

BIM Navigation with Hand-Based Gesture Control on Sites

Chao-Chung Yang¹ and Shih-Chung Kang²

¹Department of Civil Engineering, National Taiwan University, Rm. 611, No.188, Sec. 3, Xinhai Rd., Da'an Dist., Taipei City 106, Taiwan (R.O.C.); PH (886) 2-33664346; email: r01521604@ntu.edu.tw

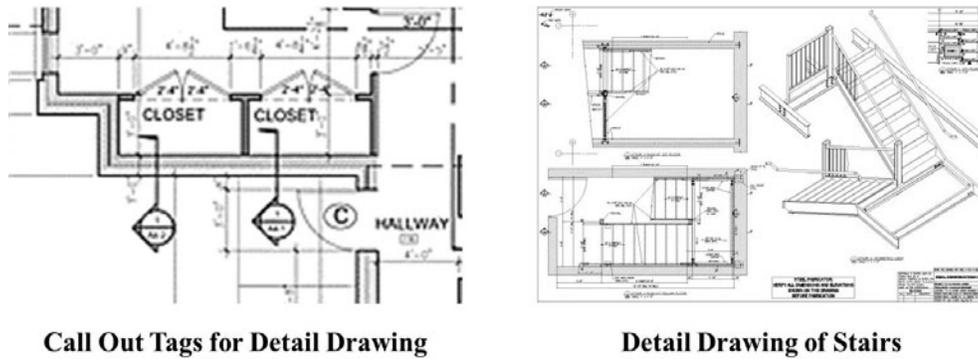
²Department of Civil Engineering, National Taiwan University, Rm. 607, No.188, Sec. 3, Xinhai Rd., Da'an Dist., Taipei City 106, Taiwan (R.O.C.); PH (886) 2-33664346; email: sckang@ntu.edu.tw

ABSTRACT

Building Information Model (BIM) technology combines a "virtual building" with extended information, removing the necessity to rely on imagination or experience to foresee the completed building. Viewing a construction project in 3-D includes spatial information that is difficult to convey in 2-D drawings and can reduce errors during construction. In addition to geometry, extended information such as the properties of building components and the relationship between them can be also included in models. However, 2-D drawings are still commonly used on sites. One main reason is that the manipulation of BIM models using a mouse is not intuitive and is often not appropriate in the site environment. This research focuses on developing an interactive interface for site workers to retrieve information from BIM models, allowing them to break through the limitations of 2-D drawings. The main feature in this research is a BIM viewer with Natural User Interface (NUI). The interface provides hand-based gesture control that allows BIM models to be navigated with ease. The main contribution of this research is to define gestures that are natural to users, and to make the manipulation of BIM intuitive, by fine-tuning the sensitivity between real movement and virtual movement. The next step of the research is to conduct a user test to validate the usability of our design.

INTRODUCTION

During the construction phase, different construction sectors mainly execute construction projects according to 2-dimensional drawings. In traditional construction drawings, detail drawings are often used to supplement details, which cannot be described on limited drawing areas. Figure 1 shows the "Call Out Tag" noted on drawings to indicate that related detail drawings are available, and an example of a detail drawing of a stairs. The detail drawing contains multiple views of target components and detail dimensions. However, it is still difficult to understand geometrically complex components well on 2-D drawings. 3-D views allow a better understanding of shapes and space and present the information in a more efficient manner. (John *et al.* 2001).



Call Out Tags for Detail Drawing

Detail Drawing of Stairs

Figure 1. Detail Drawing

Building Information Modeling (BIM) refers to the digital representation of a building throughout the project life-cycle. In comparison with a traditional building design process using 2-D drawings, BIM not only contains geometrical information in 3-D, but also covers functional characteristics for further analysis in digital format. Shen and Chua claim that BIM is an enabling technology that aids coordination among various specialists, reducing fragmentation through the employment of a 3-dimensional geometrical model with integrated value-adding services for each phase of project development (Shen and Chua 2011). However, the usage of BIM models in the construction phase is still limited; it is still common for site workers to retrieve design details from 2-D drawings. Despite this, many researchers still make great efforts to implement BIM on site.

Davies and Harty propose the concept of "SiteBIM" and describe an empirical case-study of the implementation of the innovative system on a tablet PC (Davies and Harty 2013). Olbrich et al. describe a new Augmented Reality framework that offers on-site access to BIM information and user centric annotation mechanism, merging virtual and real buildings on mobile devices (Olbrich *et al.* 2013). However, the above devices might be vulnerable on site due to the dusty environment. On the other hand, the target users of the above tools are mainly project managers and designers on-site. On-site workers usually need to wear work gloves and it is therefore inconvenient to use contact controllers or devices. Furthermore, due to the lack of degrees of freedom of a standard mouse for 3-D manipulation, many researchers worked on developing 3-D input devices, like trackball, 3-D mouse, dataglove and evaluating their performance (Hand 1997).

RESEARCH BACKGROUND

Natural User Interface (NUI) refers to the human-machine interface that is effectively invisible. Steve Mann uses the word "Natural" to refer to an interactive method that comes naturally to users, and the use of nature itself and the natural environment (Mann and Rowe 2002). Multi-Touch technology is a well-known example of one such interactive method, and gesture control allows users to interact with a computer naturally without the need for a controller. Atia and Tanaka focused on developing a tilting gesture for direction control in ubiquitous environments and their system was shown to perform well (Atia 2010). Thus, this research takes advantage of the non-contact features of these technologies along with their inherent intuitive properties and applies these concepts to the manipulation of BIM models.

RESEARCH GOALS

Our research aims at developing a hand gesture-based BIM model viewer, targeted toward onsite workers. The interface allows users to retrieve building data with ease. The two main goals of this study are outlined below:

1. Define the hand gestures for model manipulation and information retrieval
2. Tune the sensitivity

The structure of the paper is as follows: first, the defined hand-gestures for BIM models and manipulation functions are introduced; then, we describe how we implemented the proposed method. Last, we give the conclusions of the current research and discuss future work.

METHODOLOGY

In this research, five model manipulation functions were targeted: zoom, rotate, orbit, pan, and select. Table 1 shows the requirements of each function for different types of task on site. Each of the functions corresponds to defined control gestures and is described below:

Table 1. Tasks and Functions

Functions \ Tasks	Zoom	Rotate	Orbit	Pan	Select
Construction Scheduling	●		●	●	●
Design Interpreting	●		●	○	
First Person Navigation	●	●		○	○

●: Required ○: Dispensable

Manipulation Functions

- (1) Zoom

In a virtual environment, the zoom action changes the distance between the models and the camera, moving along the view vector of camera. In equation (1), p_i is the position of the camera and v_i is the view vector. Variable s determines the zooming speed and p_i' is the new camera position after zooming.

$$p_i' = p_i + sv_i; i = x, y, z; \tag{1}$$

- (2) Rotate

The rotate function is used in first person mode. Users control the view direction of camera by a defined gesture. In this research, quaternion-based camera rotation is used. The view quaternion is described as quaternion \mathbf{v} in the equation (2), which is a combination of the x, y and z values of the view vector and w of the rotation angle for the scalar component.

$$\mathbf{v} = v_x\mathbf{i} + v_y\mathbf{j} + v_z\mathbf{k} + w \tag{2}$$

The rotate axis is also described in quaternion format in equation (3), including four variables: components of axis vector u_x, u_y, u_z and a rotation angle θ .

$$\mathbf{q} = e^{\frac{1}{2}\theta(u_x\mathbf{i}+u_y\mathbf{j}+u_z\mathbf{k})} = (u_x\mathbf{i} + u_y\mathbf{j} + u_z\mathbf{k}) \sin \frac{1}{2}\theta + \cos \frac{1}{2}\theta \quad (3)$$

In the equation (4), \mathbf{v}' is the new view quaternion after rotating.

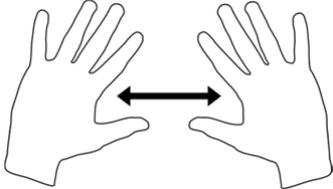
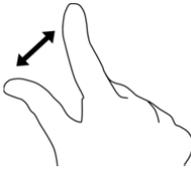
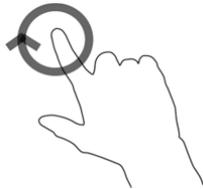
$$\mathbf{v}' = \mathbf{q}\mathbf{v}\mathbf{q}^{-1} \quad (4)$$

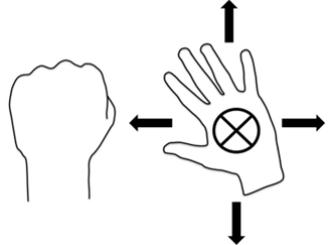
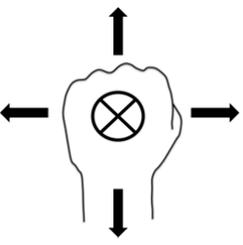
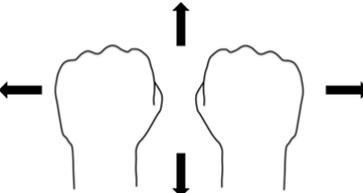
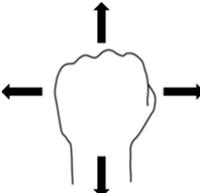
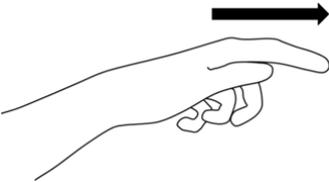
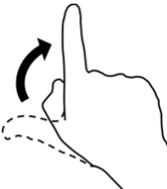
- (3) Orbit
The orbit function makes the camera revolve around a target model or component. The camera is positioned in spherical coordinates and further detail of its movement will be described in the implementation section.
- (4) Pan
The pan function allows the camera to move horizontally or vertically along the right vector and up vector of the camera.
- (5) Select
To retrieve further information, users are allowed to select objects using this function.

Gesture Definition

In Table 2, the gestures for the five functions mentioned are defined.

Table 2. Functions and Gestures

Functions	Gestures		
Zoom			
Rotate			

Orbit		
Pan		
Select		

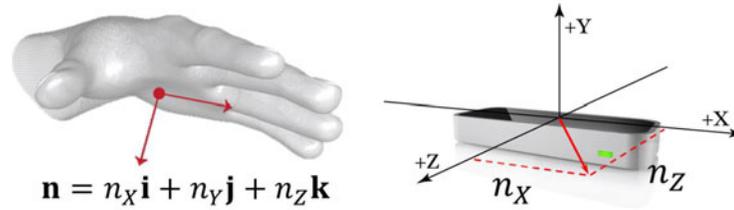
IMPLEMENTATION

To implement the proposed interface, Leap Motion API and Navisworks API are used to develop the software in the C# programming language (figure 2). Leap Motion is a computer sensor device, detecting the motions of hands, fingers and finger-like tools as input, and the Leap motion API allow developers to get tracking data for further uses. Navisworks project review software for AEC industries integrates building models and data, and API supports camera control in a virtual environment. Currently, we have implemented three of the functions mentioned above: rotation, orbit, and zoom.



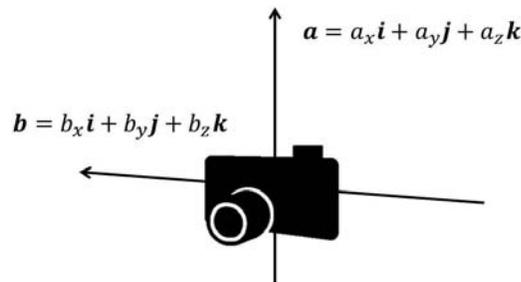
Figure 2. Field Test

- (1) Rotation
 "Palm tilting" gestures are defined to control rotation of the camera. First, vector **n** is the normal vector of the palm (figure 3) in the real world, and vector **a** and **b** refer to the up and right vectors of a virtual camera (figure 4).



$$\mathbf{n} = n_x \mathbf{i} + n_y \mathbf{j} + n_z \mathbf{k}$$

Figure 3. Normal Vector of Palm



$$\mathbf{a} = a_x \mathbf{i} + a_y \mathbf{j} + a_z \mathbf{k}$$

$$\mathbf{b} = b_x \mathbf{i} + b_y \mathbf{j} + b_z \mathbf{k}$$

Figure 4. Up Vector a and Right Vector b

Then, we use the x and z component (n_x, n_z) of vector \mathbf{n} to determine the addition ratio of vector \mathbf{a} and \mathbf{b} in equation (5) below, and to determine the rotate angle in equation (6) below. The addition result of \mathbf{a} and \mathbf{b} is the rotate axis vector \mathbf{u} in equation (3).

$$\begin{bmatrix} u_x \\ u_y \\ u_z \end{bmatrix} = \begin{bmatrix} a_x & b_x \\ a_y & b_y \\ a_z & b_z \end{bmatrix} \begin{bmatrix} \frac{n_x}{|n_x| + |n_z|} \\ n_z \\ -\frac{n_z}{|n_x| + |n_z|} \end{bmatrix} \tag{5}$$

$$\theta = \sqrt{n_x^2 + n_z^2} \tag{6}$$

(2) Orbit

The orbit function needs to be transformed from Cartesian coordinates in the Navisworks environment to spherical coordinates. d_x, d_y and d_z are distance components between the camera and the target model, and are transformed into three parameters with spherical coordinates as shown in figure 5 and equation (7)

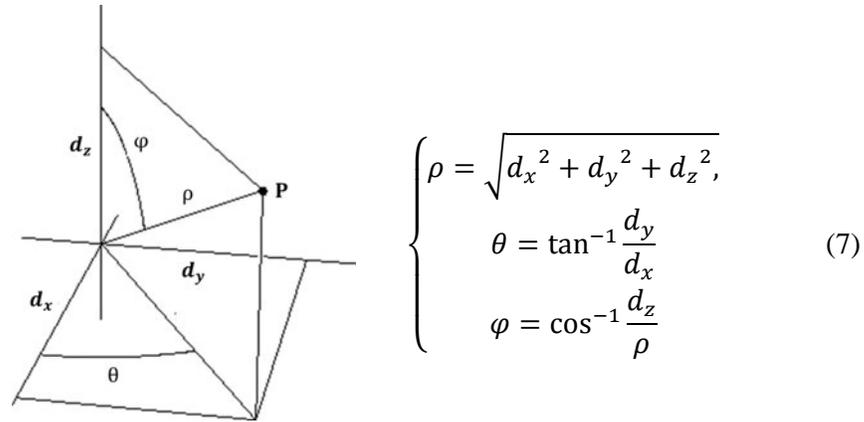


Figure 5. Coordinate Transformation

Equation (8) shows that the position change of a palm in the x and z direction is translated into change of θ and φ to orbit around a model. R_θ and R_φ is the tuning ratio, and in this research, we tune the value to 30 and 50.

$$\Delta\theta = \frac{\Delta X}{R_\theta}, \Delta\varphi = \frac{\Delta Z}{R_\varphi} \tag{8}$$

(3) Zoom

To adjust the zooming speed in equation (1), two factors are considered: the change in distance between the center of two palms or finger points and the distance between objects and the camera in equation (9). R_z is tuned to 50 here.

$$s = \Delta D \frac{\rho}{R_z} \tag{9}$$

$\Delta D = \Delta(\text{Distance between Center of Two Palms or Finger Points})$

CONCLUSION

This research aims to develop a hand gesture-based interface for BIM model navigation. The two main contributions of this study are control gesture defining and sensitivity tuning. However, an assessment of actual user experiences should be gleamed from real users. In the future, we plan to validate the usability of this system by conducting a user test on a construction site for on-site workers. Five points of usability would be of concern: (1) Learnability: difficulties when encountering the design for the first time, (2) Memorability: difficulties with reestablishing proficiency after a period of not using the system, (3) Efficiency: performance after users have learned the system, (4) Error: number and severity of errors and (5) Satisfaction: comfort level of the design.

The remaining two manipulation functions will also be completed in future works. In addition, we intend to develop the distance estimation function, which

makes users' hands visible in the virtual environment when exploring, allowing the actual distance between users and objects to be experienced.

REFERENCE

- Atia, A. (2010). "Interaction With Tilting Gestures In Ubiquitous Environments," *arXiv preprint arXiv:1007.5158*.
- Davies, R. and Harty, C. (2013). "Implementing 'Site BIM': A case study of ICT innovation on a large hospital project," *Automation in Construction*, 30(0), 15-24.
- Hand, C. (1997). "A survey of 3D interaction techniques," *Proceedings of Computer graphics forum*, 269-281.
- John, M. S., Cowen, M. B., Smallman, H. S. and Oonk, H. M. (2001). "The use of 2D and 3D displays for shape-understanding versus relative-position tasks," *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 43(1), 79-98.
- Mann, S. and Rowe, P., *Intelligent image processing*: IEEE, 2002.
- Olbrich, M., Graf, H., Kahn, S., Engelke, T., Keil, J., Riess, P., Webel, S., Bockholt, U. and Picinbono, G. (2013). "Augmented reality supporting user-centric building information management," *Visual Computer*, 1-13.
- Shen, L. and Chua, D., K.H. (2011). "Application of Building Information Modeling (BIM) and Information Technology (IT) for Project Collaboration," *Proceedings of International Conference on Engineering, Project, and Production Management*.