement is at least partially a function of the well-developed grid pattern of streets in most downtowns and partially a function of the increased demand that results from two-way operation on the formerly one-way corridors. If there is sufficient excess capacity on these parallel corridors, the increase in traffic could make them more attractive for redevelopment [7].

Planners in Austin, TX (pop: 680,000), estimated that converting one-way couplets to two-way operations would increase travel delay by 23% overall [7].

A study published in the Journal of the Institution of Transportation Engineers found an average speed differential of 4 to 5 mph when comparing the one-way and two-way networks; this differential would equate to an additional 6 minutes of travel time per half-mile [5]. Edmonton, Alberta, reported reduced vehicle speeds along corridors recently converted from a one-way to a two-way operation. In most cases (e.g., West Palm Beach, FL (pop: 1,049,000), and Louisville, KY (pop: 1,005,000)), post-conversion capacity was not an issue [7]. The survey of conversions in other communities revealed that, in some cases (e.g., New Haven, CT (pop: 126,000), and Hickory, NC (pop: 36,000)), the conversion to two-way operations resulted in less confusion for out-of-town visitors and in a more "user friendly" image for the community [7].

2.5 Economic Impacts

There is an agreement among literature resources that two-way street conversions have positive impacts on economic growth. West Palm Beach, Florida (pop: 85,000), reported a dramatic increase in new retail shops, restaurants, and residential use after such a conversion. They attribute the change to exchanging mobility (i.e., vehicle speed) with access brought about by the two-way circulation as well as with livability through streetscape design. Property values increased as well [7]. Toledo, Ohio (pop: 323,000), reported that long time vacant buildings were now being occupied or sold to developers for new shops and restaurants [7]. Merchants in Lafayette, Indiana (pop: 50,000), were very concerned about the loss of traffic at first but found that business traffic actually picked up after the conversion [7]. Charleston, SC (pop: 95,000), experienced a dramatic increase in new retail and service businesses in the area [7]. Lubbock, TX (pop: 200,000), reported that the City has not received any unfavorable comments, and the general consensus is that the conversions have been beneficial to the central business district (CBD), which started experiencing growth after several years of decline [7]. In Kitchener, Ontario (pop: 209,000), office vacancies declined from 35% to 11%, new housing units increased by 700 units, and street front retail vacancy rates declined from 12.4% to 7.8% [7]. Albuquerque, NM (pop: 678,000), reported that automobiles do not move as quickly out of the downtown area but that more people seem to be staying downtown after hours. These finding further confirm the economic potential of two-way conversions for downtown revitalization.

2.6 Environmental Impacts

The findings of a study conducted by Meng L.K et al in 2004 [5] indicated that a one-way street generates lesser emissions (0.28 gm/mile of HC emission compared to 0.21 gm/mile on two-way streets, 20.8 gm/mile CO emission against 23.4 gm/mile, and 1.10 gm/mile NO_x emission compared to 1.26 gm/mile) and reduces fuel consumption (34.8 miles/gallon compared to 27.1 miles/gallon on two-way streets). According to mathematical calculations of planners in Austin, TX conversion of several one-way streets to two-way could increase their downtown air pollution by 10 to 13 % [9]. However, it is suggested that efficient synchronization of signal systems could assist a great deal in creating a stable traffic flow and thereby in reducing adverse impacts on the environment.

3. PROJECT BACKGROUND

3.1 Project Objective

Prior to the implementation of one- to two-way street conversion, a need exists to investigate the street conversion options, their feasibility, and their potential effects on existing and future traffic operations in downtown Birmingham. Hence, this study focuses on analyzing the possibility and impacts of converting selected one-way streets to two-way operations as a strategy for creating a sustainable transportation system in the Birmingham City Center and thereby improving accessibility and downtown circulation and providing a range of transportation options. Particular attention has been given to ensuring that the new system would improve accessibility to new and evolving land uses and would better serve the existing and future transit and non-motorized transportation needs while maintaining overall street network flow.

3.2 Research Methodology

In order to meet its objectives, the study carried out several sequential tasks in a predetermined fashion. Following a systematic study of one to two-way street through a literature review, an extensive data collection and processing effort was executed to acquire insights into the existing traffic operations and demand characteristics at the study site. Details are provided in section 4.1 below.

Once the existing conditions were determined, a detailed traffic impact analysis was performed. This analysis involved traffic simulation modeling and signal optimizations and was undertaken to assess the potential implications of two-way street conversions at and around the test sites.

4. SIMULATION MODEL DEVELOPMENT

4.1 Data Collection

In order to develop the base case study networks, extensive geometric data (i.e., number of lanes, lane widths, turning restrictions, medians etc) and traffic data (traffic counts by movement type, pedestrian activity, transit/heavy vehicle counts) were collected from a series of field visits and observations. Existing signal timing information and traffic change patterns were obtained from the City of Birmingham.

4.2 Simulation Model Selection

For a successful completion of this project, it was imperative that the selected simulation software replicate the existing conditions as close to reality as possible by incorporating as many details as possible about the traffic operations, geometric elements and land use information. Equally important was the software's capability to handle proposed future geometries and traffic conditions. Another consideration was the types of outputs from the candidate models. Careful consideration of requirements and capabilities of various models led to the selection of SYNCHRO as the simulation software of choice for the study.

SYNCHRO, developed by Trafficware Inc, is a complete software package capable of modeling and optimizing traffic signal timings. SYNCHRO Version 5.0, which was used for this study, implements methods of the 2000 Highway Capacity Manual (2000HCM), for capacity analysis. Besides calculating capacity, SYNCHRO can also optimize cycle lengths, splits and offsets on the basis of delay and stops reduction principles. Thus the software eliminates any need to try multiple timings plans in search of the optimum. Additionally, for coordinated intersections, such as those in our case, SYNCHRO explicitly generates progression factors that assist in determining the sufficiency of provided signal timings. SYNCHRO is also fully interactive and any changes made to input values will automatically update the results thus reducing the network development and simulation time.

4.3 Network Creation

The study traffic network was created in SYNCHRO with the help of links and nodes with links defining streets and nodes representing intersections. Lane data, traffic volumes and signal timings were then entered by clicking on these links. Care has been taken to re-create actual distances at the site with the assistance of digital photos, Google maps, Google earth, and field observations. The SYNCHRO network creation screen is shown in Figures 1a and 1b under one- and two-way operations respectively.

Figure 1(a): Sample SYNCHRO Network for Existing Scenario (One-way Operation).

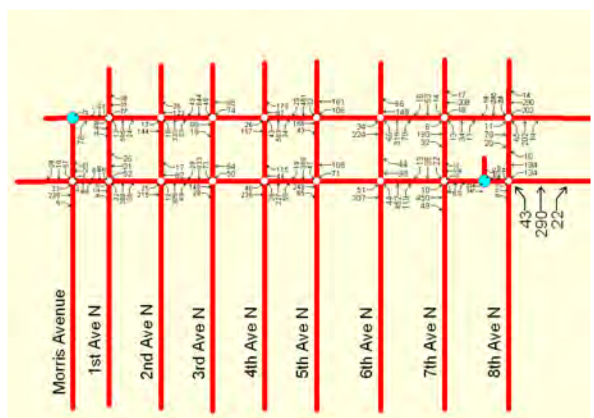


Figure 1 (b): Sample SYNCHRO Network for Two-Way Operations.

The lane input window in SYNCHRO makes it possible to specify lane groups, lane width, grade, area type, storage length, detector locations and other information. Storage length data, measured at site, are useful when it comes to detecting blockage problems. Traffic volumes for each lane group, peak hour factor, percentage of heavy vehicles, and other related details were entered by using the traffic volume input sheet.

Signal timing information such as left turn type, phase number, lead/lag assignments, minimum and maximum splits, lost times etc was entered in the timing and phasing windows.

4.4 Design Scenarios

The expected changes in geometric design, traffic demand and signalization plans resulting from the introduction of the two-way operations were considered during the development of design scenarios. Three scenarios were applied for network creation and simulation [1], as follows:

- **Base Case – Existing One-Way Conditions**
  To facilitate comparisons, a base case was modeled that accurately represents the existing lane configurations, traffic conditions and signal timings. In this scenario, 7 study streets operate as one-way, namely 2^{nd} Ave N, 3^{rd} Ave N, 4^{th} Ave N, 13^{th} & 14^{th} Str N and 17^{th} & 18^{th} Str N.
- **Case 1 (LC1)– Phased Conversion (3 out of the proposed 7 corridors considered for conversion)**
  Careful thought yielded the idea that a potential phased conversion would be a more practical approach. Hence Case 1 scenario studied the impacts from the conversion of just 2^{nd} Ave N and 13^{th} and 14^{th} Street pair to two-way operations. Those corridors appeared to have obvious advantages as potential two-way street candidates especially since they were free of on- and off-ramps to expressway and highways, contrary to their counterparts (i.e., 3^{rd} Ave N, 4^{th} Ave N, 17^{th} Str and 18^{th} Str).
- **Case 2 (LC2)– Full Conversion (All the 7 study streets converted to two-way operations)**
  Case 2 performed analysis assuming future two-way conversion of all 7 study corridors of the downtown grid. However, in order to eliminate potential confusions and considering the design complexities, two-way conversions have been terminated at the intersection prior to the on/off ramp to/from highways.

4.5 Two-Way Design Considerations

4.5.1 Geometric Design

Under the 2-way operation the currently existing Right of Way (ROW) was maintained at all locations so that no buildings, landscaping or parking were to be compromised, should the proposed two-way design be implemented. The outer lanes were designed to have a width varying from 13 to 15 ft so that they could be utilized by bicycles as well. Any additional ROW available was recommended to be utilized for central medians, landscaping, or diagonal parking.

In order to facilitate the left turns without affecting the through vehicle movement, a central two-way left turn lane (TWLTL) was provided wherever turning volumes were high or considered necessary. These lanes were designed for a width varying from 10 to 12 ft. Also care has been taken to maintain consistent lane widths and number of lanes along the entire stretch of a corridor so as to eliminate any serpentine motion of vehicles. A sample corridor layout under the proposed two-way design is illustrated in Figure 2.

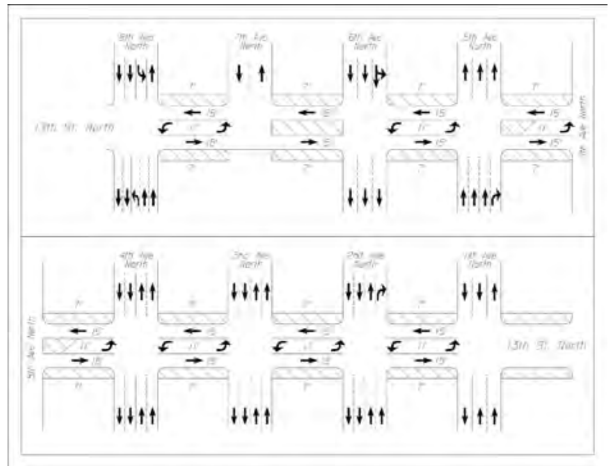


Figure 2: Proposed Geometric Design of 13^{th} St N under Two-way Operations

4.5.2 Parking Considerations

The proposed future geometry retained the existing parking facilities wherever made possible by the ROW and design priorities. Parallel parking at some locations was converted to diagonal and vice versa, depending on the ROW availability and/or limitations. Additionally, some extra on-street parallel and diagonal parking was introduced at a few locations. The width of existing on-street parallel parking on the study streets was 7 ft. The same was maintained in the proposed design. Any diagonal parking indicated has a ROW requirement of 13 ft.

4.5.3 Traffic Demand Considerations

The *Base Case* scenario exhibits the existing volume of vehicles that was recorded at site. In order to better evaluate the pattern in which two-way traffic would behave immediately after conversion, driver behavior/expectancy was considered during the assignment of traffic to the future design scenarios.

In *Case 1* it was assumed that since 13^{th} and 14^{th} Strs currently carry vehicles SB and NB respectively, drivers familiar with the site would be more likely to travel on 13^{th} Str to go south and 14^{th} Str to go north immediately after the conversion. Bearing this in mind, it was decided to retain greater volumes of traffic on streets respective to their current flow direction and thereby distribute volumes in 3:2 proportions. Hence, it was assumed that, under two-way operations, 13^{th} Str would carry 60% of the currently existing demand of SB vehicles, and transferred the remaining 40% to 14^{th} Str N. Similarly, 60% of NB vehicles were retained on 14^{th} Str N and the remaining transferred to 13^{th} Str N.

Moreover, 2^{nd} Ave N currently operates as one-way WB and hence when converted to a two-way street the traffic pattern is bound to change. Drivers who were forced to travel on 2^{nd} Ave previously to reach their destination post conversion will have the option of traveling on 1st Ave N and turn left/right or go on 4th Ave N (which also carries vehicles WB) and turn left. Considering this, it was decided that sharing 20% of 2nd Ave N WB traffic with 1st Ave and another 20% with 4th Ave N would represent the future operations appropriately.

Since 2nd Ave N carries vehicles both ways, it was assumed that a portion of vehicles would shift from their normal route on 3rd Ave and cross over to 2nd Ave N in order to reach their destination and avoid additional maneuvers.

In *Case 2* all the 7 study streets are proposed to be converted to two-way operations. The same assumption of driver expectancy is considered here as well. It is assumed that since the above streets currently carry vehicles SB and NB, drivers are more likely to travel on 13th and 17th Str to go south and 14th and 18th Str to go north even after the conversion. Hence, it was decided to distribute volumes in 3:2 proportions. That is, it was assumed that, under two-way operations, 13th and 17th Streets will carry 60% of the currently existing demand of SB vehicles and 14th and 18th Streets N will retain 60% of NB vehicles. The remaining 40% of volumes were transferred to their street pairs.

The 3rd Ave N and 4th Ave N act as a pair, one carrying vehicles EB and the other taking them WB. When converted to two-way streets their traffic patterns will be altered. Drivers who were forced to travel on any one street just because of its one-way nature would now have the leisure of traveling on any street of their preference and turn left/right to reach their destination. Nonetheless, driver expectancy factor is still very much in effect. Hence it was decided that allocating 20% of the total EB traffic that now runs on 3rd Ave N onto 4th Ave (which was WB one way initially) and transferring 40% of WB movement from 4th Ave to 3rd would adequately represent the future operations.

4.5.4 Traffic Control Considerations

The *Base Case* network was modeled to incorporate all the traffic signal features that were obtained from the City of Birmingham. The downtown grid contains intersections that operate with signalized and unsignalized controls. Most of the signalized intersections operate as pre-timed and few locations are controlled by coordinated signals. The current downtown network has an existing optimum cycle length (CL) of 80 sec. In order to provide a basis for comparison and accurately represent the existing scenario, the base case was not optimized for intersection splits or network CL.

As far as *Case 1* and *Case 2* networks are concerned, 7 one-way streets are converted to two-way operations. For the affected corridors it is imperative that the signal timings be altered, in order to ensure smooth traffic flow and effective operation of the grid.

5. RESULTS

The existing and proposed alternate networks were modeled by using SYNCHRO and simulation runs were carried out in two basic steps:

a) Intersection split optimizations, and
b) Network cycle length optimizations.

Because the study streets are a part of a greater downtown grid, changing CL and offsets would have potential undesirable impacts on signal timings of the entire downtown system. For this reason, it was considered appropriate to retain the present values of CLs and offsets for every intersection. Hence all proposed scenarios, including the base case, have been optimized for splits alone.

SYNCHRO provided an optimum CL value of 80 seconds for all scenarios. Hence, the usage of 80 sec signal CL for future operations to match with the remaining downtown street networks is validated.

5.1 Operational Impacts

The results presented in Table 1 clearly demonstrate the advantage of the two-way designs over the existing one-way operations. Operational impacts such as signal delay, travel time, number of stops, average speed and distance travel show an improvement under the proposed two-way operations (*Cases 1* and *2*) as compared to current operations (*Base Case*). For example, due to optimized signal timings, signal delay per vehicle is expected to fall by 22.2% under two-way operations compared to the Base Case.

Table 1: Comparison of Various MOEs for Entire Network as Provided by SYNCHRO

Network-Wide MOEs	Base Case	Case 1	Case 2
Signal delay / veh (sec)	9	7	7
Total signal delay (hrs)	222	174	180
Stops / veh	0.62	0.53	0.58
Stops	54,139	46,649	50,947
Average speed (mph)	16	18	18
Total travel time (hrs)	480	431	437
Distance Traveled (mi)	7,761	7,719	7,698
Queuing Penalty (veh)	317	167	18
Performance Index	381.1	307.8	322.2

Moreover, a 5.9% reduction in stops is expected in Case 2 along with a 12.5% increase in average speed.

A comparison between Case 1 and Case 2 shows that the partial conversion plan proposed in Case 1 yields slightly better results. Still the differences are not significant and thus either option appears to be a good candidate for implementation.

Overall, the findings are reassuring that the network will perform well in the event of conversion of the proposed streets to two-way and no adverse impacts on traffic circulation should be expected.

5.2 Environmental Impacts

Table 2 shows that fuel consumption, CO emissions, NO_x emissions and VOC emissions are predicted to reduce upon conversion of the proposed streets to two-way operations (from 6.5% to 10%). This can be attributed to the fewer stops and reduced delays that vehicles would have to encounter. Also, the distance that a driver is expected to travel upon conversion is found to reduce considerably thereby facilitating avoidance of any additional maneuvers required by a vehicle to reach its destination. Overall, no negative impacts on the environment are expected due to the proposed conversions.

Table 2: SYNCHRO Output – Environmental MOEs for Entire Network for All Scenarios

Network-Wide MOEs	Base Case	Case 1	Case 2
Fuel Consumed (gal)	781	703	731
Fuel Economy (mpg)	9.9	11.0	10.5
CO Emissions (kg)	54.61	49.13	51.07
NO_x Emissions (kg)	10.63	9.56	9.94
VOC Emissions (kg)	12.66	11.39	11.84

5.3 Parking Impacts

At a public meeting preceding the project, citizens and local business owners expressed concerns about the potential loss of parking, should a two-way street conversion be implemented in downtown Birmingham. With the design considered in this study existing parking has been retained under the two way operation with the exception of one city block (4th Ave between 17th and 18th Str N). Moreover, a number of locations (10 city blocks total) gained additional parking spaces on one or both sides whereas some locations retained the existing parking spots but benefited from increase of parking width. Overall there is clear evidence that the on-street parking situation in the Birmingham City Center is expected to improved, shall a conversion of the study corridors to two-way streets takes place.

6. CONCLUSIONS

After careful analysis, it was found that the proposed two-way street conversions in downtown Birmingham, AL are feasible for the given study network. No adverse impacts on traffic operations or the environment are expected from potential conversions. However, modifications to the existing signal timings are essential in order to bring about these conversions and to ensure smooth and efficient traffic flow through the streets in the downtown grid.

A summary of study conclusions and recommendations follows.

- Many locations across the nation have successfully converted one-way corridors to two-way. In earlier studies two-way streets have been found to increase accessibility to businesses and promote economic growth, offer a traffic calming option in downtown areas, provide better internal circulation, and improve neighborhood livability and sustainable development.
- Conversion of all study streets to two-way operations is feasible and is not expected to result in deterioration of service along any of the study corridors or any adverse environmental impacts. Moreover, no loss in parking spaces is expected by implementing the two-way facility design options studied in this project.
- The conversion may be implemented either all at once, or in phases starting with conversion of 2nd Ave N and 13th and 14th Street (Phase 1), and following with conversion of the remaining facilities (Phase 2). Advantages of the latter approach include the improved chances in securing funding and moving forward in a timely fashion, straightforward implementation of the two-way operation along corridors that are free of on- and off-ramps to expressway and highways, and the opportunity for re-evaluation of the need to convert the other four corridors after the completion of Phase 1 given feedback from the public, and traffic operations in downtown Birmingham at that time. The main shortcoming of this strategy is that signal timings need to be evaluated and programmed twice (i.e., once for Phase 1 operations, and once again under Phase 2). This is a time consuming and costly process that may prove to be a less practical and desirable strategy as compared to the conversion of all 7 corridors at once.
- In any event, a comprehensive follow up signal timing study is required to determine the signal timings of all affected signalized intersections in the City Center. This should expand beyond the 7 corridors as signal progression and redistribution of

traffic due to the two-way conversion will impact the signal timings of intersections in the adjacent corridors. The signal timing study will require (a) detailed field data collection of traffic volumes, geometric data, traffic control, parking, and other relevant data and (b) analysis using an appropriate simulation and signal timing optimization model such as SYNCHRO.

- The successful implementation of two-way operations will depend on the ability for the City of Birmingham to secure funding for recommended improvements and the support from local businesses and the public.

REFERENCES

1. Gonzalez-Strength and Associates Inc., and TRIA Inc., Feasibility Analysis of Converting Certain One-way Streets to Two-way traffic in Birmingham's City Center, Technical Report Prepared for the City of Birmingham, October, 2009, Birmingham Al.
2. "Why is the City conducting the Central City Two-way Conversion Study?" Central city two-way conversion study – Purpose of study. City of Sacramento, Department of Transportation. http://www.cityofsacramento.org/transportation/engineering/fundingcentralpurpose.html
3. "Downtown Streets: Are we Strangling Ourselves on One-way Networks?" by G. Wade Walker, Walter M. Kulash and Brian T. McHugh, Glatting Jackson Kercher Anglin Lopex Rinehart Inc. TRB Circular E-C019: Urban Street Symposium.
4. "Feasibility Study: Restoring Two-Way Traffic to North Front Street" by James R. Flechtner, 2005. City of Wilmington, NC.
5. "A Microscopic Simulation Study of Two-way Street Network versus One-way Street Network" by Lum Kit Meng and Soe Thu, 2004, Journal of The Institution of Engineers, Singapore, (44) 2.
6. "Applying New Urbanism Street Principles in Downtown Milwaukee, WI, USA", by Kenneth H. Voigt, May 2006, ITE Journal. http://findarticles.com/p/articles/mi_qa3734/is_/ai_n17177161
7. "Fargo-Moorhead Downtown Framework Plan Update" Report by Fargo-Moorhead Council of Governments, January 2007, City of Fargo & City of Moorhead.
8. "No Two Ways About It: One-Way Streets are Better than Two-Way", by Michael Cunneen and Randal O'Toole, February 2005, Independence Institute, Colorado, Issue Paper 2.
9. "Should Cities Convert One-way Streets to Two-way?" from Thoreau Institute Website. http://www.ti.org/vaupdate30.html

Virtual Driving and Eco-Simulation

VR City Modeling, Drive Simulation, and Ecological Habits

Christopher J. Grasso,
Forum8AZ LLC, USA
chris@forum8az.com

Michael J. McDearmon,
Arizona State University, USA
michael.mcdearmon@asu.edu

Yoshihiro Kobayashi,
Arizona State University, USA
ykobaya@asu.edu

Keywords: Drive Simulation, Traffic Visualization, Virtual Reality, Driving Behavior, Vehicle Emissions

Abstract

This paper introduces a VR city model developed for research in driving simulation, driver behavior, and vehicle emissions. The project is a part of an interdisciplinary multi-year academic research grant at Arizona State University. First, the outline of the research is explained. The methods of VR city modeling are then introduced. The modeling process of creating road and intersection networks and traffic flow is explained. The system integration with a PC, modeling and VR software, and a Drive Simulator is illustrated. A research study of driver behavior and vehicle emissions is detailed. Next, a case study linking research in ecological driving with a simulator and Japan's National Agency for Automotive Safety & Victims' Aid is examined. Lastly, computational tools to extract driving behavior data and future endeavors are discussed.

1. INTRODUCTION

Virtual Reality (VR) as it pertains to this study is a simulation of a specific environment, real world or fantasy. It enables an interactive experience with computer generated 3D models in a particular situation or scene. Virtual 3D cities have been recently available to the roughly 1.5 billion internet users across the world free of charge (www.internetworldstats.com). Two of these programs include Google Earth and Microsoft Virtual Earth [Google 2009], [Microsoft 2009]. Both programs allow users to navigate the globe and immerse themselves in 3D terrain and virtual city environments. On a professional and academic level, these VR programs are vital to the fields of architecture, urban planning, and geography. 3D visualization is an integral part of the workflow and decision making process in these disciplines.

A drive simulator is a hardware tool that can be used for entertainment, training, education, or research purposes. It consists of a vehicle interior, with seat, steering wheel, brake and accelerator pedals. A screen mounted atop the dashboard shows an out-the-windshield view of the drivable VR scene. It is an immersive environment with realistic driving conditions and controls as well as vehicle sounds. Drive simulators can monitor driver behaviors and train drivers to operate different kinds of vehicles in a safe VR environment. For instance, Arizona Department of Transportation uses a simulator to train drivers of large snow plow trucks. Since it only snows two months a year and the job is hazardous, simulator training helps drivers to be better prepared for the winter season [Kihl et al. 2006]. Using additional hardware and software, it is also possible to track a user's head and eye position for studies on fatigue, inattention, and lane changing habits [Ji and Yang 2002].

This paper presents the relatively new concept of conducting research using a driving simulator in conjunction with a realistic VR model and road network of an actual city (Phoenix, Arizona). The city and road network were created using multiple 3D modeling packages, as well as an off-the-shelf driving simulation package. The resulting VR city environment contains traffic simulation capabilities that emulate reality as closely as possible. Computational tools were developed that allow for the extraction of driving behavior data from the simulation for analysis and application on several other research endeavors.

2. BACKGROUND

A VR city model of Phoenix, Arizona was created as part of the Digital Phoenix project. This grant project was collaboration between architecture, urban planning, design, and high speed computing at Arizona State University. The project's goal was to visualize the past, present and future of the core downtown area while creating a tool to better understand growth patterns, policy changes and urban dynamics of the city of Phoenix and surrounding Maricopa County. A 3D model of the metropolitan area was a large part of this project. Urban planning and high speed computing used a software tool, UrbanSim, to run macro simulations of social, economic, and environmental scenarios. This data was integrated and visualized to better understand the impacts of policy choices of the growing urban area. The broader research objective of the project as a whole is to develop computational tools and digital contents to visualize and simulate several urban issues at a 3D immersive theater.

A VR city model of a one mile square of downtown Phoenix was created using a variety of off-the-shelf software packages. It included approximately 500 buildings quickly modeled and textured using photogrammetry. Approximately 70 roads, 200 intersections, and a section of the 1-10 (a major US Interstate that runs through the heart of downtown Phoenix) are also included in the model (*see Figure 1)*. In addition to the existing buildings, many large-scale construction projects scheduled for completion by 2011 were created by hand using renders from several architecture and development firms.

Figure 1. Photo of I-10 traffic flow (top) and realistic VR simulation of I-10 traffic flow (bottom)

The VR Phoenix model was visualized in a traffic simulation and civil engineering software package called UC-win/Road [Forum8 2009]. Popular in Asia, this software had never been used to simulate such an extensive area. A built-in function of the software is automatic traffic generation on a created road network. Each traffic element (cars, motorcycles, buses, and trucks) behaves as an intelligent agent, stopping at red signals and changing speed and lanes corresponding to the surrounding vehicles. These vehicles have the ability to be controlled by an exterior input device. A gaming pad, a gaming steering wheel, or a full-size driving simulator, for example, can be used to control any vehicle in the model for an in-car perspective.

The research donation of a driving simulator by Forum8 allowed the Digital Phoenix Development Team to explore real world traffic scenarios while driving through the city at various times of day. It also enabled software customization to extract driving behaviors from the simulator and the ability to calculate CO_2 emissions, gas consumption, and other statistics from each drive through the model.

3. METHODS

3.1. Creating the VR City Model

Three pieces of modeling/VR software were used in the creation of the virtual city environment: Nverse Photo, 3D Studio Max, and UC-win/Road. In order to save time in both modeling and texturing of buildings, photogrammetry was the chosen method. First, a pilot and photographer were hired to fly over the city of Phoenix at 6000 feet. The photographer was directed to take one photograph straight down, and four angled shots from north, south, east, and west.

The negatives from these pictures were scanned in at 2400dpi, resulting in each .tiff image being larger than a gigabyte. A Precision Lightworks product, Nverse Photo, was used for the rapid 3D modeling of the urban environment [PrecisionLightWorks 2009]. This is a plug-in for 3DS Max that utilizes photogrammetry for quick modeling and texturing. The five aerial images are imported into the software, and linked together by tagging the same reference points in multiple shots. Buildings are quickly and efficiently modeled through an extrusion and automatic texture wrapping process as shown in Figure 2. This allows for an extensive urban area to be modeled in a fraction of the time it would take to do this manually. A two mile by one mile area of downtown Phoenix was modeled, including over 500 buildings in a two week period by two individuals. 3DS Max was used to hand model and texture future buildings.

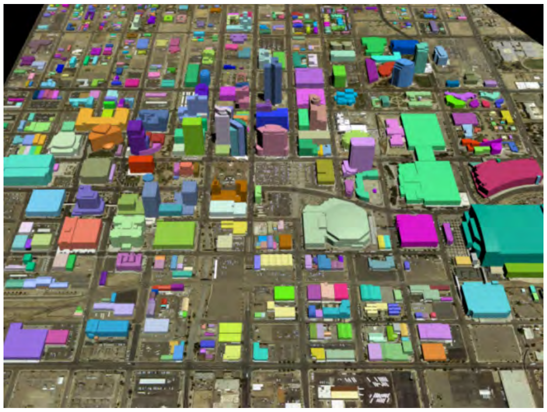


Figure 2. Ground Plane and Building Geometry Modeled Using Nverse Photo

UC-win/Road is a VR environment that allows users to navigate in a 3D space. The software was developed as a traffic simulation and visualization tool for civil engineering. The first step in creating the virtual city environment is to import the ground texture maps. To achieve this, an orthorectified satellite image is tiled and mapped onto the ground plane of the VR space. Adobe Photoshop was used to tile this image. The second step is to import the hundreds of buildings created and textured by Nverse photo. They were exported out of 3dsMax in quadrants of 24 buildings. The final step is to place the quadrants on the correct location of the ground map. UC-win/Road has built in functions to move, scale, and rotate imported objects in 3D space, which makes it easy to line up the quadrants properly (shown in Figure 3).

Figure 3. Buildings Imported into VR Environment (UC-Win/road)

3.2. Road Network Creation and Traffic Controls

UC-win/Road has an intuitive function to create road networks and automatically generate intersections. Once a road network is created, the software will generate 3D modeled traffic on the roads. The traffic will respond to traffic signals, road hazards, and other cars as it progresses through the VR environment.

In a 2D overhead view, the satellite ground map is used as a reference. Vector lines are traced over the satellite image in the approximate location of the roads. Each road can be edited to change the number of lanes, road texture, size of the sidewalk, cutting, banking, and road height. Where these lines cross, default intersections are automatically generated.

Once created, the intersection has many editable features. Lines can be added to the texture to represent stop lines and crosswalks. The intersection can identify when a 3D model of a traffic light has been placed in its vicinity, and will control the light accordingly. The timing of the lights and the signal phases can all be manually set. Drive paths of vehicles through the intersections as well as stopping points are also fully controllable.

Once back in the 3D view, the software will generate fully functional roads with textures. The software comes built in with a vehicle library including many varieties of cars, busses, trucks, motorcycles. The amount of vehicles per hour can be set for the start and end of each road in the model. At the click of a button, traffic is simulated and will drive throughout the road network in the model. This feature, coupled with the importing of 3D building data, creates a lifelike virtual urban environment. By ctrl+alt clicking any vehicle on any road, the user is given an in-car perspective of the 3D scene (see Figure 4). If a driving simulator or gaming wheel is attached to the PC, the user can actually drive the vehicle around the model.

Figure 4. In-Car Perspective While Driving Through the City

3.3. System Integration

Because of the vast amount of texture data associated with the VR city model, a high-performance PC is required to run it smoothly. The PC currently running the model and simulator has dual Xeon processors, 4 GB ram, an Nvidia Quadro fx4500 video card (512 MB dedicated memory), and operates on the Windows XP platform.

The drive simulator connects to the PC with two USB cables for data transfer. The 45" LCD display connects to the computer with a 30ft VGA cable. The drive simulator has built in 5.1 surround sound speakers, with a subwoofer located under the seat to feel vibration. An external Creative soundcard also connects to the PC via USB 2.0, and to the simulator with three 3.5mm jack speaker cables. The simulator components and the speaker amplifier plug into a standard outlet with two universal PC power cables.

Once both the computer and the drive simulator are powered up, UC-win/Road is opened. This connects the two devices, and the simulator turns the wheel automatically to calibrate it. The user sits in the driver's seat. Buttons on the dashboard let a user scroll through options for driving. A predesigned scenario can be run with constraints to a given area and vehicle, or the user can freely drive on any road with any vehicle. Once the selection is made, a first person view of the roadway appears on the LCD screen. The user must start the vehicle with the key by turning the ignition. The user then depresses the brake pedal and puts the vehicle in 'drive' using the gear shift.

Once in drive, the vehicle performs true to life. The simulator is built with the seat, pedals, and steering column directly out of a Hyundai so it looks, feels, and responds like a real car (see Figure 5). The software still has complete modeling and editing functionality using mouse and keyboard even while connected to the simulator.

Figure 5. Drive Simulator in Use

3.4. Implementation of Eco-Drive

UC-win/Road offers several functions to extract driving data. One such function is the Eco-Drive plug-in. Although this plug in is still under development, it provides an accurate result based on a driver's habits while using the simulator. This plug-in is based on a carbon dioxide emission model of an average gasoline powered sedan. The software calculates fuel consumption and carbon footprint using a formula based on the results of a driving log. According to the research by Oguchi, the CO2 emission amount can be estimated using the following formula:

$$E = K_C(0.3T + 0.028D + 0.056\sum_K \delta_k (v_k^2 - v_{k-1}^2))$$

E is the CO2 amount (kg-c), T is the traveling time, D is the travel distance (meters), K is the number of speed measure points, δ k is the speed incremental value (1 or 0), and vk is the speed (m/sec) at the point k (Oguchi 2002). Figure 6 shows the CO2 calculation panel (the Eco-Drive plug in results).

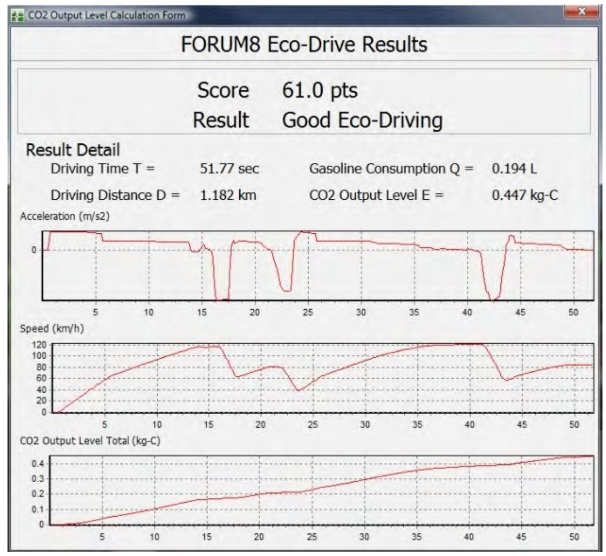


Figure 6. Eco-Drive Results Screenshot

As seen in Figure 6, the Eco-Drive Results chart is shown on the screen after a driver completes a designated scenario. This chart shows a main score (from 1 to 100), and also a result based on this score. The chart also shows readouts of the driving time, distance traveled, gasoline consumption, and CO2 output level. Three graphs are also shown (acceleration/time, speed/time, and CO2 output/time).

The main readout, the Eco-drive score, is calculated using the following formula:

Score = MIN [(Distance(km)/Fuel (L)) : 10] *10

Fuel consumption is determined by the distance that can be driven on one liter of gasoline. Metric measurements are used as this software is developed out of Tokyo, Japan. Eco points will be high if the extent of the consumption of gas is low. The scale of the eco score awards drivers with a low fuel consumption. 10 km/L will equal the highest rating of 100 eco points. This score is calculated by dividing the distance by fuel consumed, and multiplying by 10. The result shown is based on the score, and can be Poor Eco Driving (1-49pts), Good Eco Driving (50 to 79pts), or Excellent Eco Driving (80-100pts).

Similar research has been done using VR models and drive simulators to measure fuel efficiency. One study used a prototype fuel efficiency support tool, which advised participants on how to change their behavior to decrease fuel consumption as they drove in the simulator (van der Voort, Dougherty, and van Maarseveen, 2000). Simax, a company based in Spain, produces a hardware/software solution to test ecological driving. This product enables a user to drive through a generic simulated urban environment, testing for fuel consumption, co2 emissions, car performance, and driver behavior (Simax, 2009).

The digital Phoenix eco-drive study is unique in that the 3d VD urban environment is very similar to the real driving conditions in downtown Phoenix. A section of a major interstate, as well as a square mile of inner city roads are provided to drive through. The NASVA Driving Behavior Assessment also used a 3D scene based on a real world environment in Tokyo, Japan. The goal was to make the VR environment as realistic as possible. This was important in making the simulations more accurate to an area the users had driven in the past.

4. CASE STUDIES

4.1. Digital Phoenix Eco-Drive Testing

The Eco-Drive plug-in was developed in spring of 2009, and the Digital Phoenix urban model was used as a beta test. This model was chosen because of the highly detailed road network and realistic downtown driving environment. It was important to replicate a situation where a driver would be forced to switch between the gas and brake pedals extensively in a relatively short distance. Realistic city driving would also require many other cars on the roadway to incur frequent starts, stops, and distractions.

A scenario was developed to drive down a mile segment of Jefferson St. through the core of downtown Phoenix. This scenario starts the driver in a mid-sized sedan. The driver has to navigate down the road, with traffic volume set at 5000 cars per hour, and follow all traffic signals. The traffic amount is meant to simulate mid-afternoon weekend flow. There are twelve traffic lights that must be navigated through before the scenario ends and the Eco-Drive results window is automatically shown on screen.

In this initial beta testing, three drivers ran through the scenario a total of twenty times each. The purpose of this study was to see if the Eco-Drive results were valid and repeatable. The first driver was instructed to drive the designated course normally, making sure to avoid other traffic and stop at all the signals. The second driver was given the same instructions about traffic and signals, but told to get to the end of the course as quickly as possible. He had to stay under 50 mph while doing so. The third driver was specifically told to be extremely cautious with gas and brake pressure in order to reduce carbon emissions.

As seen in Table 1, the results of the Eco-Drive plug-in reflected the driving behavior of each tester. The first driver, only instructed to obey traffic laws, had a mean score of 66.4. Most of his results were in the range of "good eco-driver" (50-79 eco points). The second driver was told to get to the end as fast as he could while still obeying signals and staying under 50 mph. Because of these restrictions, his use of the gas and brake pedals was hard and jerky. His scores reflected this aggressive style of driving. All 20 of his scores fell into the "poor eco driver" category (0-49 eco-points). Driver 3 was instructed to drive gently as this would impact his overall CO_2 emissions. All but three of his scores were in the "excellent eco-driver" range (80-100 eco-points). His scores reflected the informative nature of the concept of eco-driving. He was made aware of carbon footprint impact by better gas and braking habits, and the results of the plug-in indicated improved scores over the other two drivers.

	Driver 1	**Driver 2**	**Driver 3**
Min	44.1	18.0	72.1
Mean	66.4	31.8	85.8
Max	78.7	41.2	99.9

Table 1. Eco-Score Results of Beta Test

Beta testing of the Eco-Drive plug-in for UC-win/Road was important in the development of this tool. A drive simulator paired with a realistic downtown VR model provided a platform and an environment to further research on informing drivers of bad ecological habits.

4.2. NASVA Driver Behavior Assessment

The same technology and methods used to create the Digital Phoenix model and Eco-Drive (outlined in section 3) were utilized for Japan's National Agency for Automotive Safety and Victim's Aid (NASVA) in creating the NASVA NET system, a VR driver behavior assessment course taken online (NASVA, 2009). This study was done independently of the Digital Phoenix Eco-drive case study, but used the same software platform. Prior to adopting a VR platform for assessing driver safety, NASVA required participants to visit a local office for a traditional pen and paper exam. The nature of assessing driver behavior with this method was inherently flawed, as NASVA associates were unable to determine if a poor exam score actually translated to unsafe driving practices.

Using UC-win/Road, an interactive driver behavior course was developed that grades participants on defensive, hasty, considerate, and eco-driving behaviors. Every action the driver performs throughout the simulation is recorded and analyzed so that appropriate strategies can be suggested for safer driving. The VR assessment provides valuable information that the previous pen and paper exam could not illustrate. The simulation can also be taken either online or at a NASVA office with a driving simulator, gaming controller, or gaming steering wheel (Figure 7).

Figure 7. A Gaming Wheel Can Be Used With the VR System.

The NASVA simulation uses a customized version of UC-win/Road to calculate a participant's safety score. The driver is taken through a series of courses that explore different driving environments: expressways, shopping areas, tollgates, pedestrian crossings, and other potentially hazardous environments (see Figure 8). Following the course, participants receive an aptitude diagnostic sheet based on the data collected from the simulations. If the test is taken at a NASVA branch, guidance is given to the driver individually or in a group.

Figure 8. Drivers navigate through common road environments in NASVA NET simulations

After integrating VR simulations into their driving assessment program, NASVA has experienced several positive outcomes. 1) Time and travel costs can be saved because the test is available 24 hours a day online. 2) Driving characteristics can be analyzed more concretely than with the previous testing method. 3) The integration of Eco-Drive promotes fuel economy in addition to driver safety. 4) Lastly, the new technology and streamlined assessment process has had positive effects on customer service and public relations for NASVA.

For future development of the NASVA NET system, NASVA hopes to integrate more advanced 3D imagery to heighten the realism of simulated driving in hazardous conditions. Further development of UC-win/Road and Eco-Drive will also allow more detailed information to be collected from drivers using the system. Finally, hardware integration with eye-tracking technology will allow greater analysis of driver distraction and attentiveness.

5. CONCLUSION AND FUTURE WORK

This paper introduced a VR city model developed for driving simulation. The VR model covered a one mile by two mile area of downtown Phoenix with all buildings, all intersections with controllable traffic signals, and all streets with the exact lane conditions. The process of making buildings, street networks, and traffic data was described. By using the VR city model and plug-in tool, commute times in different traffic conditions and Eco-Drive results were analyzed.

A future project is to develop a more flexible system to extract specific driving data in VR environments using the drive simulator. For example, 1) the evaluation system of good driving, safe driving, or eco-friendly driving; 2) a cognitive system to analyze driver vision by integrating with eye tracking systems; and 3) interactive traffic control systems to change the traffic amount interactively without restarting.

VR simulations and computer models have been created to visualize design output for only the last decade. Some projects use these kinds of models to show the results of scientific analysis such as energy effect, aerodynamics, and carbon footprint information. However, the human cognitive aspect has not yet been analyzed well using this framework. A future goal is to innovate by implementing the VR framework outlined in this paper to bridge the gap between the cognitive sciences and the design field.

Based on successes with NASVA and the Eco-Drive system, an "Eco Driver Awareness Program" is currently being developed to educate drivers on their current habits and their impact on CO2 emissions and fuel consumption. Although the relationship between cautious driving and fuel conservation have been well documented, the goal of this program is to provide drivers with a more tangible score that they can work to reduce in their daily habits.

ACKNOWLEDGEMENTS

The authors wish to acknowledge the support received from Forum8 Co. Ltd. This project is sponsored by the Digital Phoenix project at Herberger Center for Design

Research at the College of Design, Arizona State University.

REFERENCES

Forum8 Co. LTD. 2009. "Ecological Drive Plug-in." Retrieved May 16, 2009 from http://www.forum8.com/topic/up78-p11-e.htm

Forum8 Co. LTD. 2009. "UC-win/Road product information." Retrieved May 2009 from http://www.forum8.com/english/uc-win/ucwin-road-e1.htm#30403

Google. 2009. "Google Earth: Explore, Search and Discover." Retrieved May 6, 2009 from http://earth.google.com/index.html

Ji Q., and Yang X. 2002. "Real-time eye, gaze, and face pose tracking for monitoring driver vigilance." Real-Time Imaging, 8 (5): 357-377.

Kihl, M., Herring D., Wolf, P. McVey, S., and Kovuru, V. 2006. "Snowplow Simulator Training Evaluation." No. 585. Phoenix, AZ: Arizona Department of Transportation.

Los Alamos National Laboratory. 2009. "TRANSIMS: transportation analysis simulation system." Retrieved April 9, 2009 from http://www.ccs.lanl.gov/transims/

Microsoft. 2009. "Microsoft Virtual Earth: The integrated Mapping, Imaging, Search, and Location Platform." Retrieved May 6, 2009 from http://www.microsoft.com/virtualearth/

NASVA. 2009. "National Agency for Automotive Safety & Victims' Aid, Safety Guidance Department." Retrieved May 16, 2009 from http://www.nasva.go.jp/mamoru/english/index.html

Oguchi, T., Kattakura, M., and Taniguchi, M. 2002. "Carbon Dioxide Emission Model in Actual Urban Road Vehicular Traffic Conditions." Journal of Civil Engineering No.695, IV-54:125-136.

PrecisionLightWorks. 2009. "Nverse Photo product information." Retrieved May 6, 2009 from http://www.precisionlightworks.com

Simax, 2009. "SImax ECO-driving Simulator." Retrieved May 16th, 2009 from http://www.simaxvirt.com/Products/EcoDriving.aspx?lang=2

Van der Voort, Dougherty, and Van Maarseveen, 2000. "A Prototype Fuel-efficiency Support Tool." *Transportation Research Part C* **9** (2001), pp. 279–296.

Superstructure

Associative modelling of Multiscale Fibre Composite Adaptive Systems

Maria Mingallon[1], Sakthivel Ramaswamy[2], Konstantinos Karatzas[3]

Architectural Association School of Architecture
36 Bedford Square, London WC1B 3ES
www.aaschool.ac.uk

maria.mingallon@arup.com, sakthivel.ramaswamy@gmail.com, k.karatzas@gmail.com

Keywords: responsive architecture, biomimicry, form finding, fibre optics, shape memory alloys.

Abstract

This paper reports on the digital simulation involved for the development of a smart material system capable of sensing and adapting to new environmental conditions. Its technology is based on incorporating fibre optics for sensing and shape memory alloys for actuation into a fibre composite material. Material, geometry and structure together with sensing and actuation had to be computed simultaneously requiring the use of in-house written codes as well as finite element models. The results demonstrate the potential of simulation for the development of smart materials and their application into full scale architectural constructions.

1. INTRODUCTION

'Thigmo-morphogenesis' refers to the changes in shape, structure and material properties of biological organisms that are produced in response to transient changes in environmental conditions. Architectural structures endeavour to be complex organisations exhibiting highly performative capabilities. They aspire to dynamically adapt to efficient configurations by responding to multiple factors such as functional requirements and environmental conditions. Existing architectural smart systems are aggregated actuating components externally controlled, whose process of change is essentially different from that of thigmo-morphogenesis, where sensing and actuation are integrated within the element itself. Emulating such morpho-mechanical system with sensors, actuators, computational and control firmware embedded in a fibre composite skin was the aim of our thesis. A thorough investigation on the newest fibre materials available in the market provide us with a firm starting point from which the geometry and structure of the system could be developed. The technology behind our material system includes the use of shape memory alloys as actuators, fibre optics as sensors and glass fibres as structural elements (see Figure 1); these are held together by a resin based polymer [1-5].

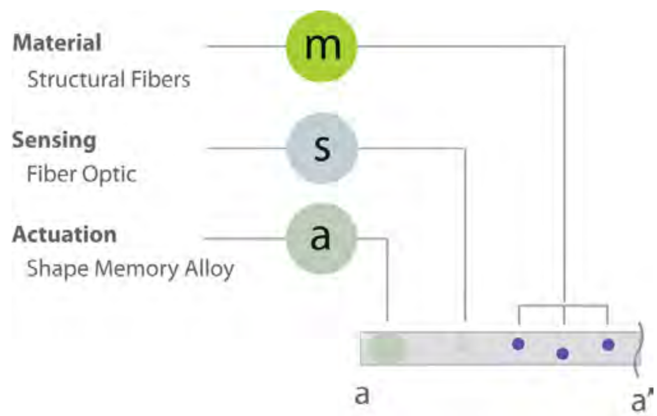


Figure 1. Material configuration diagram.

2. BIOMIMICRY

Living cells have been proved to be stimulated by pressure to increase and multiply: pressure and by extension strain, are direct encouragement of growth. D'Arcy Thompson has extensively written about this concept which he named 'growth under stresses'; he used a large number of examples present in the natural world such as the discussion held on bone constructions [6]. Bone tissue is first deposited as woven bone in a disorganised immature structure present in young bones and in healing injuries. Woven bone is weaker, with a small number of randomly oriented coarse collagen fibres; it forms quickly and it is used by the organism temporarily while it is rapidly replaced by lamellar bone, which is highly organised in concentric sheets. Lamellar bone is stronger and needs to be able to resist the main type of forces which will act on the bone throughout its life: bending, torsion and compression. The first two forces require the fibres to run diagonally with different angle orientations and in opposite directions, forming a dense outer layer. Due the compression forces exerted by the weight of the body, cancellous bone tissue is developed in the interior of mature bones following the direction of functional pressure according to Wolf's Law (see Figure 2).

Wolff's law is essentially the observation that bone changes its external shape and internal architecture to produce minimal-weight structures that adapt in response to the stresses acting upon it. The process starts with trabecular tissue being secreted and laid down fortuitously in any direction within the substance of the bone. If it lies in the direction of one of the lines of principal stress, it will be in a position of comparative equilibrium. However, if it is inclined obliquely to the lines of principal stress, the shearing force will tend to act upon it and since fibres are linear elements which cannot withstand shear forces, the fibre will be moved away until it coincides with a line of principal stress. Bone tissue is continuously being recycled and it is capable to adapt to new loading conditions.

Bone growth and most importantly 'growth under stresses' recap the structural concept behind our material system where both external and internal configurations are the result of the forces acting upon them.

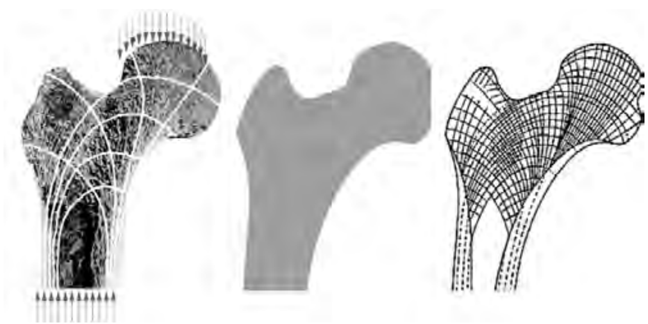


Figure 2. Principal lines of tension and compression in a femur. Trabecular bone tissue spreads in curving lines from the head to the hollow shaft of the femur.

3. GEOMETRY

The geometry was approached from two different scales: local and global. The local geometry emerges from the topological definition of a single cell as the smallest unit within the material system. The proliferation of these single cell topologies had to respond to the set of rules dictated by the material system. The following subsections discuss the intricacies of modelling such configurations which required the development of in-house written codes in both visual basic and vb.net, as well as the use of finite element models and associative modelling software such as Grasshopper for Rhinoceros.

3.1. Global Geometry

In the global geometry the strategy was to follow the rules present in the formation of bones which are capable of achieving such lightweight but rigid constructions. Two main rules are: fibre directionality and fibre density. Fibre directionality accounts for the orientation of the internal loads that would need to be withstood by the material: 'force travelling paths'. Fibre density responds to the variable magnitude of those forces across the bone, being denser where there is a concentration of load as higher load will need to travel through. These mechanisms act in the micro level, but are also responsible for the overall form. Smooth or sharp curvatures in the macro scale are the result of the fibre organisation in the micro level.

Our global geometry aims to reproduce this system. In the micro level fibres are oriented following the direction of the lines of principal pressure which by extension are the lines of principal strain also known as principal stresses. In other words, the direction of the fibres responds to the direction of the internal forces present in the material. Also in the micro level, fibre density is increased or decreased according to the scalar value of the forces locally. There would be areas where there is a concentration of load and thus, a higher density of fibres will be supplied. On the contrary, material should be saved in areas where loads are lower. The macro level would be the result of the two mechanisms explained above.

This process was simulated digitally through the use of two algorithms written to fit the specifications of the project. Due to technical requirements the theoretical process described above needed to be reorganised. The process started by 'form finding' the overall form, for which the optimum fibre organisation will be subsequently found.

The first algorithm consists of an adaptation of the dynamic relaxation problem. Usually dynamic relaxation is used with materials that can only withstand tensile forces such as membranes. The variation introduced here is possible due to the use of a material capable of compression. Arches can be designed to experience only axial forces if their geometrical layout is made to coincide with the inverse of the bending moment diagram. For example, a catenary is the curve resulting when a cable is hanged from its two extremes, i.e. only self-weight is involved. While in this configuration the cable can be made rigid by applying a material capable of taking compression; if the cable is then flipped vertically, it will form an arch which would not have any bending moment. These configurations are called 'anti-funiculars' and constitute the base of form finding techniques. This logic was extrapolated to 3D and the result was a digital form finding algorithm for the formation of shell structures.

For this particular case, we decided to apply the algorithm to the design of a pavilion. Such structure could be simplistically seen as a roof and a vertical support. Roofs are usually formed by horizontal spanning elements which require vertical supports to bring loads to ground level. The resulting morphology integrates both structural roles within one continuous surface (see Figure 3).

The simulation starts with a horizontal plane located at the height of what would have been a planar roof. Following the strategy previously discussed on 'anti-funicular' structures, vertical loads are introduced in the opposite direction as to how they would actually act in the final structure. Therefore, gravity is applied upwards while reactions coming from the vertical central column are applied downwards. The points around the perimeter are fixed. Finally, the algorithm starts running and the roof starts curving as a result of the loads; there is a time when the points defining the surface start oscillating and a few seconds later, equilibrium is found. The roof surface has curved in response to the applied loads and it now integrates both roof and column within a single continuous surface (Figure 3a).

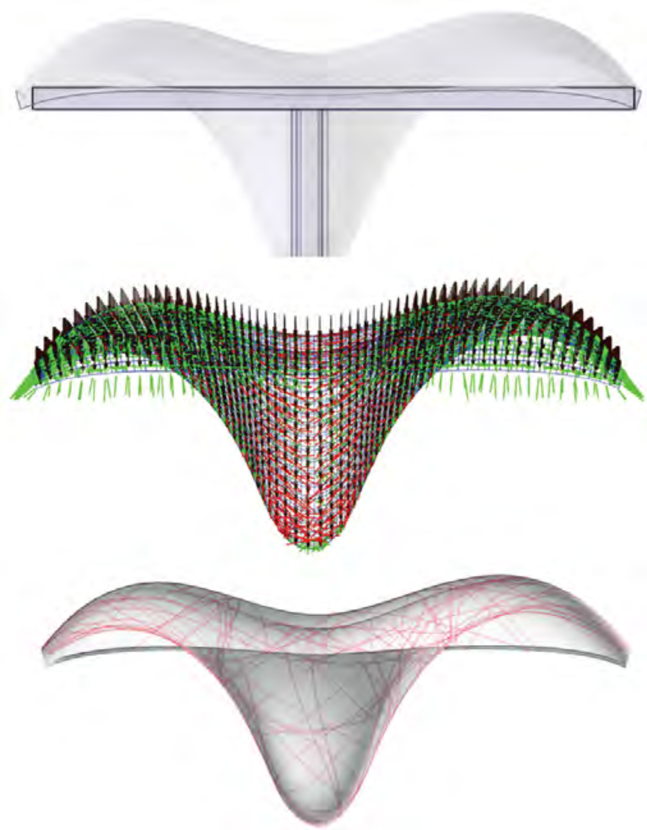


Figure 3. (a) Top, diagram showing the comparison to a conventional structure using horizontal spanning elements and vertical central support; (b) middle, principal stresses vectors from FEM in GSA; (c) bottom, resultant fibre distribution on form found geometry.

The second algorithm was written to simulate the growth of the fibres. In bones, fibres are distributed following the direction of the lines of principal pressure whose equivalent in structural engineering are the principal stresses. The principal stresses are obtained by rotating the 'Cauchy stress tensor' into such a coordinate system that shear stresses are zero and only normal stresses remain. Shearing forces are those which produce 'angular distortion', causing particles to slide over one another. While shearing stresses can by no means be suppressed, the danger of rupture under shearing stress is lessened the more we arrange the fibres along the lines of principal stresses since there is no shearing stress along or perpendicular to these lines.

The digital simulation of the fibre growth combines the use of a Finite Element Model (FEM) with an algorithm written in Visual Basic for Rhinoceros. The process starts by calculating the principal stresses on the previously form found structure which is exported into GSA (in-house Arup structural analysis software) and converted into a Finite Element Model (see Figure 3b). The values of the vector principals obtained in GSA are exported into an Excel file which is later computed by the algorithm. Once these preparations are finalised, the algorithm starts running. First, it selects the surface previously form found and continues by importing the vector principals' data stored in the Excel file. Before fibres can be laid, the algorithm needs to sort out the lists of data by direction and magnitude. Fibres are then laid by 'linking' points where the direction of the principal stresses is similar. In addition, the higher the stress value the greater the number of fibres that are laid.

The result is a compendium of fibres wrapped around the periphery of the central column previously modelled through form finding to simulate a traditional vertical support. Fibres curve and cross orthogonally with one another, following the lines of principal stress formerly extracted from GSA. They distribute around the cylindrical trunk, becoming denser the closer they are to the end of the support. In the contrary, fibres spread out as soon as they abandon the column zone, travelling across the extension of the roof and fading around the corners (Figure 3c).

The distribution of fibres together with the overall form created a single continuous surface featuring different degrees of flexibility. The corners were found to be the most flexible zones in comparison with the area surrounding the vertical support. This was essential in order to integrate the actuation technology within the whole system. The geometry both in local and global levels had to be conceived in such a way that actuators could be 'activated'. The corner areas being more flexible than the surrounding zones were intended to have intense actuation, with a range of different apertures and closures within the local geometry. On the contrary, the area around the column required higher densities and thus, actuation had to be restricted to the minimum.

3.2. Local Geometry

Fibre composites are characterised for achieving high ultimate strength values which lead to the production of extremely thin surface-like geometries. These are very useful for the development of shell structures where all the external loads travel across the surface and thus, there are not out-of-plane forces to be considered or these can be neglected. For those cases where out-of-plane effects cannot be neglected, reduced thicknesses would not contribute to transferring the loads and thus, the strategy needs to be reconsidered. Instead of increasing the surface thickness which will lead to a bulkier design, a more convenient strategy is to carefully reassess the geometry. The aim must be to generate 'structural depth' in the out-of-plane dimension; in other words, the material should be re-distributed away from the hypothetical 'neutral axis' originally positioned half way through the surface thickness. This arrangement can easily be achieved using corrugations where the stiffness of the structure mostly depends on the height of the 'waves'. Varying the local height of the corrugations would therefore lead to a structure with stiffer differentiated areas which ultimately contribute to the structure as a whole. In essence we are looking for a methodology to create a space truss out of a compendium of surfaces -rather than individual members- where the structural height varies depending on where loads are applied.

Our investigation of space trusses formed by continuous surfaces led us to the work of Erwin Hauer, one of the main contributors to the 1950s movement known as 'modular constructivism'. His sculptures are particularly interesting for achieving infinite patterns of repetition which are used to create limitless, basically planar, screen-like formations, employed to produce multidimensional structures. The work of Erwin Hauer inspired the conception of our topology [7]. The aim was to turn a thin surface into a tissue-like webbing pattern interconnecting convex and concave surface curvatures in multiple layers. A set of thin surfaces prone to buckling was therefore transformed into a robust structure by means of altering its topological definition; a series of interwoven surfaces defined now a continuous space truss capable of withstanding a wider range of loading conditions. Finally, after pursuing an extensive study of the topologies developed by Erwin Hauer, we developed a variation built upon the following logic (Figure 4):

1. The first level consists of two hexagonal grids displayed in two levels: top and bottom and with certain offset in between each other (Figure 4a).
2. These grids are then linked in between with two set of 'tripods'; legs extending from top grid to bottom and vice versa (Figure 4b, c).
3. The tripods hold the control points defining the edge curves of the final surface (suture curves) (Figure 4d). The entire topology of these curves is based on the spatial arrangement of the tripods and their geometrical association with the hexagonal grids.
4. Removing the auxiliary grids helps visualising how the suture curves construct the space-frame. Lofting in between these curves renders the final interwoven surface (Figure 4f).

The next step was to establish the strategy of 'adaptation' at the local level. This was carefully structured into different fields of action or what we have referred to as 'actuation variables' which constitute the set of parameters that can be altered within the material system in order to adapt to novel situations:

- size of openings, which would regulate the light and air transferred through (Figure 5),
- overall curvature, which allows the material system to adapt to doubly curved surfaces if necessary (Figure 6),
- and finally, the height of the undulating pattern which from the structural point of view is the structural depth of the space-truss (Figure 7).

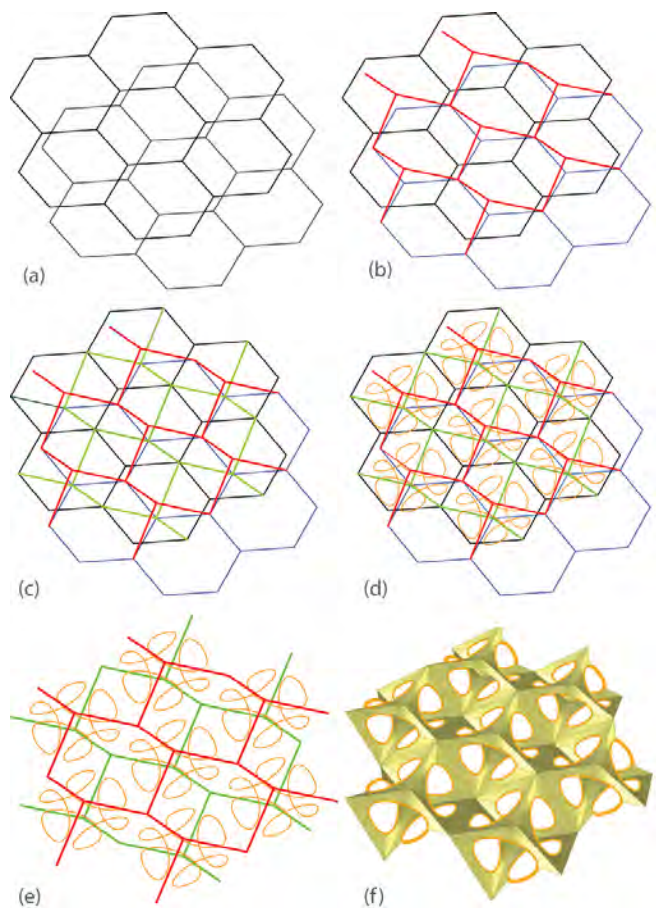


Figure 4. Local geometry topology; (a) double layer hexagonal grid; (b,c) tripods; (d) suture curves; (e,f) final geometry.

The 'suture curves' are the main element defining the final surface. Altering these curves or the points that control them would alter the geometry of the system at the local scale. Therefore, we decided to strategically locate the actuators (SMAs) along these curves, establishing the negotiation logic in between Hauer static topologies and our fibre composite adaptive system. Varying the size of the openings will result on increasing or decreasing the amount of light and air transferred (Figure 5). The topological definition of the 'suture curves' allows for such variation by simply altering the position of their control points along the tripods defining each cell.

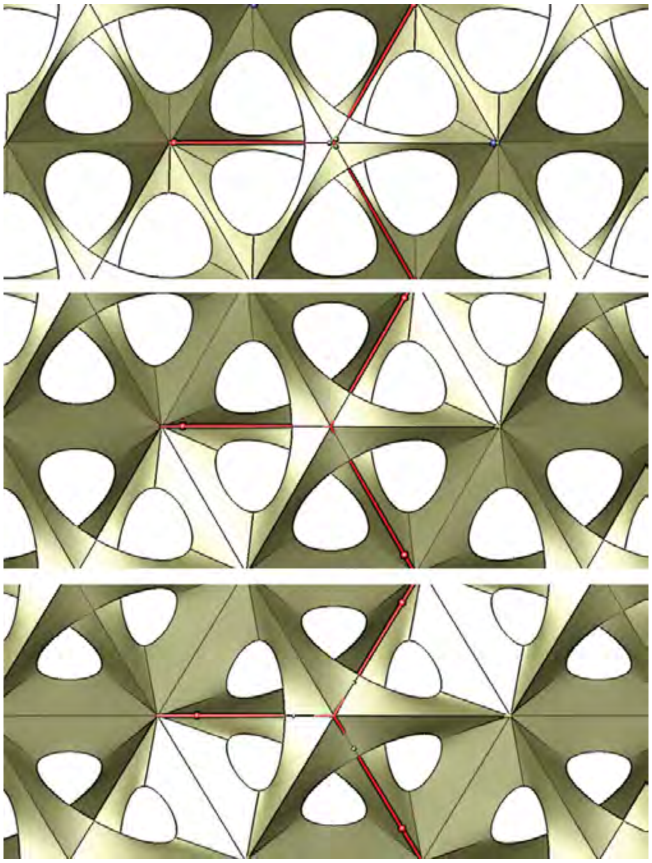


Figure 5. Various opening sizes for different lighting and temperature conditions.

This results in altering the curvature locally and hence, reducing or increasing the size of the openings. However, in order to avoid unrealistic curvature changes, the movement of these control points had to be limited by grouping them in four sets of three. The first set is highlighted in red in Figure 6. The second, third and fourth sets are shown in black, green and blue respectively. The movement of the red set is coupled with the movement of the blue set and thus, when the height is reduced the red set of points travel downwards along the tripods while the blue set is travelling upwards. The same rule is applied to the black and green sets of points. In addition, the length of the tripod legs is maintained through the process and hence, when the height is reduced the horizontal spacing of the legs increases (tripod expands) and viceversa when the height is increased (Figure 6).

On the other hand, the overall curvature is controlled by the distribution of the grid points in the space. Through a controlled distortion of the points forming the primary hexagonal grids, the overall spatiality of the space-truss can be altered. This 'controlled distortion' is realised by the use of attractor curves which mark the changes of curvature that the originally flat space-truss needs to undertake to transform into a doubly curved geometry. This process starts by extracting the isocurves from a 'mother' surface which defines the overall form of the global geometry. These isocurves are then set to be the attractor curves. The points on the grids travel in the z direction until their distribution matches that of the attractor curves. Also, the closer the points are to the curves the higher would be the attracting force. In addition, the effect of this force can be altered by increasing or reducing the attractor factor (Figure 7). This tool is vital for the combination of the local and global geometries and it also enhances the adaptability of the entire system to a wide range of global forms.

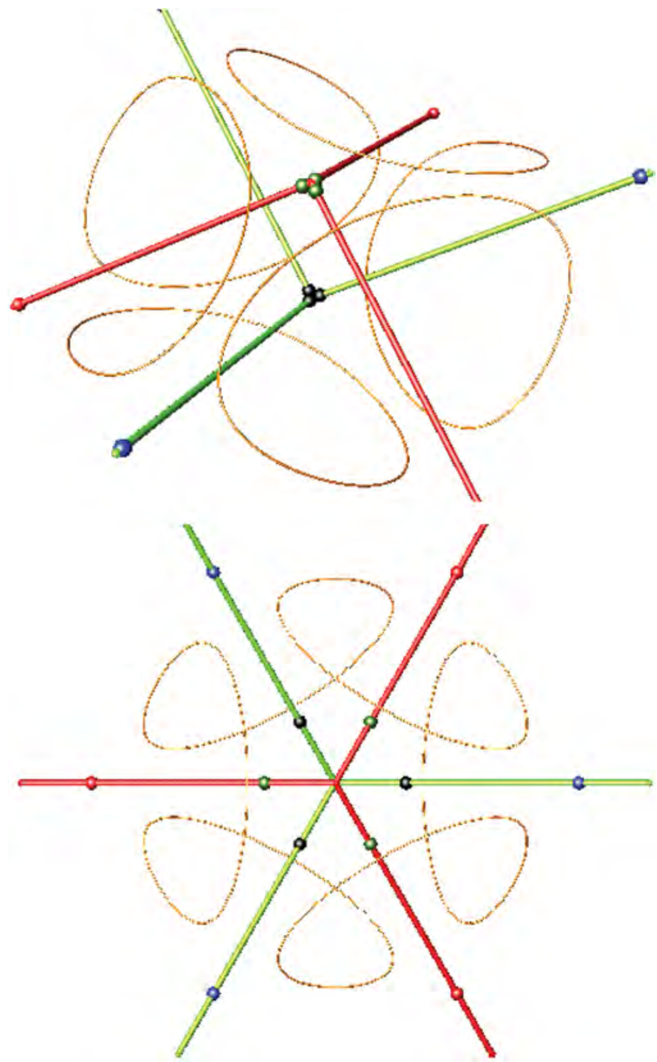


Figure 6. Perspective and top view of a single cell with the two set of tripods and the location of the control points.

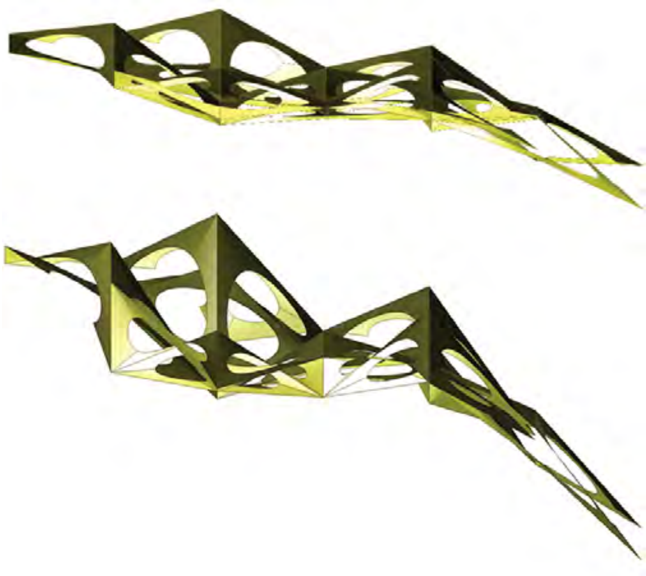


Figure 7. Local geometry following principal curvature lines of a given surface, in this case extracted from digital form finding.

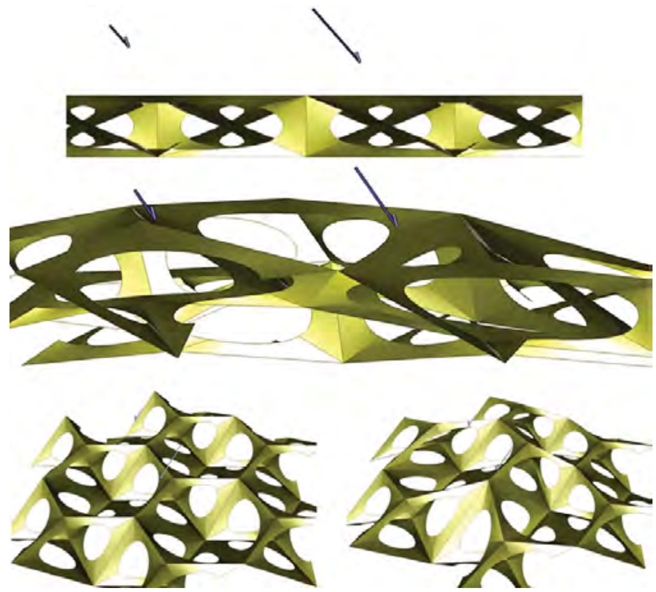


Figure 8. Initial and resultant geometry as a response to the applied loading. Note how the structural depth has locally increased where the point loads are applied and thus, adapted to the new loading conditions.

Finally the algorithm also allows for the manipulation of the height of the space-truss or in other words its structural depth. This is perhaps the simplest mechanism and it consists of varying the spacing in between the two hexagonal grids in the z direction. This coincides with the height of the tripods which is also coupled with the opening and closing of the tripod legs. The logic behind a change on the structural depth is regulated by the external forces that act upon the system. For instance, if there is a wind load acting on a certain area, the sensing devices would identify this and the actuators would be set to act accordingly (Figure 8). The actuation will increment the structural depth locally where the action has been identified. In the digital model this has been simulated with the help of attractor points. The affected points of the grid will re-organise accordingly to provide a greater structural depth in the direction of the load vector which would be proportional to the magnitude of the force inherited. The structural depth of the space-truss grows bigger towards the mid-span areas of the roof, yet capable of adapting its height in response to any change in the loading conditions.

The resulting structure is a space-truss made out of an interweaving of continuous surfaces which in the local scale acquires stiffness from combining convex and concave curvatures (Figure 9).

4. FURTHER RESEARCH

The geometrical definition of the topology has a versatile potential for efficient functioning in several environmental contexts. The system could be ‘activated’ by energy acquired from the environment, rather than using external sources. Hot and dry climates which have a considerable diurnal variation in temperature would be the most appropriate to exploit the environmental energy for the functioning of the adaptive system since the surface temperature of the shape memory alloys could easily reach an actuation temperature of 35°C. The diurnal temperature variation could be exploited and easily incorporated into the digital model in order to maintain a comfortable micro climate inside buildings. This strategy would help with minimising the need of artificial energy and external control.

5. CONCLUSIONS

The proposed fibre composite adaptive system possesses multiple organisational levels with different assembly logics, which contribute to its emergent behaviour and integrated functionality. An increase in the ambient temperature alters the layout of the suture curves to ‘open’ a hole through which air can circulate. Similarly the openings are closed when there is a need to maintain a certain temperature inside the pavilion. If the wind increases or changes direction, the structure re-organises adopting a more efficient configuration against the new loading conditions. In essence, form, structure, geometry, and behaviour created a cohesive synergetic whole and constituted the most important consideration for the construction of the digital model. Such ambitious requirements required the combination of several software packages initially intended for different professional disciplines with in-house written codes such as the form finding algorithm and the fibres growth script. The result is a system which has the potential to adapt and self-organise efficiently to transient changes in the environmental conditions, illuminating the way towards the development of smart architectures.

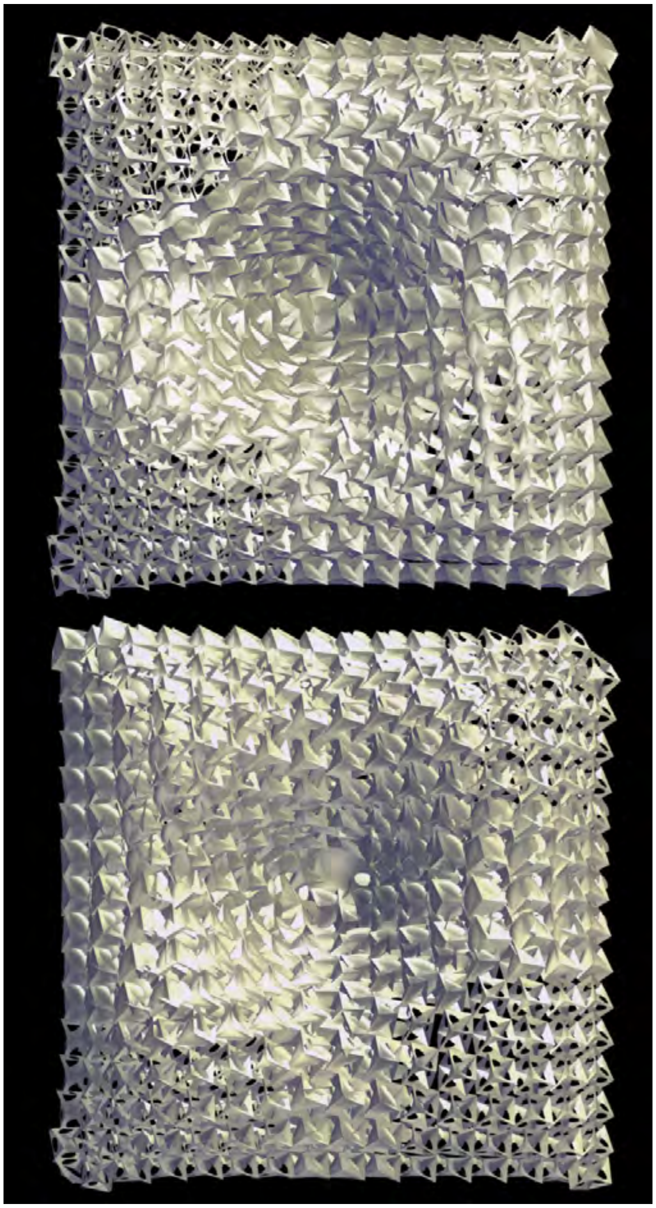


Figure 9. Combination of local and global geometries. Top view of final design showing different configurations for the roof structure.

Acknowledgments

We would like to thank our tutors Mike Weinstock for his involvement and invaluable guidance, and Michael Hensel for his help and encouragement. George Jeronimidis, for his immense help in developing our research and performing the physical experiments. We would also like to thank Stylianos Dritsas, for his essential input in evolving the geometry. Finally, we would like to thank Prof. Kevin Kuang and W.J. Cantwell for sharing their knowledge and research on shape memory alloys and fibre optics.

References

[1] Kuang, K. S. C., Cantwell, W. J., "The use of plastic optical fibre sensors for monitoring the dynamic response of fibre composite beams", Measurement Science and Technology, Volume 14 (2003), pp. 736-745.

[2] Kuang, K. S. C., Quek, S. T., Cantwell, W. J., "Morphing and Control of a Smart Fibre Metal Laminate", Sampe Europe, 30th, International Jubilee Conference and Forum, Session 7C (2009), pp.481-486.

[3] Kuang, K.S.C., Cantwell, W.J., Zhang, L., Bennion, I., "Process monitoring of aluminium-foam sandwich structures based on thermoplastic fibre–metal laminates using fibre Bragg gratings", Composites Science and Technology, Volume 65 (2005), pp. 669–676.

[4] Kuang. K.S.C., Quek, S.T., Tan, C.Y., Chew, S.H., "Plastic Optical Fibre Sensors for Measurement of Large Strain in Geotextile Materials", Advanced Materials Research, Volumes 47-50 (2008), pp. 1233-1236.

[5] Kuang, K. S. C., Cantwell, W. J., Scully, P. J., "An evaluation of a novel plastic optical fibre sensor for axial strain and bend measurements", Measurement Science and Technology, Volume 13 (2002), pp. 1523–1534.

[6] Thompson, D.A., "On Growth and Form", Ed. Cambridge University Press, First Edition 1961; reprinted 1994.

[7] Hauer, E., "Continua--Architectural Screens and Walls", Princeton Architectural Press, New York, 2004.

Illustration Credits

Figure 2: Diagram constructed using an original illustration from: Thompson, D.A., "On Growth and Form", Ed. Cambridge University Press, First Edition 1961; reprinted 1994, page 232, Fig. 100 "Head of the human femur in section, from a photo by Professor A. Robinson".

*Note: Any other figures and diagrams presented herein are original work of the authors of this paper.

LibreArchi: Library of Interactive Architectural Models Containing Exploratory and Didactic Simulations

Ivanka Iordanova & Temy Tidafi

Université de Montréal, Canada
ivanka.iordanova@umontreal.ca, temy.tidafi@umontreal.ca

Keywords : Design Exploration, Interactive Models, Simplified Simulation, Library, Passive Performative Architecture.

Abstract

This paper presents a library of interactive architectural referents, LibreArchi, which can play a twofold role, exploratory and didactic, during the conceptual phases of an architectural project. The library is composed of interactive models of chunks of architectural know-how represented in a multimodal way: through an interactive model including simulations, pictures, text, video, etc. The concept of distributed intelligence provides a methodological basis for LibreArchi in two ways: a precedent or a new project can be informed by or include several different chunks of know-how, thus allowing for higher complexity and flexibility; and being open for qualified input of interactive models of know-how, it is constructed by the community. This paper is an invitation to reuse and share.

1. INTRODUCTION

The optimization of energy consumption of a building is now often a subject of interest in architectural education. This article offers a didactic and exploratory shared environment which integrates knowledge and know-how from arts and sciences linked to architecture. It is possible to include here structuring, structural and environmental forces, which could influence the future building in order to optimize its performance, but also to ameliorate the general quality of the living environment. An analogy can be made with Nature here, which in its role of an 'architect' optimizes the forms in respect of material and energy use. As puts it d'Arcy Thompson in "On Growth and Form" (1917), the form of an object is a 'diagram of the forces' which have acted or act on it. Thus, the climatic and environmental considerations should be seen as part of a larger group of abstract non geometric parameters which can influence the architectural form. This implies a holistic approach to architectural conception (Abalos 2007, Seebohm 2007; Tidafi 2007).

2. DIGITAL SIMULATION AND THE CONCEPTUAL PHASE

The digital means, now ubiquitous in architecture and engineering as well as in universities, have passed through different development phases by now, but are still not adapted for integrated design (Seebohm 2007). The tests on structure and performance are normally performed by specialists, and after the architectural conceptual phase. This way, they cannot play a role as 'form-finding' approach and contribute for innovative and optimal design.

The term 'performative architecture' was introduced to indicate the use of the building performance as a dominating principle of design (Kolarevic 2003, Oxman 2006). This strategy to architectural design is based on simulations guided by abstract data and information. These can come from rather different 'worlds': financial, spatial, social or cultural, or yet: structural, thermal and acoustic, etc. In consequence, we can expect that the higher requirements to building performance will now trigger a redefinition of the expectancies to design, its process and practice.

In the last decade more and more authors reported form-finding techniques based on abstract (non-geometric) data. Digitally, this is introduced through parametric models of performance or simulations. Thus, Turrin et al. (2009) use a parametric model to investigate the structural geometry and, according to the responsiveness required to support passive heating and cooling, also integrating openable modules in the structural morphology. According to the authors: "while aiming at a performance oriented conceptual design, parametric modeling serves to support the design exploration on an architectural element". And also: "The choice for using parametric techniques is based on their capacity for creating design alternatives while managing the complex system of relationships". With the same objectives, Attar et al. (2009) report realizations of physics-based generative design, which, according to the authors, can represent "a paradigm shift from the traditional primacy of object to an exploratory approach of investigating interacting elements, interdependencies and systems". The same arguments have brought us, some six years ago, to the beginning of the composition of a library of interactive models of architectural know-how serving as a didactic tool and as a space for sharing.

3. LIBRARY OF ARCHITECTURAL REFERENTS

Taking into account architects' preference to work from precedents and metaphors (here commonly named 'referents'), we propose chunks of environmental know-how encapsulated in interactive referents. They represent either architectural precedents or objects and phenomena coming from other spheres of knowledge or from the nature. For easier access and use by the architectural studio students, they are organized in a digital library thus creating a complementary learning environment. The digital methods used to model the encapsulated know-how can be explicated. This represents a fundamental difference compared to the existing databases of precedents which do not communicate the methods of creation of the architectural object.

The referents which could take place in such a library are extremely numerous. Various architectural themes can be represented (Figure 1).

Some environmental categories presently available in the library are:

- wind simulation: visualization of the direction and the speed of the wind (through particles); defining the 'wind shadow' created by buildings;
- deformation of a surface under the force of the wind (Figure 2): This procedure can be used for obtaining a more aerodynamic form;
- simulation of snowdrift in front and behind a building (depending on wind's direction), see Figure 3a.
- sound propagation and acoustics in a built space depending on its form (materials are not taken into consideration in these models);

As the interactive precedents provide a direct link between the form of an architectural object and its behavior in respect to phenomena from the environment, the library helps designing Passive Performative Architecture where the first optimization of performance comes from the form.

4. INTERACTIVE SIMULATION MODELS

These models are not entirely scientific but provide a reasonable approximation of the simulated phenomenon and its effects on the built environment. Structural themes of exploration are proposed too (Figure 1): some are based on the method of concrete structural elements formation in tissue molds under the force of gravity (West 2001); others are metal triangulated structures, etc. Advanced digital techniques non-designated for architecture are used for creating these models, for example particles and reflectors (for modeling wind, sound and snow), dynamics (for the snow, as well as for the simulation of the molding of concrete), kinematics (for some structures), etc. This approach increases the complexity of elaborating the model but gives it a greater liberty for various interpretations and usages. These have been identified as conditions for creative use of the digital medium (Fischer 2007). The interaction of a student with a model is facilitated by the architectural terminology used for the names of the parameters in the models of LibreArchi. After manipulations and explorations, a re-used model can be directly integrated in the project at hand (or a simulation can be performed directly on it).

The Library of interactive architectural referents (LibReArchI) is conceived as an open space for sharing referents and know-how. The web-page interface of the library proposes a general view of graphical representations of the models. In our opinion, it is not necessary to classify the referents by theme because an unintentional order would better stimulate creativity. A search function can help finding relevant referents according to desired criteria or keywords.

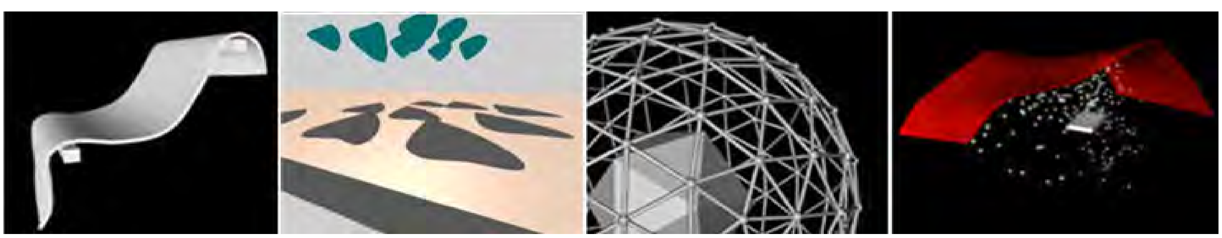


Figure 1: *Examples of different architectural themes represented in the library: form (created under the weight of the material); shadows (minimize or maximize the shadowed area); structure; sound propagation, acoustics.*

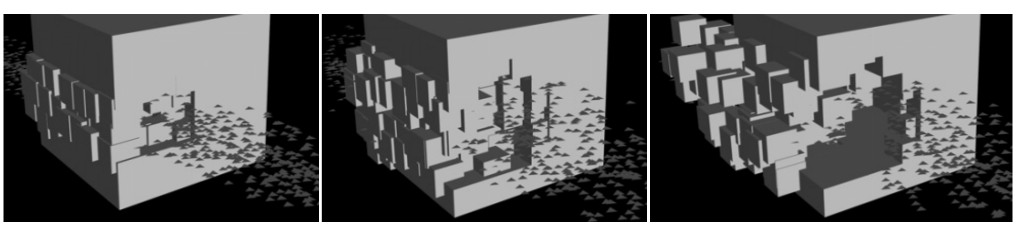


Figure 2: *Deformation of an object under the force of wind (model from the library)*

5. VALIDATION AND VALUES TO THE COMMUNITY

LibReArchI was used a learning assistant in architectural studios with students from 3rd year or in masters' of architecture. The examples presented in this paper are taken from a masters' studio with term project (Museum of Inuit Art in the Great North) where the environmental and climatic aspects were playing a major role. The library was introduced into the program of the studio through preparatory exercises on the different climatic aspects and their interaction with the building and its occupants (sun, shadow, wind, sound, cold). The timing of the exercises coincided with the beginning of the work of the students on their term projects. This way, the proposed methods were directly assimilated and used for the project. Some aspects were more developed by the students, depending on the different concepts chosen for their projects (Figure 3b). The referents concerning aerodynamics, link to wind and snowdrift were the most appreciated ones, because there are hardly any other accessible tools for their simulation during the conceptual phase of the project.

A more profound and exhaustive validation of LibReArchI was realized as a doctoral research (Iordanova 2009). It was found that, through parametric explorations and watching the explicative videos, the students acquire a good understanding of a referent. There were numerous cases of re-use of know-how encoded in interactive models for the purposes of a new project. The contribution of the parametric and generative methods proposed in the library of referents to the creative processes was important. It plays a role in the transfer of domain know-how too, but to a limited extent, given the relatively small number of referents in the prototype.

Environmental approaches and simulations were seamlessly integrated in the design process, given that the students did not have to interrupt the process in order to change software or conceptual environment (Figure 4). Most of the students tested the simulations on their projects and took them into consideration for the development of the form of the building. Some of them influenced the conceptual form through the impact of a given environmental phenomena on it, with the purpose to obtain a better performance.

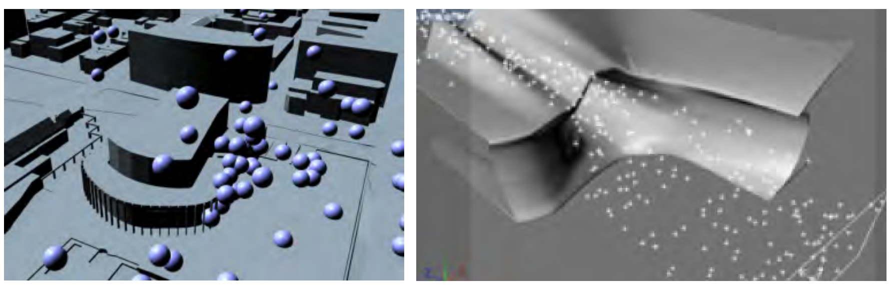


Figure 3: *(a) Snowdrift forming (model from the library);*
(b) formation of an opening for the wind turbine in a building (exploration of a student)

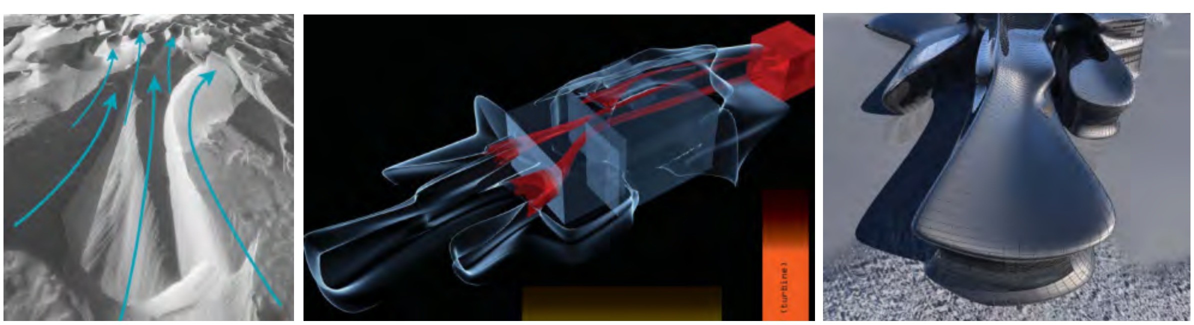


Figure 4: *Project of a Museum of Inuit Art in the Great North (V. Audy):*
inspiration; conceptual form with integrated wind turbine; final project.

The video recording of design sessions showed different levels of re-use of the referents: from a full assimilation leading to a free and most creative reuse, through a creative adaptation, and till a clumsier patchwork. According to a questionnaire given to the students at the end of the term, the manipulation and the reutilization of chunks of know-how were evaluated as enriching and stimulating for the creativity. Nevertheless, some difficulties were noticed: parametric modeling needs more concentration on the elucidation of design intentions, and requires a different way of thinking. Even more, in this approach the computer is allowed to 'participate' in the form-finding process, thus taking a part of the creative role of a designer. These cognitive and ethical issues are not easy to accept or overcome.

6. HOW TO USE THE LIBRARY

The prototype of LibreArchi is accessible via a website: http://www.arclab.umontreal.ca/LibreArchi/. The access for download is controlled through a password only in order to keep track of the interest in the community and the use of the library. The access for upload is given after identification to qualified members of the community. The authors and the community of students in our School of Architecture would be happy to share this asset with a larger academic community. At the same time, we will be grateful to be able to enrich the creativity-stimulating and the learning-assisting potential of LibreArchi through your contribution.

7. CONCLUSIONS AND PERSPECTIVES FOR THE DEVELOPMENT

From a didactic point of view, the library of referents has a double objective: teach to students a new way to think of and explore an object in design (based on processes and not on a result only); and provide architectural know-how concerning the design process and the building performance. As a Conceptual Digital Space, the library offers a first collection of referents of a new type: interactive models and simulations. It proposes a method to organize these referents in order to be able to reuse them later on. The library is 'open' and allows the addition of an unlimited number of new referents or chunks of knowledge, linked or not linked to the already existing ones. The type of content and the representation formats are not limited either (as long as they remain digital). The perspectives for the development of LibreArchi and the digital approaches to architectural design include the addition of new referents and the optimization of the existing models. An ontology network is considered to be used as an 'intelligent' organization of the referents, instead of the presently available web-page interface.

Bibliography

Ábalos, I. (2007). *Sustainable?* Colloque au CCA Study Center, Montréal, 2007.

Akin, Ö. (2002). Case Based Instruction Strategies in Architecture. *Design Studies*, *23*(4), 407-431.

Attar, R., Aish, R., Stam, J., Brinsmead, D., Tessier, A., Glueck, M. and Khan, A. (2009) Physics-Based Generative Design. *In:* T. Tidafi & T. Dorta, eds. Proc. of *Joining Languages, Cultures & Visions: CAADFutures'09*, 231-244.

Fischer, T. (2007) Enablement or Restriction? On supporting others in making (sense of things). *Proc. CAAD Futures 2007*, Springer, 585-598.

Iordanova, I. (2009) *Assistance de l'enseignement de la conception architecturale par la modélisation de savoir-faire des référents*, Ph.D Thesis, Université de Montréal, 390. http://hdl.handle.net/1866/2798

Iordanova, I., & Tidafi, T. (2007). Referents modeling for the architectural design studio: Cognitive bases. *In*: Paoli, G.D., Zreik, K., et R. Baheshti, eds*: Proc. of EuropIA-11*, 171-184.

Kolarevic, B. (Ed.) (2003). *Architecture in the digital age*, New York: Spon Press

Leclercq, P., & Heylighen, A. (2002, July, 2002). 5,8 analogies per hour. *In:* Proc.of: *AID'02 Artificial Intelligence in Design* Cambridge, UK. 285-303.

Oxman, R. (2006). Theory and design in the first digital age. *In: Design Studies*, *17*(3), 229-265.

Seebohm, T. (2007). Paradigm Shift. *In* : De Paoli, G., Zreik, K., et R. Baheshti, eds: Proc. of *EuropIA-11*, Montreal, 3-17.

Tidafi, T. (2007). New Digital Cultures Facing Complexity. *In:* De Paoli, G., Zreik, K., et R. Baheshti, eds: Proc. of *EuropIA-11*, Montreal, 131-145.

Turrin, M., A. Kilian, R. Stouffs, S. Sariyildiz (2009) Digital Design Exploration Of Structural Morphologies Integrating Adaptable Modules : A design process based on parametric modeling, *In*: T. Tidafi & T. Dorta, eds.: Proc. of *Joining Languages, Cultures and Visions: CAADFutures 2009*, 231-244.

West, M. (2001) Fabric Formed Concrete Structures. *In:* Proc. of *First International Conference on Concrete and Development*, Teheran

BIM-based Building Performance Monitor

Ramtin Attar, Ebenezer Hailemariam, Michael Glueck, Alex Tessier, James McCrae, Azam Khan
Autodesk Research
210 King Street East
Toronto, ON, Canada M5A 1J7
{firstname.lastname}@autodesk.com

Keywords: Building Performance Monitor, Visualization, Sensor-Network, Sustainability, Building Information Modeling.

Abstract

This video presents a set of visualization techniques for displaying real-time building behavior and usage, for the purposes of energy minimization, in the context of a semantically-rich building information model. Ambient occlusion shading is used to visualize the 3D space of the building, over which other visualizations may be layered. Specifically, occupancy, power usage, heat sources, air flow, and temperature are presented, as well as aggregated data sets such as “activity”.

1. OVERVIEW

Real-time building performance monitors, that visualize information collected from sensors distributed throughout a building, typically show the raw data values as simple text labels on 2D floor plans. This video shows how the same data can be better visualized in 3D (see Figure 1) in the context of the specific zone configuration of the building under consideration (see Figure 2). For example, while the temperature in a room is a valuable piece of information, seeing the temperature visualized as surface shading in the room together with seeing physical features that may contribute to the current temperature value (the air conditioning vents, air flow pattern, the size of the windows, placement of radiators, etc.) will provide much more contextual information to the building operator looking at the data (see Figure 3).

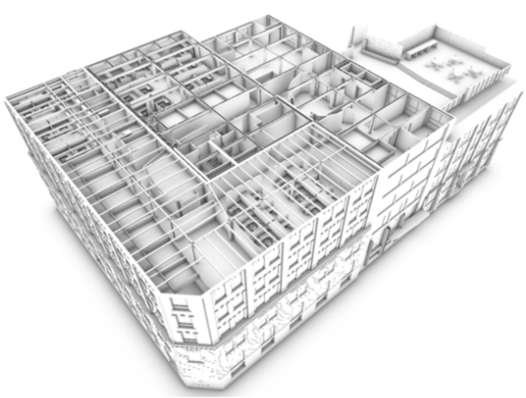


Figure 1. Real-time ambient occlusion base rendering of building surfaces.

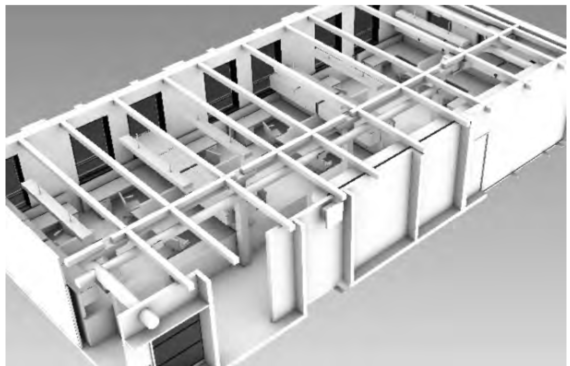


Figure 2. Specific zone on 5th floor of building, showing air handling units, windows, cubicle layouts, etc. obtained from the building information model.

Figure 3. Real-time visualization of occupancy (chair highlighting), heat source (computer glowing and red arrows), power usage (text label and meter), and aggregated “activity” status (peach shading of background).

References

Hailemariam, E., Glueck, M., Attar, R., Tessier, A., McCrae, J., Khan, A. 2010. Toward a Unified Representation System of Performance-Related Data. eSim 2010 Proceedings. (In press).

Project Metropolis: Digital Cities

Richard D. Howard
Autodesk Inc.
111 McInnis Parkway
San Rafael, CA 94903, USA
richard.howard@autodesk.com

Keywords: Urban Design, Visualization, Sustainability, Building Information Modeling, Digital City.

Abstract

This video presents a conceptual framework for a Digital City simulation and analysis software application that combines data from geographic information, building information modeling, traffic simulation, and sunlight simulation datasets. Informal studies with potential users were performed.

1. OVERVIEW

Project Metropolis, the codename for Autodesk's Digital City initiative, is an application and development framework concept for aggregating design models and GIS data within an immersive 3D geospatial environment for the purposes of visualizing, analyzing, and simulating trends, phenomena, state, and events in ways that are not possible within conventional 2D geographic information systems (GIS) or existing 3D application software. Metropolis is intended to be used in conjunction with existing GIS data produced by commercially available products such as those from ESRI and the Autodesk geospatial applications, and then augmented with highly detailed models from design applications like AutoCAD Civil3D for terrain and roads, Autodesk Revit for building information models (BIM), and Autodesk 3D Studio and Maya for other relevant models such as street furniture, bridges, etc. The software proposes a plug-in architecture enabling extensible and advanced 3D visualization, data analysis, and time-based simulation. Our goal was to illustrate a number of potential scenarios of use via a highly interactive and graphical technology mockup. Scenarios include interactive analysis of underground infrastructure, sun/shadow simulation for specific time-of-day and day-of-year (see Figure 1), the simulation of buildings and materials and urban sound pollution, simulation of energy usage over short and long intervals, carbon footprint and greenhouse gas emission analysis, and transportation simulation for differing times and conditions (see Figure 2). The Metropolis mockup was used to both assess market potential via live focus group demonstrations in twelve cities worldwide and also to validate specific functional innovations via structured one-on-one studies with potential users.

Figure 1. Solar study interface design from the Project Metropolis video.

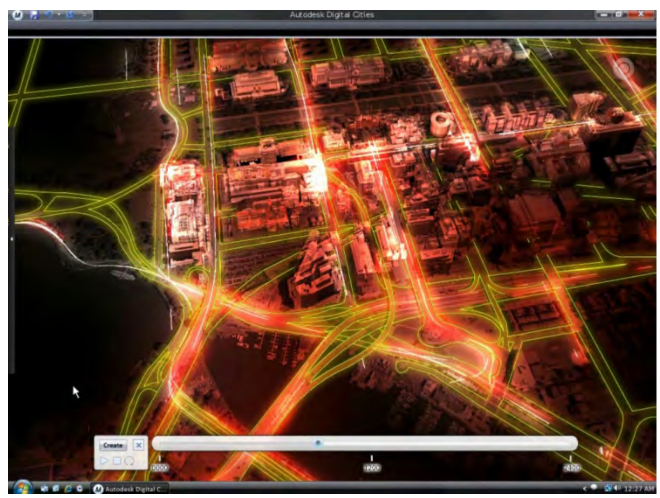


Figure 2. Traffic flow visualization from a simulated traffic dataset.

Acknowledgements

Special thanks to Noble Studios, Dave Arsenault, Liam Speden, Robert Bray, and Chris Andrews.

Organizers

General Chair

Azam Khan
Autodesk Research

Program Committee

Robert Aish
Autodesk Inc

Ramtin Attar
Autodesk Research

Ellen Yi-Luen Do
College of Architecture & College of Computing, Georgia Tech

Mark D. Gross
CoDe Lab, Carnegie Mellon University

Kasper Hornbæk
Department of Computer Science, University of Copenhagen

Michael Jemtrud
School of Architecture, McGill University

Judit Kimpian
Head of Sustainability and Advanced Modelling, AEDAS

Branko Kolarevic
Faculty of Environmental Design, University of Calgary

Barry Leonard
Air Staff, National Defence Headquarters

Greg Madey
Research Professor, University of Notre Dame

Rivka Oxman
Vice Dean, Faculty of Architecture and Town Planning, Technion ITT

Gabriel Wainer
ARS Lab, Carleton University

Sponsors

Sponsored by

Autodesk®

In co-operation with

acm

SCS

SpringSim

Cover Image Credits

145 **Virtual Driving and Eco-Simulation**

CHRISTOPHER J. GRASSO, MICHAEL J. MCDEARMON and YOSHIHIRO KOBAYASHI
Forum8AZ and Arizona State University

169 **Project Metropolis: Digital Cities**

RICHARD D. HOWARD
Autodesk Inc

163 **LibreArchi: Library of Interactive Architectural Models Containing Exploratory and Didactic Simulations**

IVANKA IORDANOVA and TEMY TIDAFI
Université de Montréal, Canada

97 **Intuitive Structures: Applications of Dynamic Simulations in Early Design Stage**

ANDRZEJ ZARZYCKI
New Jersey Institute of Technology

05 **Multi-Objective Optimization In Architectural Design**

IAN KEOUGH and DAVID BENJAMIN
Buro Happold Consulting Engineers & Columbia University Graduate School of Architecture, Planning, and Preservation

05 **Multi-Objective Optimization In Architectural Design**

IAN KEOUGH and DAVID BENJAMIN
Buro Happold Consulting Engineers & Columbia University Graduate School of Architecture, Planning, and Preservation

31 **Explorations of Agent-Based Simulation for Architectural Design**

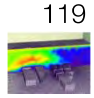
NICK PUCKETT
University of Kentucky

61 **Space Perception and Luminance Contrast: Investigation and Design Applications through Perceptually Based Computer Simulations**

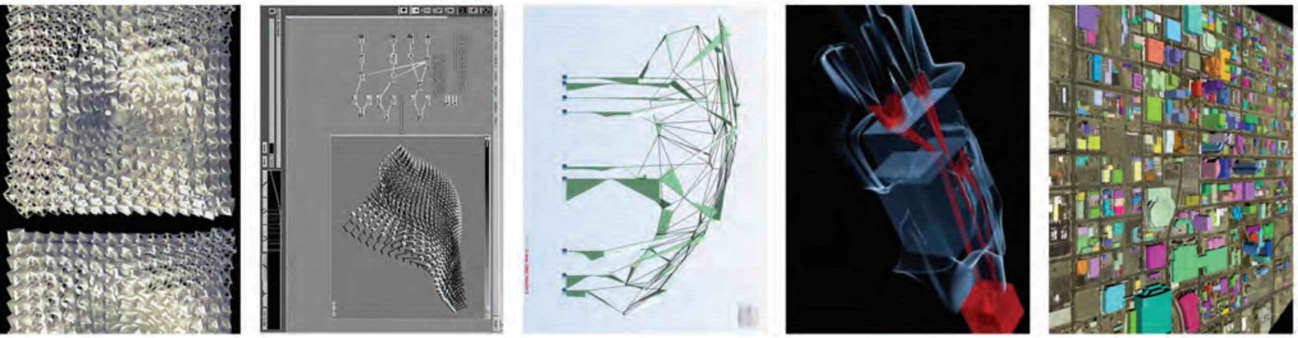
NAN-CHING TAI and MEHLIKA INANICI
University of Washington, Department of Architecture

155 **Associative modelling of Multiscale Fibre Composite Adaptive Systems**

MARIA MINGALLON, SAKTHIVEL RAMASWAMY and KONSTANTINOS KARATZAS
Architectural Association School of Architecture

119 **Augmented Reality Framework supporting Conceptual Urban Planning, enhancing the Awareness for Environmental Impact**

HOLGER GRAF, PEDRO SANTOS and ANDRÉ STORK
Fraunhofer Institute and TU-Darmstadt

Index of Authors



Symposium on Simulation for Architecture and urban Design 2010