The Modellers' Apprentice - Gesture-Based 3D Design in Immersive Environments

François Rioux, Frank Rudzicz, Michael Wozniewski
Centre for Intelligent Machines,
McGill University
3480 University Street
Montreal, Quebec Canada
{frioux, frudzi, mikewoz}@cim.mcgill.ca

ABSTRACT
We describe a system that takes advantage of two-handed interaction in an immersive environment. The system aims to improve visualization of 3D (CAD-like) design applications. The toolglass metaphor is employed as an interface widget allowing users to control the system with gesture. The gesture set is designed with the rules of two-handed interaction in mind, offering guidelines for the most natural usage. The development process relies heavily on user testing and feedback, evolving from a paper prototype to a fully immersive gesture-driven environment. This results in an iterative process, where users play the most important role in the improvement of the system. We document the evolution of this project from theoretical conception to working beta system.

Author Keywords
Immersive environment, transparent toolglass, bimanual interaction, 3D Modelling

ACM Classification Keywords
H.5.2 User Interfaces - I.3.6 Methodology and Techniques (Interaction techniques)

INTRODUCTION
The Modellers' Apprentice is a fully immersive 3D computer-augmented environment aimed to improve visualization for 3D modelling applications. Users can simultaneously navigate through a world of their own creation, while placing and modifying 3D models in real time. The interface is designed with the properties of two-handed interaction in mind, and makes use of the transparent toolglass metaphor.

Many potential target audiences could benefit from such a system including users of computer assisted design (CAD) technologies, 3D game designers, and any architect, designer or engineer with a need for 3D layout tools. The focus of the initial prototype described in this paper is on the tasks of the interior designer, creating an application to facilitate the 3D layout of interior spaces. Users are able to place models of furniture, decorations, and appliances in a room and then apply various actions such as translation, rotation, texturing, etc.. Various configurations can be saved and loaded, and the designer can then bring in his/her client to experience the space before construction begins.

The system is built on the infrastructure of the Shared Reality Environment (SRE) at McGill University [5]. This is an un-tethered environment whose only source of input is video from a series of well-positioned cameras, and hence the primary form of interaction employs the tracking of physical body gestures. The SRE displays output by three back-projected screens which surround the user and immerses their field of vision as shown in Figure 1.

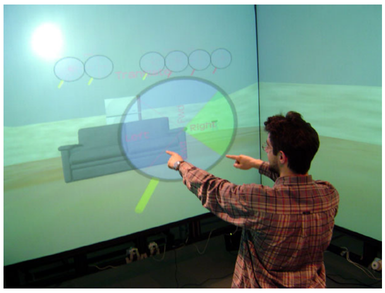


Figure 1. *The Modellers Apprentice* running in The Shared Reality Environment.

RELATED WORK
Several new interaction techniques have recently emerged for 3D visualization and modelling with two hands. Grossman et al. [6] consider a slightly lower-level approach to two handed interaction on large displays. They create 2D planes using a technique called 'tape drawing' which 3D modelling is then base upon. Hinckley et al. [8] and Stoakley et al. [12] use interface props that a user holds in his non-preferred hand and can set a frame of reference or camera position to

view a 3D model/world. In some cases, users can manipulate the world using their preferred hand. Balakrishnan et al. [1] similarly explore moving a camera around a virtual world with the non-preferred hand while manipulating 3D objects with the preferred hand, yet only use a keyboard and mouse. The project that has been developed makes use of Toolglass interface widgets which were proposed by Bier et al. [3].

SYSTEM OVERVIEW

The Toolglass Metaphor

An important feature of *The Modellers' Apprentice* is the use of toolglass widgets, as originally proposed by Bier et al. [3]. The toolglass interface is composed of a semi-translucent circular menu that is angularly partitioned into pie-like wedges. Each wedge represents a unique action within the context of its toolglass. An action is applied by "clicking through" a wedge, and the action is applied (in general) to objects directly behind the center of the toolglass. Figure 2 shows a concept diagram of a typical toolglass that might be used in the system.

Figure 2. A toolglass for placing models

The toolglass has been designed with the additional affordance of a handle, that should hopefully stimulate the user to 'grab' it. The fact that the handle is on the left hand side is purposeful, and should likewise urge users to grab the toolglass with their left hand. The reason for this has to do with the properties of two-handed interaction (described below).

Using gestures described later in the gestural interface section, a user will select a toolglass from the rack of available toolglasses and position it over a target area. With the other hand, a wedge is selected and the associated action for that wedge is invoked. The target object or position for the action is found by a projection of the user's view through the 'crosshair-like' center of the toolglass.

Bimanual interaction

There is an obvious disparity in the utility of our left and right hands. This is often described as *handedness*, where one ***preferred hand*** is favored over the other for certain tasks. Yves Guiard [7] defines some rules governing this phenomenon:

- *Preferred-to-non-preferred Reference:* the preferred hand finds its spatial references in the results of non-preferred hand motion.
- *Asymmetric scales of motion:* the preferred hand has higher temporal and spatial scales of motion. Non-preferred hand has more infrequent, coarse movements.
- *Non-preferred hand precedence:* the non-preferred hand moves first to set a frame of reference, then the preferred hand applies actions relative to that reference.

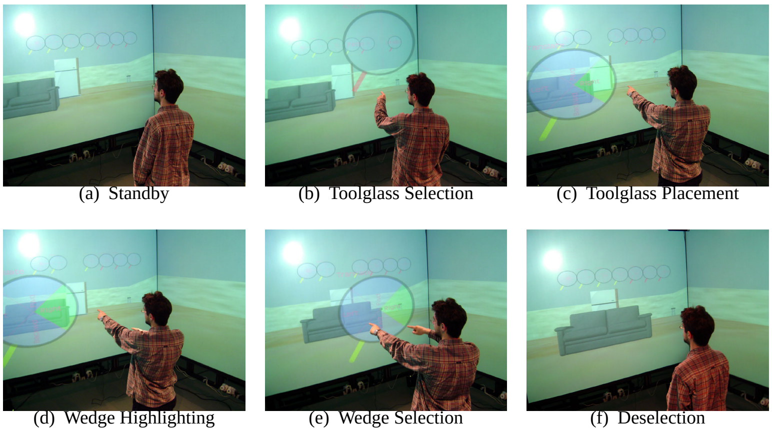
Several studies [4, 8, 10, 9, 1] have illustrated the benefits of well designed bimanual tasks. Not only does bimanuality aid in speed and efficiency, but the structure of the task can be more complex. More information is inherently available by means of the kinaesthetic feedback provided by the muscles in our arms and hands; a user of a bimanual interface subconsciously knows the position of his hands in space and the distance between them. Hence less artificial feedback is needed and in the end, the cognitive load required by the user is lessened.

These rules and benefits can offer important constraints and guidelines for the design of new interaction techniques with two hands. In particular, Guiard's principles are considered to help design the gesture set for interacting with toolglasses. The user is biased to grab toolglasses with his/her non-preferred hand, thus setting the frame of reference, while the preferred hand is used to apply actions. The preferred hand always moves relative to that reference, and if the user makes several small adjustments to models, the frequency of the preferred hand will be greater than the non-preferred hand. These guidelines should result in natural and efficient interaction for most tasks available in the environment.

Gesture Set

The use of toolglasses allows a fairly simple gesture set since all actions on the system can be performed through the use of these menus. To begin, it is assumed that a single user stands in the center of the SRE, having both hands at his/her side. The gestures available to the user are as follows:

1. *Pointing*: In general, pointing is invoked by holding an arm partially (but not fully) extended. It has the effect of moving cursors about the screen (in an arc about the user) without altering the state of the environment. Pointing however results in different behavior depending on the hand used:
   (a) *Non-preferred hand*: When no toolglass is held the user points to a toolglass rack to *indicate* which one he or she would like to grasp. Once a toolglass is selected, it will follow the motion of the pointing hand and hence allows the user to position it over a target. See figure 3(b).
   (b) *Preferred hand*: When a toolglass is held by the non-preferred hand, pointing *indicates* which wedge from the toolglass he or she would like to select. See figure 3(d).



(a) Standby (b) Toolglass Selection (c) Toolglass Placement

(d) Wedge Highlighting (e) Wedge Selection (f) Deselection

Figure 3. Gestural interaction with toolglass metaphor

2. *Selection*: The selection gesture occurs when either hand fully extends away from the user's body. It has the effect of selecting the toolglass or wedge that is currently *indicated* by the pointing gesture:

    (a) *Non-preferred hand*: The toolglass that the hand is currently pointing at will become 'grasped', and attach itself to the cursor corresponding to the hand. See figure 3(c).

    (b) *Preferred hand*: The action corresponding to the wedge that the hand is currently pointing at will be invoked. See Figure 3(e).

3. *Deselection*: Deselection occurs automatically when users returns their hands to their sides, resulting in any selected toolglasses returning to the rack. See Figure 3(f).

4. *Navigation*: Navigation occurs when users offsets their center of mass away from the center of the SRE, and is used to move the user's view around the environment in the direction of that offset. For instance, users moving forward results in sustained forward movement in the environment, until they return to their original physical position. Note that this action may also be achieved via the toolglass interface.

DESIGN STRATEGY

In order to maintain consistent vision and direction throughout the iterative development of *The Modellers' Apprentice*, the design team adhered closely to important principles of HCI and design. Decisions were always user-centered and endeavored to create a natural interaction paradigm. These guidelines allowed for a consistent and unified development process, which is described in the proceeding sections.

HCI Principles

The affordances and expressiveness provided by the chosen interaction techniques were rich, and required that careful attention be paid to their effects on other principles of usability. This attention followed an overall design strategy throughout the project, defined by several usability heuristics [11] that materialized over time, as follows:

- *Aesthetic and Minimalist Design*: There is no irrelevant information in the interface, and any forcing functions are made 'invisible' to the user.
- *Visibility of System Status*: The status of the environment and objects within that environment are immediately visible to the user, and all actions that can be performed are distinguished from those that are not.
- *Recognition Rather than Recall*: Cognitive load on users is kept to a minimum using affordances provided by the toolglass interface and bimanual interaction techniques.
- *Flexibility and Efficiency of Use*: Experienced users can modify the system configuration to remove, adjust, or add interface functionality and model sets, according to their tastes to alter their task domain or accelerate their experience.

Design Techniques

Three HCI techniques provided the principal guidelines involved in the design and implementation of *The Modellers' Apprentice*, namely:

1. <u>Early Prototyping and Testing</u>: Early focus on users concentrated the direction of the project. Preliminary paper prototypes, interviews with potential users, and a minimally functional prototype allowed for the discovery of new affordances, and important constraints for the system.

2. <u>Iterative design</u>: The project underwent several iterations of design, implementation, testing, and re-design according to formative user feedback. At each stage of the process, subsequent versions of the system quantitatively im-

proved upon their predecessors in terms of decreasing error reports and interface complaints from the users.

3. User-Centered design: The feedback was obtained from three groups: an initial pool of users, a novice evaluation team and an expert evaluation team. At each stage, their feedback was carefully analyzed and transformed into targets that were resolved in the subsequent release.

USER FEEDBACK, ANALYSIS, AND INCORPORATION

The focus of design was placed on the user from the outset, and at each step of the iterative design process, their concerns were compiled through various means, analyzed, and acted upon. User feedback was drawn from three groups of testers, the first being untrained in principles HCI, the second being proficient, and the third expert. The type of feedback, and the manner in which it was elicited, was qualitatively different for each group, as described below.

Initial User Feedback

Before any code was written, a pool of almost a dozen people who were unfamiliar with the principles of HCI (some were barely computer literate), were interviewed in order to get a non-technical and unbiased sense of the most basic interface issues the initial plan would encounter. Interview techniques ranged from informal qualitative discussion to guided paper protototyping, and some initial concerns included the following:

- *Toolglass rack* Semantically dividing the tasks into multiple toolglasses, and having them arranged linearly reduced and separated cognitive load, and was somewhat reminiscent 'drop-down' menus.
- *Handle attachment* Early designs for the toolglass excluded the handle attachment, but many of those interviewed suggested that they would not initially know what to do with a mere circular toolglass without some other visual clue. Adding the handle also benefits us technically, as it aids to naturally separate the user's hands in the SRE.
- *Wedge highlighting* To highlight the selected toolglass wedge before invoking its action would make the users more comfortable and aware of the state.
- *Labels* Textually labelling toolglasses by category, and wedges by function, would also provide an extra degree of user assistance.
- *Redundancy* Although the initial group seemed enthusiastic on the whole with regards to the toolglass metaphor, several felt that gestures besides pointing and selecting would enrich the experience.

Prototype

Coded in OpenGL, the prototype system was actually relatively functional for a prototype. It allowed the modification of OpenGL primitives (sphere, cube, torus, and teapot), in a 3D environment, with 3D toolglasses, as shown in 6(a). Functions available to the user included object placement, movement, rotation, scaling, and colouring. Other toolglasses also allowed the user to exit the system, 'reset' the environment, and other so-called 'session' actions. It ran on a PC architecture (Windows and Linux), using a keyboard-and-mouse interface. The contrast between the features of the prototype interface and of the the target interface were stark, and the test users were asked to concentrate on other elements of the interface in their evaluation.

Evaluation of Prototype

An evaluation team consisting of other system developers with a knowledge of HCI principles were assigned with a set of laboratory experiments and asked to provide formative feedback. The experiments were objective measures of the prototype interface and satisfied each of the evaluation criteria mentioned above. Formative feedback included the following suggestions:

- *Visual Feedback*: Despite the affordances provided by the toolglass 'crosshairs', additional feedback indicating which object was selected was desired.
- *Error Prevention/Recovery*: The ability to undo undesired or mistaken actions, and confirmations for potentially 'dangerous' operations were urgently desired features.
- *Toolglass Classification*: A way to differentiate between toolglasses that operated on the environment, and those whose functionality was 'global' to the session, was desired.
- *Global Session Parameters*: Evaluators requested the ability to modify global parameters that would allow the control of general, high-level settings.

As it happened, most of the evaluation team's feedback echoed directions the development team had already decided to take, but their suggestions, and results from the experiments, indicated that the design was on track.

Alpha System

Once the feedback from the first iteration of development was compiled and analyzed, it became a requirements list for the alpha release of the project. Although the most important changes between the prototype and the alpha system were reactions to user feedback, others were self-initiated according to a long-term design arc, as described below.

Reactive Improvements

Simple chromatic object highlighting was implemented to satisfy the requirement of visual feedback (although it was not the ideal solution, as rectified later), and a new 'Undo/Redo' toolglass was added that held a stack of invertible actions for error recovery. Furthermore, XML configuration files were created that allowed modification of high-level properties such as the nature of toolglasses, available models, or the interaction technique itself. Classifying toolglasses according to semantic property was accomplished according to colour differentiation on the toolglass handle, as shown in Figure 4.

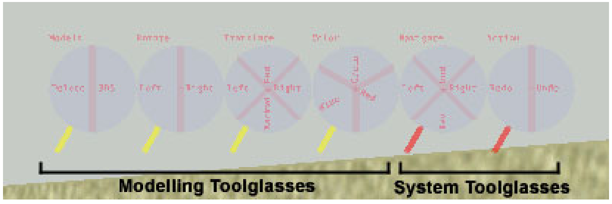


Figure 4. Toolglass Rack with task differentiation by colour.

Self-Initiated Improvements

Improvements were also made in the alpha release according to the developer's own design arc. These included replacing the primitive model shapes with more complex 3D graphic models (3DS) (shown in Figure 5), adding collision detection with the walls of the environment to better match the real world, and adding functionality to 'save' and 'load' configurations of the world to disk. Other general improvements to the graphic aesthetic and responsiveness of the interface were made during this iteration.

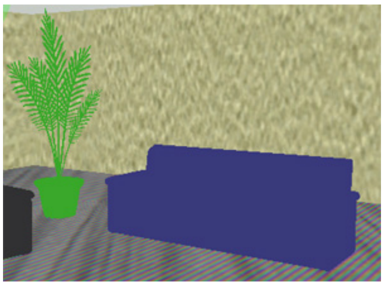


Figure 5. Initial and partial support for 3DS models.

Evaluation of Alpha

An evaluation team consisting of relative experts in HCI (Master's students whose theses were in this field) were asked to evaluate the alpha system according to their own standards and guidelines which were undisclosed to us. Their feedback came in the form of action items that they suggested should be rectified before the beta release, and focussed primarily on the accompanying documentation rather than the software. However, they did point out technical problems of lighting conditions in the room, and suggested a more striking visual cue was necessary for the toolglasses on the rack for them to 'stand out'. Unfortunately, these evaluators cited a difficulty recognizing *which* object in the environment was selected by a toolglass at any time, indicating that the chromatic highlighting implemented after the prototype was *not* the optimal solution to this problem.

Beta System

Given critiques from the experts in HCI, changes have been made in order to solve the problems they encountered. The software has also been integrated into an immersive environment available in the *Shared Reality Environment* [5].

Reactive Improvements

Like in the previous version of the system, the problem was the feedback provided when a toolglass pointed out at an object was not obvious enough. Therefore, a new mechanism has been implemented which draws a bright green bounding box around the object the toolglass is pointing at. Some other changes have been made in order to avoid a user to making an irreversible action. The toolglass tree has been configured so that critical actions like "Quit", "Reset" and "Load Session" will pop a confirmation toolglass to confirm the action the user is trying to perform. Several improvements have also been made in order to make the toolglasses more visible on their rack by putting a black rim around them and by separating them along the screen. This update can be seen on figure 7. A lot of effort has been put on the appearance of the scene as well, so it looks as realistic as possible. To extend features available to the user, the texturing of 3DS models and submodels have been implemented as can be seen in figure 7.

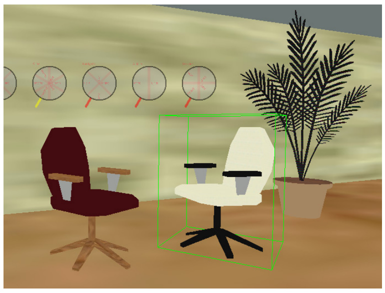
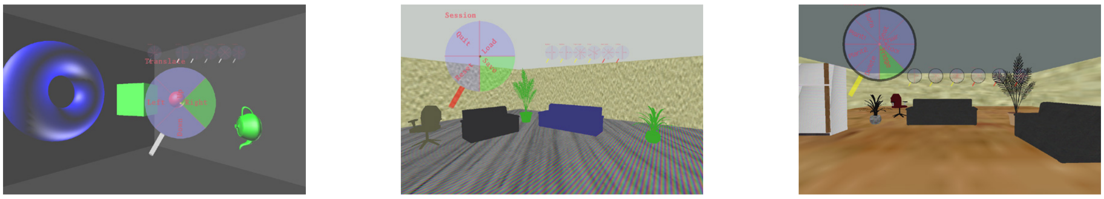


Figure 7. Improved object highlighting, toolglass visibility, and subobject texturing in the Beta system.

Integration in the SRE

The prototype and alpha versions of the system were implemented using mouse and keyboard as input gestures and a computer screen as a "3D" output. However, for the beta version of *The Modellers' Apprentice*, the system has been implemented in an immersive environment composed of three screens of 8 by 8 feet with rear projection display. Therefore, the user is surrounded by the world he created in order to have an immersive experience of the world he is designing. In the beta system, gestures were used as input to the system as described in the gestures list section. The hands were tracked with a simple segmentation and tracking heuristic in order to provide their position in three dimensions. For gesture recognition, a simple finite state machine was implemented. There was no need for a more complex method of recognition since the gesture set is simple enough to represent each gesture command with a finite number of states.



(a) Prototype (b) Alpha system (c) Beta system

Figure 6. Evolution of design: a) Prototype, b) Alpha system c) Beta system

CONCLUDING REMARKS & FUTURE WORK

Overall, the design decisions have produced a system that feels natural to use. The iterative design process and user-centered design has solved many confusions along the way and revealed several weaknesses that have since been repaired. Involving users early in the process also saved time in the long run.

The gesture set was found to be usable, yet some users commented that it would be nice to bypass the toolglass interface for quick and simple gestures like moving and rotating a model. Currently, the limitation for implementing such gestures is the tracking and recognition system, which only provides rough estimates of hand position. One particular area of future work is to implement more robust tracking such that more advanced gestures will be possible. Another area of improvement that will increase efficiency is to replace text labels on toolglasses with pictograms. This will allow users to gain a quicker overview of the toolglass functions.

To validate the benefits of the system, some formal usability tests should be performed such as exploring whether two hands versus one hand improves efficiency and lightens the cognitive load of certain tasks. Also, tests should be created to see how people compare the toolglass metaphor compared to the more traditional metaphor of menus and a pointer.

With more robust tracking and continued iterative design, the system should provide a valuable ground for the exploration of new interaction paradigms.

ACKNOWLEDGEMENTS

The authors would like to thank Yves Boussemart, Nadia Hilario, Jerome Pasquero, Daniel Sud, and Jeremy Cooperstock for their assistance and advice.

REFERENCES

1. Balakrishnan, R., Kurtenbach, G. Exploring Bimanual Camera Control and Object Manipulation in 3D Graphics Interfaces. *Proceedings of ACM CHI 1999 Conference on Human Factors in Computing Systems*. 1999. p.56-62.
2. Bier, E.A., Stone, M.C., Fishkin, K., Buxton, W., and Baudel, T. A Taxonomy of See-Through Tools. In *Proceedings of CHI '94*, p.358-364.
3. Bier, E.A., Stone, M.C., Pier, K., Buxton, W., and DeRose, T. Toolglass and Magic Lenses: The See-Through Interface. In *Proceedings of SIGGRAPH '93*. **27**: p.73-80.
4. Buxton, W., Myers, B.A. A study in two-handed input. *Proceedings of the SIGCHI conference on Human factors in computing systems*. 1986. p. 321-326.
5. Cooperstock, J. Shared Reality Environment. http://www.cim.mcgill.ca/ jer/research/sre/, March 2004.
6. Grossman, T., Balakrishnan, R., Kurtenbach, G., Fitzmaurice, G., Khan, A., Buxton, B. Interaction Techniques for 3D Modeling on Large Displays *Symposium on Interactive 3D Graphics*. 2001. p.17-23.
7. Guiard, Y. Asymmetric Division of Labor in Human Skilled Bimanual Action: The Kinematic Chain as a Model *Journal of Motor Behavior*, 1987, **19**(4): p.486-517.
8. Hinckley, K., Pausch, R., Proffitt, D., Kassell, N.F. Two-Handed Virtual Manipulation *ACM Transactions on Computer-Human Interaction (TOCHI)*. 1998. **5**(3). p.260-302.
9. Kabbash, P., MacKenzie, I.S, Buxton, W. Human Performance Using Computer Input Devices in the Preferred and Non-Preferred Hands. *Proceedings of SIGCHI conference on Human factors in computing systems*. 1993. p. 474-481.
10. Leganchuk, A., Zhai, S., Buxton, W. Manual and Cognitive Benefits of Two-handed Input: An Experimental Study *ACM Transactions on Computer-Human Interaction*, 1998. **5**(4). p.326-359.
11. Nielsen, J. Heuristics for User Interface Design. http://www.useit.com/papers/heuristic/heuristic_list.html, April 2004.
12. R. Stoakley, M.J. Conway, and R. Pausch. Virtual reality on a WIM: Interactive worlds in miniature. In *CHI '95 Conference Proceedings*. 1995. pages 265–272, 1995.