# 3D User Interface Combining Gaze and Hand Gestures for Large-Scale Display

**ByungIn Yoo**
Multimedia Lab, SAIT,
Samsung Electronics Co., LTD.
Yongin, Gyeongi, Korea
byungin.yoo@samsung.com

**Sungjoo Suh**
Multimedia Lab, SAIT,
Samsung Electronics Co., LTD.
Yongin, Gyeongi, Korea
Sungjoo.suh@samsung.com

**Jae-Joon Han**
Multimedia Lab, SAIT,
Samsung Electronics Co., LTD.
Yongin, Gyeongi, Korea
Jae-joon.han@samsung.com

**Dusik Park**
Multimedia Lab, SAIT,
Samsung Electronics Co., LTD.
Yongin, Gyeongi, Korea
dusikpark@samsung.com

**Changkyu Choi**
Multimedia Lab, SAIT,
Samsung Electronics Co., LTD.
Yongin, Gyeongi, Korea
Chankyu_choi@samsung.com

**Changyeong Kim**
Multimedia Lab, SAIT,
Samsung Electronics Co., LTD.
Yongin, Gyeongi, Korea
cykim@samsung.com

**Kwonju Yi**
Multimedia Lab, SAIT,
Samsung Electronics Co., LTD.
Yongin, Gyeongi, Korea
Kj.yi@samsung.com


*CHI 2010*, April 10–15, 2010, Atlanta, Georgia, USA.
ACM 978-1-60558-930-5/10/04.

## Abstract
In this paper, we present a novel attentive and immersive user interface based on gaze and hand gestures for interactive large-scale displays. The combination of gaze and hand gestures provide more interesting and immersive ways to manipulate 3D information.

## Keywords
Gaze Gesture, Freehand Gestures, Immersive UI, Attentive UI, Interactive Wall Display

## ACM Classification Keywords
H.5.2. [Information Interfaces and Presentation]: User Interfaces-Interaction styles, Input devices and strategies.

## General Terms
Design, Human Factors

## Introduction
Large-scale displays come into reality, as displays become larger and thinner. Among the usages of the large-scale display, it is expected to be used as new media and information windows in theaters, offices, and even house-holds with the form factor of a wall. The large-scale displays provide a rich experience and some benefits to users [3]. Therefore, adding an intuitive natural interaction method to the large scale display provides a more immersive experience to the users. New types of the following applications are expected [9]:

- Information Browser
- Virtual Tour

Information browser provides various types of representation of information utilizing the large scale screen and an UI which users can manipulate easily by treating each of the digital contents on the screen as the real object in the real world. Virtual tour is a virtual environment which reconstructs a famous place virtually. It provides a sensation to users as if the users visit the place in advance. The above two applications at the interactive wall type display require an attentive and immersive 3D UI by tracking the gaze and hand gestures to provides a natural interaction to users [12].

The wall-type display, which is larger than a reach of each human, requires a new type of interaction methods other than the conventional keyboards and mice. Multi-touch [10] is a potential interaction method but has a drawback when some widgets on the screen are located "too far to reach" relative to the users. In addition, while touching the screen, the whole screen is not visible to the user [8]. To solve these issues, some researchers provide solutions such as "Extending Touch [1]", and "Shadow Reaching [8]" to facilitate manipulation over long distance. Luc and Jasper also presented two-handed interaction for presentation in an open space [6]. However, the pointing metaphor of 2D interaction with the two hands only does not provide enough immersive experience. "Depth Touch [7]" and "Interaction in the air [2]" shows hand based 3D gestures that manipulate 3D digital object with spatial touch information. On the other hand, it gives more freedom to the users on manipulating the 3D object but does not increase attentiveness to the users due to the lack of the knowledge about users' intention. To improve the pointing accuracy, Kai and Reiner adapted head orientation to pointing gesture recognition [11], but it isn't intended to apply to a user's attention. Moreover, since the wall-type interactive display has a wide range, it is necessary to move control points such as cursors quickly to the area of user's interest. Therefore instead of moving the pointer or other types of control points based on the hand only, the systems need to respond attentively where the user is looking.

The paper proposes a novel fusion interaction at about three meter distance off the screen, which combines the gaze and the hand gesture for the user to manipulate the structured 3D contents immediately and its UI method which is attentive and immersive.

## Interactive Wall Display

The wall type interactive display of the paper is composed of two components, i.e., a display unit and a sensing unit. The display is a rear projection display (4m x 1.8m x 3m). The sensing unit is composed of three CCD color cameras and a Time of Flight (TOF) depth camera (Figure 1).

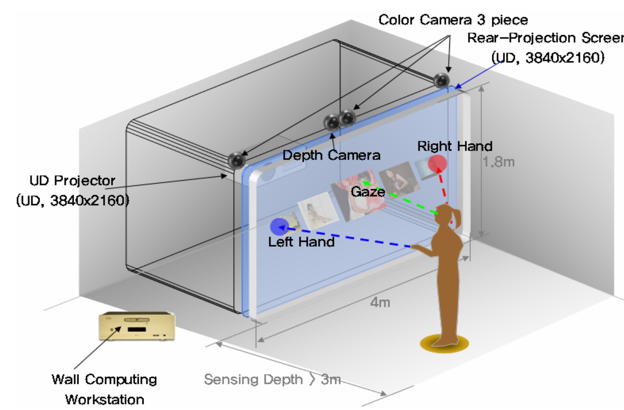


**Figure 1.** Interactive Wall Display Architecture

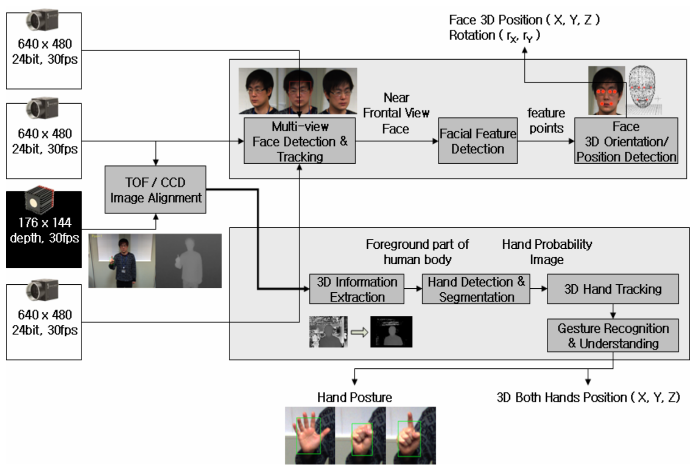


**Figure 2.** Gaze and Freehand Gesture Sensing Framework

The sensing S/W framework is shown in Figure 2. There are two main modules in addition to TOF/CCD image alignment module. After aligning the two images from the centered TOF/Color cameras, the depth information from the TOF camera is used to segment the user's body silhouette from the background. The module on the upper side of Figure 2 estimates the face 3D position (x, y, z) and orientation (pitch, yaw). The module adapts the Support Vector Regression algorithm. The resulting performance of the algorithm can achieve 10 degrees of accuracy and the range of the head orientation from -45 to 45 degrees for yaw angle, and from -30 to 30 degrees for pitch angle. The other module on the lower side of Figure 2 tracks the hands position (x, y, z) and recognizes three kinds of hand postures. The Camshift algorithm is used for tracking hand position. The fingertip detection and the improved Shape Context features are combined for recognizing postures. Both hands and gaze information can be sensed in about 15fps simultaneously [5].

## Gaze and Freehand Interaction

The orientation and position of the head and the positions of the two hands which are acquired by the sensing S/W framework are transferred to the UI with formatted abstract events. The first type of the abstract events is a gaze input.

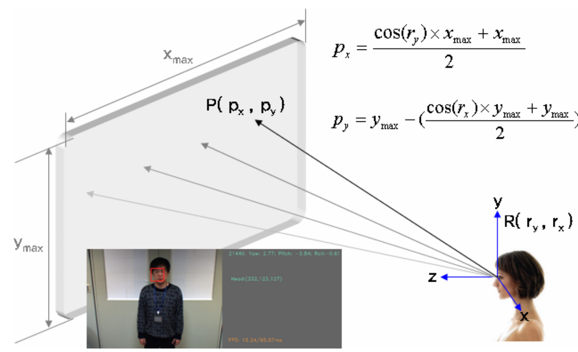


**Figure 3.** 3D gaze gesture

Figure 3 shows that a head orientation as a gaze gesture. The head orientation is first converted to a position on the screen coordinate by the equations in the figure and then used as the gaze input.

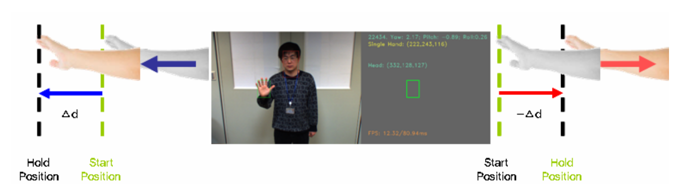


**Figure 4.** 3D push-pull hand gesture

The second type of the abstract events is a zooming and movement control input. The 3D push-pull hand gesture in Figure 4 is chosen as the input because humans engage machines by pushing buttons and pulling the levers to execute a command to a machine. The difference (▵ d) between start and hold position in Figure 4 is used for controlling the acceleration of zooming or movement.

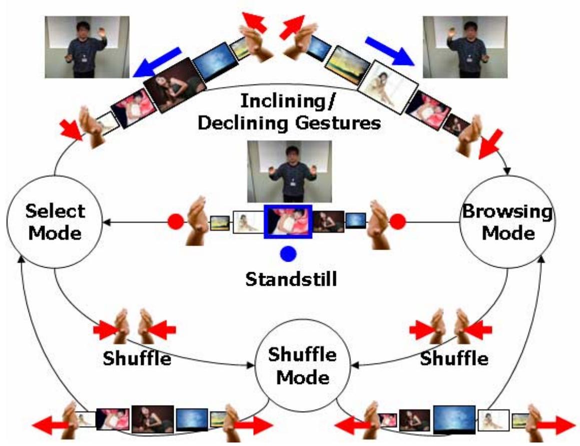


**Figure 5.** 3D bimanual gestures

The third type of the abstract events uses 3D bimanual gesture which is a natural interaction where a user can grasp the both ends of 3D digital object bundle virtually like a real card deck with both hands (Figure 5). There are 3 states including selecting, browsing and shuffling. The state of the events changes depends on the 3D positions of both the hands.

## 3D User Interface

The designed gestures are applied to UI applications. Attentive virtual tour (Figure 6) is an application that a user can travel virtually using gaze and hand gestures. The user can select a region of interest in 2D map with a search light (a bright circle in Figure 6) by 3D gaze gesture, and then control the zoom of the region of interest by 3D push-pull gesture. After diving into 360° panoramic view mode, the user can change a viewpoint by gaze and zoom by push-pull. The user experience is similar to flying over the region immersively without any physical walking and change the viewing angle attentively with respect to the gaze.

**Figure 6.** 2D map with a search light / 360° panoramic view

**Figure 7.** 3D hallway browser / 3D card browser

Immersive information browser is an application with which a user can browse and navigate the digital information using gaze and both hands. Figure 7 shows two kind of immersive information browsers for wall type displays. The one on the left-hand side of Figure 7 is a 3D hallway browser, which can change viewpoint by the 3D gaze input and walk forward/backward along media contents hallway by 3D push-pull gesture. The one on the right-hand side of Figure 7 is a 3D card browser by using 3D bimanual gesture, the 3D cards bundle can be placed various positions including z-position in a 3D virtual space. The view point of 3D card browser is changed perspectively and scrolled automatically according to the spatial positions of both hands. When the both hands are at a standstill horizontally (Figure 5), digital objects aren't scrolled and stopped. The user feels an immersive browsing experience in a 3D virtual space with natural gestures without having to physically walk.

## User Study

We conducted a user study aimed at evaluating the proposed gestural interaction and UI. A total of five people participated in the initial experiment. The first task is that participants choose the specific region on a 2D map with the search light and then browse a 360° panoramic view. The second task is that participants find certain pictures in the 2D file browser, 3D hallway browser and 3D Card browser. All of the participants execute the two tasks at three meters away from the screen with a wireless keyboard and a wireless mouse. Then, the same tasks were performed with the proposed user interface combined with the gaze and hand gestures. After the experiments, the participants were asked to score points from 1 to 9 for the following four questions. How convenient to use? How much do you feel fatigue to interact? How interesting is each of the interactions? Which interaction method do you prefer?

The result of the experiment on the "Attentive virtual tour" shows in Figure 8 that the proposed interaction method makes users more interested and is preferred, while the level of convenience has shown little difference. The result of user study on the "Immersive information browsers" shows in Figure 9 that 3D hallway browser is the most interesting and preferring browsing method.

Note that the users felt more fatigue on the interaction methods using gaze and hand gestures than the conventional WIMP-based interaction method such as keyboards or mice. Especially users felt more fatigue on 3D Card Browser. It is due to the fact that the users were required to stretch out both hands to manipulate digital objects. Therefore it is recommended to find a method minimizing the fatigue when a hand gesture interaction is considered as an input.

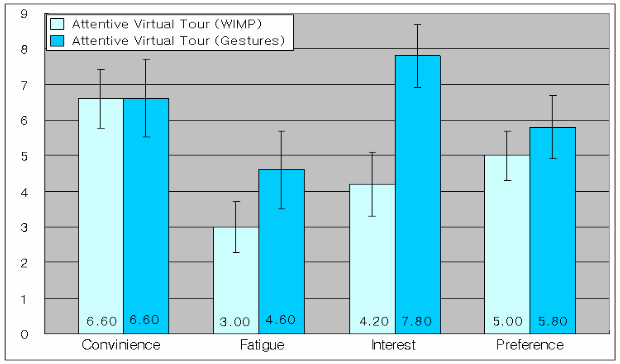


**Figure 8.** The result on Attentive virtual tour

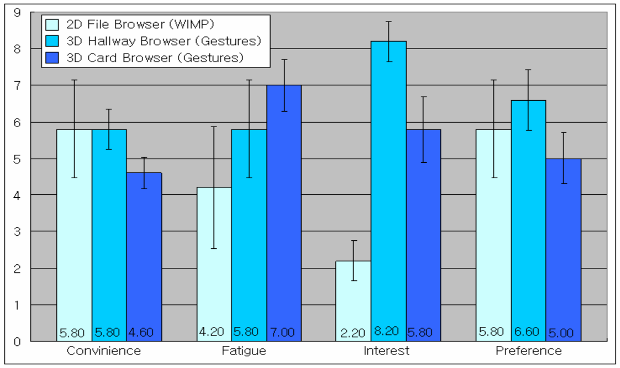


**Figure 9.** The result on Immersive information browser

## Conclusion

In the paper, we proposed three meters away gestures and 3D UI for interactive large-scale displays. The results suggest that users prefer the UI combining gaze and hand gestures manipulating 3D structured contents on the screen due to the fact that the combined UI makes more attentive and immersive than the conventional UI. We are currently designing more experiments to support the results and doing more analysis to test if the proposed methods improve usability. There are still remaining issues to solve such that users feel fatigue while using hands in the air, which will enable the gestural interaction to be commonly used.

In the future work, we will investigate a novel gestural interaction for 3D large-scale displays to improve immersiveness and a sense of control and communion when a user manipulates 3D objects on the screen.

## References


[1] Alexander Schick, Florian van de Camp, Joris Ijsselmuiden and Rainer Stiefelhagen, Extending Touch: Towards Interaction with Large-Scale Surfaces. ACM International Conference on Interactive Tabletops and Surfaces 2009, Banff, Alberta, Canada.

[2] Otmar Hilliges, Shahram Izadi, Andrew D. Wilson, Steve Hodges, Armando Garcia-Mendoza and Andreas Butz, Interaction in the Air: Adding Further Depth to Interactive Tabletops. ACM Symposium on User Interface Software and Technology 2009.

[3] Xiaojun Bi and Ravin Balakrishnan, Comparing Usage of a Large High-Resolution Display to Single or Dual Desktop Displays for Daily Work. CHI2009, Boston, MA, USA

[4] Ricardo Jota, João M. Pereira and Joaquim A. Jorge, A Comparative Study of Interaction Metaphors for Large-Scale Displays. Works in Progress CHI 2009, Boston, MA, USA

[5] Xi-Ying Wang, Hai-bing Ren, Jing Wan, ByungIn Yoo, Chang-Kyu Choi and Ji-Yeun Kim, Gesture recognition for Large Display of Next Generation. Samsung Conference 2009.

[6] Luc Vlaming, Jasper Smit and Tobias Isenberg, Presenting using Two-Handed Interaction in Open Space. IEEE Workshop on Tabletops and Interactive Surfaces 2008.

[7] Hrvoje Benko and Andrew D. Wilson, DepthTouch: Using Depth-Sensing Camera to Enable Freehand Interactions On and Above the Interactive Surface. IEEE Workshop on Tabletops and Interactive Surfaces 2008

[8] Garth Shoemaker, Anthony Tang and Kellogg S. Booth, Shadow Reaching: A New Perspective on Interaction for Large Wall Displays. ACM Symposium on User Interface Software and Technology 2007.

[9] Doug A. Bowman, Jian Chen, Chadwick A. Wingrave, John Lucas, Andrew Ray, Nicholas F. Polys, Qing Li, Yonca Haciahmetoglu, Ji-Sun Kim, Seonho Kim, Robert Boehringer and Tao Ni, New Directions in 3D User Interfaces. The International Journal of Virtual Reality, 2006, 5(2):3-14

[10] Jefferson Y. Han, Low-Cost Multi-Touch Sensing through Frustrated Total Internal Reflection. ACM Symposium on User Interface Software and Technology 2005.

[11] Kai Nickel and Rainer Stiefelhagen, Pointing Gesture Recognition based on 3D-Tracking of Face, Hands and Head Orientation. ICMI'03, Vancouver, British Columbia, Canada.

[12] Roel Vertegaal and Guest Editor, Attentive User Interfaces. Communications of the ACM, March 2003, Vol.46.No.3.