I³ - EYE-CUBE

Interactive intuitive mixed-reality interface for Virtual Architecture

STEPHEN K. WITTKOPF, SZE LEE TEO
National University of Singapore
Department of Architecture and Fellow of Asia Research Institute
akiskw@nus.edu.sg

AND

ZHOU ZHIYING
National University of Singapore
Department of Electrical & Computer Engineering
g0301340@nus.edu.sg

Abstract. This paper introduces a new tangible interface for navigating through immersive virtual Architecture. It replaces the common mouse or glove with a set of tangible cubes. It includes physical architectural floor plans as contextual haptic constraints for the cubes to ensure better object manipulation compared to free space. The position and orientation of the cubes relative to the floor plan is tracked by web cameras and a newly developed program translates this into the 6-dof of the virtual camera generating a 3D view for the immersive projection of virtual architecture. This easy to use tangible interface mixes common 2D dimensional (real) with 3D immersive representation (virtual) of Architecture to overcome the problem of ‘Getting lost in Cyberspace’.

1. Introduction – Immersive Presentation of Architecture

In this paper we focus on selected aspects of architecture namely on the organization of space and representation of scale. Consequently, computational visualizations of architecture need to enable proper visual perception of space and scale. Hence visualizations of architecture on small computer screens and interaction devices such as mouse or keyboard are
deemed to be not sufficient. Fortunately the advances of technology and development of Virtual Reality have brought about large screen visualizations with stereoscopic immersive projections that ‘wrap’ the 3d space around a user eventually making him believe that he is inside the virtual world rather than looking at it from outside. This is indeed a very positive development towards better visual perception of architecture, but we still feel that the navigation is too difficult to appropriately explore architectural space.

2. Problem Statement – Lost in Cyberspace

Common navigation devices still require the user to wear a position and orientation tracker systems, whereby the movements of the viewer in the real space are synchronized with the virtual architecture. Now imagine a viewer has to explore a virtual corridor, how would he interact with the space? Intuitively one would start moving in real space, and expect the system to translate this into walking along the virtual corridor, well, until the user hits the projection screen. A regular mishap when layperson walks through virtual space for the first time.

Another persistent problem is that users tend to lose orientation exploring virtual worlds. They don’t know if they are still heading in the same direction, is the room where they came from towards the left or light, etc. This sense of orientation is very important in the exploration of architectural space. A lack of orientation is crucial, quite characteristic and is commonly results in what is referred to as ‘Getting lost in Cyberspace’.

3. Project brief – A new Tangible and Mixed Reality Interface

The objective of the project was to overcome this problem by approaching it from two directions. The first is more from the human-computer-interface point of view and concerns the user’s ability to navigate through space intuitively. An easy control of the 6 degrees-of-freedom (x,y,z, yaw, pitch, roll) is not given by a common mouse (2-dof) and gloves are too technical to use, although they allow for 6-dof. The second problem comes from an architectural point of view and expresses the concern that the 2D representations as printed floor plans, section, elevations are still common when dealing with architecture and should thus not be excluded. Hence the objective was to propose a navigation interface that links both the 3d-immersive and 2d-drawing representation.

A first brief mapped the key features of such a system which foresees a table which holds 2d drawings and a moveable tangible object that
represents the camera. The object shall become a non-wired easy to grasp interface which translates all 6-dof to the virtual camera. And the object resting on the floor plan, should establish a 3rd person view that tells the positions of user or rather virtual camera. The navigation interface should furthermore be independent of the displayed architecture, meaning it should work with other floor plans of even different scale and architecture. The procedure to synchronize the coordinate origin between the floor plans and 3d model and the extents in all three directions must be an easy procedure to be performed quickly by a standard user.

4. Mixed Reality Lab (Department of Engineering)

Current research in developing human-computer-interfaces tries to overcome these problems by developing interface devices that represent themselves less technical and are easy to handle and respond intuitively. We refer to Augmented or Mixed Reality when the three dimensional computer generated virtual space (or architecture in our case) augments the visual cues of the real work we are in, so that basically both worlds a) the Virtual and b) the Real can be seen at the same time (Milgram, Takemura, Utsumi, and Kishino, 1994). At the same time two major transitions happen to replace the traditional input and output devices. So called multimodal interfaces extend the range of possible user input by gesture, sound, speech, touch etc. (Schomaker, Nijstmans, and Camurri, 1995). The usual glove for interacting in virtual worlds is for instance such an interface which allows the user to communicate with the system by gestures expressed through finger positions or movements.

On the other hand usual output devices such as monitor screens are replaced by surrounding stereoscopic projections environments which make the user feel inside a space rather than looking at it through a window (Cruz-Neira, Sandin, and DeFanti, 1993). Wearable display systems such as Head-Mounted-Displays (HMD) are other developments in this area. One of the authors has developed several combinations of multimodal and mixed reality interfaces with one combination customized for this particular architectural usage (Zhou, 2004). Physical cubes are used as tangible user interface to interact with Augmented Reality (Ishii and Ullmer, 1997).

5. Digital Space Lab (Department of Architecture)

The interface is supposed to be integrated into the Digital Space Lab (DSL) of the Department of Architecture (Wittkopf, 2004). The DSL comprises of three systems. The commercial VR Software EON Professional is used to
import 3D-CAD models and render stereoscopic images of high resolution in real-time. The rendering is distributed over two high-performance graphic PC-workstations. The display system then blends both images together resulting in a total pixel resolution of 2304x1024. Four bright projectors beam the images from the back onto a translucent flat screen of 2m by 4.5m size. Each pair is projecting one view, meaning the left and right image overlap. This is then turned into a 3d image in the eye of the user by wearing simple polarized glasses. Figure 1 shows the PC-workstations and projection screen. For a presentations viewer would just stand or sit in front of the screen while a expert user sitting in front of the PC-workstations navigates them through the space. Alternatively the user can use a wireless mouse while directly looking at the projection screen.

This large, stereoscopic, bright image of high resolution allows users to view Architecture from within to judge on scale, space and visual connections as can be seen in Figure 2. The immersive visualization of architecture can be augmented with interactive features which eventually establish a laboratory for architectural design studies, a lab of particular importance for teaching and learning by experimenting.

*Figure 1. Working session inside the Digital Space Lab showing the PC-workstations on the right hand side and the back-projection screen behind the user*
The current navigation devices include a 2-dof and 6-dof mouse but experience has shown that the following movements are relative difficult:

- Going back/forth or left/right while looking around
- Panning vertically and horizontally along a façade
- Locking a certain angle (looking up) while panning or walking
- Jumping to one view without traveling

Two or three button mice only allow modifying two or three degrees of freedom (dof), which have to be identified upfront. The keyboard can help to activate the other dof’s but all 6-dof at the same time cannot be performed. The standard setting for instance would allow the user only to walk towards the view, which is quite in-natural since we look around wile walking. This forces a user to learn a new confusing way to navigate which is different from the natural and henceforth not intuitive.

Space mouse or 5-6-dof mice on the other hand are very touch sensitive and require additional push of buttons to switch between different dof and provide very little haptic feedback. Gloves require the user to learn a certain language of finger gestures before one can easily navigate through space. So in short, a natural navigation is characterized by movements in space (6-dof) but the supporting interfaces are not very intuitive.
5. The Interface

5.1 SYSTEM DESIGN

We name this interface system as the Eye-Cube, abbreviated as I^3 to represent an interactive intuitive interface. In most cases, the main function of first cube is to become a virtual eye of the user in virtual space, creating an immersive experience for him during the architecture visualization process. The second cube on the other hand serves as a multi-function interface device to allow the user to interact with the virtual environment in an intuitive manner, depending on how it was pre-programmed by the designer.

The core of this system lies in two cubes (7cm x 7cm x 7cm) with different patterns printed on each and every face of both cubes. We found the possible reasons for choosing cubes/blocks lie mainly in two aspects:

- As compared to a ball or other artifact in complex shapes, a cube/block has stable physical equilibriums (resting on one of its surfaces) which make it relatively easier to track/sense. In this sample system, we define the states of the cube by these physical equilibriums.
- Cubes when piled together form a compact and stable structure. This could reduce scatter on the interactive workspace.

In addition to the above mentioned, the cube is an intuitive and simple object that we are familiar with since childhood. This graspable object allows us to take advantage our keen spatial reasoning and leverages off our prehensile behaviors for physical object manipulations.

The cubes can be made of any solid and hard but relatively lightweight materials such as plastic or wood. In this case, we choose acrylic for it is easily available. The size of the cubes is chosen in such way so that it can be easily held in average adult users’ hands. There are neither wires attached nor circuitry embedded in the cubes; they are just plain solid cubes with patterns.

A table large enough to allow a floor-plan of A0 size (1189mm x 81 mm) to be placed on top of it is located 5m in front of a 4.5m by 2.5m rear projection screen. Two desktops computers are responsible for running the EON Professional visualization software to render the virtual environment the user desires to see and project it on the screen through 4 bright high-resolution projectors.

Figure 3 shows the components of the interface as part of the Digital Space Lab of the Department of Architecture. The foreground shows a table
with a floor plan on which the floor plan is mounted. Two web cameras are tracking the cubes position from above.

![Image of cubes with floor plan](image)

*Figure 3. Cubes resting on a physical drawing, vision tracked by two web cameras to generate the 3d view on the large projection screen.*

The table thus becomes the platform for the user of which he can interact with the virtual environment projected. As pointed out, this physical table surface will provide contextual haptic constraints to ensure better object manipulation compared to free space (Wang and Mac Kenzie, 2000). Based on the user’s manipulation of the cubes with respect to the floor-plan laid on the table, the virtual environment is directly influenced and affected. By this means, the tangible cubes have become a handle to interact with the physical and virtual world simultaneously. By referencing to the physical layouts, the designer will be well-instructed their location and orientations in the virtual world, hence he will not get lost in the virtual design space.

Figure 4 shows a user controlling the view with the cube showing the arrow on top, and saving the view through a rotation of the other cube. The cube size is actually depending on the resolution of the camera and its distance to the table.
Figure 4. User controlling the view with the cube showing the arrow on top and saving the view with a rotation of the other cube.

To track the movements and states of the cubes, two Unibrain IEEE1394 cameras are required to overlook the table from the top. The necessity of two cameras arises due to the limited field of view of the camera lenses to encompass the whole A0 size floor-plan. With two cameras, a volume space of A0 size by the height of 30cm can be covered with two cameras looking down from the height of 1.4 m, allowing users to move the cubes freely in such a space. Both cameras have a slight overlap vision of about 20% at the centre.

The video which has the cubes in it is captured from the camera are fed to the desktop computers via a 10m IEEE394 cable, where it will be processed by a program. This will be. Figure 5 shows how the developed program recognizes the cubes and translates this into meaningful data such as position and orientation to be channeled directly to the EON Professional visualization software in real-time.
In short, the whole system forms a close-loop feedback system, where the user’s physical input (cube manipulation) is affecting the output (virtual environment video projected on screen) and then back to the user as a feedback, allowing him to relate what he does with the cubes and what he sees on the screen. This system is shown in Figure 6. The complex communication and tedious computation process that lies underneath are completely invisible to the end-user. More importantly, the gap between the physical and virtual world become blurred and the user interacts with both worlds simultaneously and intuitively.
5.2. TECHNOLOGY

5.2.1. Tracking

The primary technology behind the I3 system lies in the field of vision tracking. As our task involves the tracking of 3D objects, we considered using ARToolkit for tracking of 3D objects (Billinghurst and Kato, 1999). However, the latest stable version of the ARToolkit runs on Linux platform, whereas our current visualization software, EON Professional runs on the Windows XP platform. Hence we used the MXRToolkit, a similar open source library package that runs Windows platform.

MXRToolkit works on the principle of tracking the position of the 2D marker with reference to the camera. However, 2D cards are relatively hard to grasp and the tracking will be difficult if our hands occlude the markers when manipulating cards. To surmount these problems, we designed an algorithm to track our 3D cube which has six different markers on each of its surfaces. The position of each marker relative to one another is known and fixed. Thus, to identify where the cube is, the minimum requirement is to track any of the six markers. This idea is similar to ‘multiple marker tracking’ in ARToolkit. However, instead of putting multiple markers on the same card, we extend and apply this idea to 3D artifact-cube (Zhou, 2004). Our algorithm ensures continuous tracking when our Figures happen to occlude different parts of cube during interaction, which is very likely to
happen. It allows an intuitive and direct handing of the cubes with very little constraints in manipulation whatsoever. This effectively bridges the gulf between the designers and the users of the \( \Gamma \) system.

5.2.2. Processing and Calibration
The tracking program is able to run at the frequency of 30Hz, allowing sufficient real-time update in the visualization program. A software average filter is also implemented in the program to smoothen the tracking data so as to reduce jittery that might arise from various factors such as unstable lighting which affects the video captured.

Ultimately, the \textit{MXRToolkit} will decipher the video captured into meaningful data to be used in the visualization software. By tracking the marker cube in the image of each video frame, the transformation matrix of each cube (if seen/tracked) with respect to the camera will be obtained through a series of calculation. This of course is not enough, because what we need is a relative position and orientation with respect to the floor-plan which the cubes are rested upon.

In order to achieve this, calibration is needed in the pre-programming process, so that the centre and boundary volume of the floor-plan is known by the program that computes the data. This calibration process is done by simply placing and marking the cube (in software using keyboard) on the centre and extreme edges of the floor-plan provided that the camera and the floor-plan remain fixed to each other, the calibration process need to be executed once and the data will be saved.

In essence, we only need to track each cube’s position and orientation, a total 6 degrees of freedom. However, for simplification we ignore the roll component so as the preserve the horizon and to restrict user to rotating the cube about the two other angles. Accidental rotation about the y-axis will be ignored by the program.

5.2.3. Communication
In order to feed the positions and orientations of the two cubes into the visualization software, we created a TCP/IP server-client for networking communication of the data. Using the \textit{EON Professional} Software Development Kit, we are able to implement the client structure to connect and retrieve data on the server that we implement on the \textit{MXRToolkit} processing program. Hence the designer user would just need to link up the positions/orientation data to the relevant section of the visualization program, for example the virtual camera. An overview of the communication flow is given in Figure 7.
Figure 7. The core processing units’ work flow of the I3 system

5.3. USAGE

In this section we will look in depth on the how the I3 system is being applied effectively as an interactive interface for architecture visualization process. We will look at three primary usages of the cubes that not only as a substitution but also in achieving features that are not possible before this using the conventional interfaces such as keyboard and mouse.

5.3.1. The Third Eye

As mentioned earlier, the primary application of the first cube is to be the user’s third eye in the virtual environment. What is novel here is not so much of how the cube represents the first person perspective of the user (as
can be done with keyboard and mouse), but how the cube actually translates the user’s action into a virtual viewfinder. Mouse and keyboard, while allowing user to roam around in virtual environment, do not provide an absolute reference frame for the user. A keyboard is static; a mouse movement is tracked based on the difference of current position with respect to previous position which is not continuous in time as the user often has to lift the mouse back to the original position. The lack of such a physical absolute reference frame often leads to user to lose his sense of position and orientation in the virtual environment, especially in instance where the surrounding looks almost the same everywhere such as inside a forest or an empty room.

A cube, on the other hand, solves this problem intuitively and elegantly. When the user moves and rotates the cube on top of the floor-plan, the first person view shown in the projection screen is directly reflected. For example, if the cube is placed facing west in the lobby of the floor-plan, the corresponding first-person view projected would be simply facing west in the lobby as well. Hence the floor-plan now becomes the absolute reference frame and the cube physical position and orientation mirrors directly what should be seen from that spot in the virtual environment. Any confusion can be cleared by just checking between what is being projected on the screen and where is the cube.

However, the advantages of using the cube interface do not stop here. Instead of using (and memorizing) different mouse button to change positions or rotations, the cube interface could not have been simpler; moving and rotating the cube in physical space corresponds directly to the movement and rotation of the view in virtual space. The mouse interface only provides freedom of movement in a 2 dimension plane, whereas the cube actually offers all the 6 degrees of freedom to the user. Text display on the screen is also possible to show the camera’s current absolute position and orientation relative to the current frame.

Of course, we should also notice that too much choices of freedom might sometimes confuse the user too. For example, the user might not be too comfortable holding and maintaining a cube in the mid air to view a building from the bird’s eyes view, but still he wishes to move around to see the every parts of the building from the current height and angle. To solve this problem, we can use the second cube to “lock” the desired position and orientation. In this case, by rotating the second cube 90 degrees clockwise about the vertical axis, the z-position and pitch angle of the first cube can be locked virtually, and thus the user can rest the cube back on the floor-plan to vary only the x-y positions and yaw.

Users can also predefine and recall up to 20 individual views for any virtual scene. In the predefinition-mode, the user would have to use the first cube to pinpoint the desired saving point. Once selected, the second cube
will be rotated clockwise for about 15 degrees to save that very point, with
text display on the screen to serve as prompt. In the recall-mode, the user
would just need the second cube and rotate in the same direction to cycle
through the views saved earlier. To extend this further, the points can be
interpolated to form a guided tour playback for the user to watch. The
switching of different modes (predefinition, recall and guided tour) can be
achieved by moving the cube to different quadrant of the floor-plan.

5.3.3. Third Person View
Sometimes first person view perspective might not give the best information
or idea how the place might actually look like. For example, the user is now
in office block somewhere in the middle of a tall building, and he wishes to
see himself from a third person perspective, something like a X-ray vision
that cuts through the building to see where he is exactly standing.

This can be done simply with two cubes. Using the second cube, we can
specify a certain rotation; say 90 degrees anticlockwise with respect to the
vertical axis to switch between the original first person view or the third
person view. Once we are in the third person view mode, what we are
actually doing is to give an offset to the original view we were in, with the
angle pointing to our original point, instead of where we were looking in the
first person view. In other words, our view is now locked towards our earlier
point at a certain distance which we now can vary through the rotation of the
first cube.

Rotating the first cube to zoom might not be sufficient to view where we
are, our representation in the virtual space could be occluded by wall or
other objects. Hence we deploy the second cube rotation about the horizontal
axis to vary the near cutting plane, assuming the far cutting plane is much
further behind, thus giving the user the ability to see through walls.

For a simpler third person viewing mode, an overhead bird eyes view that
looks directly downwards might be sufficient and intuitive enough to give a
clear picture.

5.3.4. Future applications - Customization of Objects
The application of the cubes for customization of objects yet another
interactive and intuitive way of sending commands from the user to the
virtual environment. Using one cube, the user could specify what he wishes
to customize, for example, color, texture, lighting condition or size through a
combination of movement or rotation. In order not to burden the user to
memorize different combination, the pattern printed on the cube should be
easily recognized for each mode, for example a painting brush pattern
representing switching color mode. Text and graphical denotations are also
displayed on the screen to act as additional aid to remind the user to switch between different modes.

The second cube is now in charge of varying the parameters of the attribute selected by the first cube. If the user selects the color attribute, moving in the x, y and rotating about the vertical axis could each independently influence the red, green and blue component of the color preferred. Or if the user wishes to change the external condition of a building, the second cube can now be treated as a virtual sun to shine on the building from different positions and angle and intensity. The options and possibilities for customization of objects that can be achieved with the two cubes are theoretically endless, although system complexity becomes a real problem when user gets confused about too many choices.

6. Conclusion

The paper introduced a new tangible interface for easy navigation through immersive Virtual Architecture to overcome the common problem of ‘Getting lost in Cyberspace’. It has conceptually been developed for general applications and was further developed and extended to meet the special requirements for easy navigation within Virtual Architecture. The integration of the 2D floor plans as contextual reference and constrains resulted in positive feedback particularly from new users, who usually would have problems operating with the limited 2-dof mouse or too complex glove. Operating the cubes to control the camera, set slanted references planes, save and recall views, and display a third person’s top view was also found to be easy and intuitive. This is important since the cubes are the key to access many combinations hosting different features. Features such as object modification or browsing through additional media such as still pictures and video are planned for the future. Having integrated this interface into the Digital Space Lab of the Department of Architecture will enable more non-expert users benefit from immersive visualization of virtual architectural space.

Acknowledgements

This project is funded by an Academic Research Fund of the National University of Singapore and supported by the Departments of Architecture and Electrical and Computer Engineering and Asia Research Institute.
References

EON Professional, Internet: http://www.EONreality.com/
MXRToolkit, Internet: http://mxrtoolkit.sourceforge.net/