Collimation Using Transparent Projection Screen for Augmented Environment HUDs

Divya Udayan J[*] ,HyungSeok Kim[*1] ,Mu Wook Pyeon[2]

[*]*Department of Internet & Multimedia Engineering, Konkuk University, Republic of Korea*
[1] *IMI, Nanyang Technological University, Singapore*
[2]*Department of Advanced Technology Fusion, Konkuk University, Republic of Korea*

Abstract

HUD is a visual media technology in which 3D virtual objects are integrated into the real world in less intrusive way. It can provide information about the status and location of the surrounding environment in real time. This paper proposes an AR-HUD system to visualize the "out-of-the-window" scene and spatially align the real and superimposed objects i.e. collimation in accordance with the user's line of sight. This is achieved through mapping the frames of references of the real site and the HUD system. The interaction between the HUD and the user is provided using an optimized head tracking approach. Observation of the people and vehicle movements in the outside world is dynamically represented in our AR HUD system. The system is further enhanced with the use of transparent projection screen with varying voltage regulation capability to adjust the transparency of the screen for observing distant objects. To evaluate the feasibility of this system, we have considered a construction site scenario which is populated with buildings, equipments, workers, and vehicles. Preliminary evaluation of the initial field tests quantify the performance of our system, compared with the real far view.

1. Introduction

Recently, different types of visualization systems are developed for many purposes like education, training of air pilots, museums, urban planning and so on [1,2,3]. The common aim of these systems is to produce graphical images conveying relevant information to the viewer. Augmented reality plays an important role in enhancing these visualization systems. To provide the user with the real feel of place, it requires detailed knowledge of the relationship between the frames of reference for the real world, the display screen and the user. The graphics displayed in the transparent screen should be spatially aligned with the real objects in the background, providing a user feel that the display and the background are integrated. This alignment process is known as registration which requires the spatial knowledge of the exact positions of the objects in the real environment and the user position and movements in order to compute the correct perspectives.

In this paper, we consider augmented reality based guidance system for urban planning and construction. Safety is the main concern in construction fields as construction field environments are getting more complex. Heavy equipments such as tower cranes, forklift and sand diggers can increase the risk of accidents. Currently the information regarding invisible/hidden objects and small materials are provided with hand gestures or shouting which cannot be precisely recognizable in real time on the exact position of events due to long distance and small size of gestures. Thus there arises the need for a remotely monitoring tower for identifying all materials on ground. In our system, we propose a spatially calibrated optical see-through AR-HUD system so that the information requested by the HUD is retrieved from a server and presented at the aligned position where the user is looking at. This system can also provide information about future completed buildings by presenting the 3D model of the future building at the exact position where the building is currently being constructed. The system also visualizes safe/dangerous areas, movement of workers and vehicles in the construction site through a transparent screen. Figure 1. shows the AR user interface in a typical construction site. HUDs in construction site should not obstruct the operations of ordinary workers or supervisors who require both hands for operations. Our AR-HUD differs from the conventional HUDs by providing a hands free way of controlling the system. Also, the interactions with the user are much more improved by providing a detailed knowledge about the frames of references of real scene, display and the user.

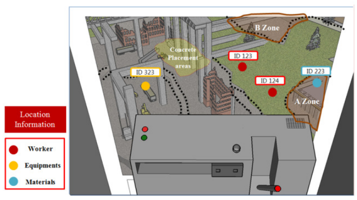


Figure 1. AR user interface in a construction site

CollabTech 2012 , August 27-29, 2012, Hokkaido, Japan.

2. Related Work

There have been many approaches for utilizing HUD system for minimally intrusive AR systems. Peterson et.al [4] conducted experiments with camera-based systems to detect driver and dynamic environments with an active visual display in a side screen. But, the side screen requires the user to look away from the road to get the information. Liu and Wen [5] have shown a reduction time of 0.8 to 1.0 secs in driver reaction time while using HUD to control their speed when compared with heads-down display(HDD). A. Doshi et.al [6] have suggested a novel laser-based wide area heads-up windshield display which provides active interface for a driver assistance system. In their work, information on speed and road are provided to driver on the front window of vehicle. They have represented dynamic environments in real-time on site. Therefore, we adopt similar ideas to our HUD system for construction site.

Wen Zahn Song et al. [7] suggested Optimized Autonomous Space In-Situ Sensor Web for volcano monitoring . They used sensor networks to monitor necessary information. Myoungjn Kim et.al [8] suggested a distributed real-time system for U-GIS informative construction which uses TMO (Time-triggered message-triggered object) for multiple processing of data. They suggested a method for getting information about equipments and situations of construction site. Tracking of user head can be achieved using predefined markers and provides a method to augment virtual objects on tracked positions [9]. But, in construction site, due to limited visual detail and its dynamic characteristics it is not suitable to use marker-based approaches. G. Klein et. al [10] suggested parallel tracking and mapping for indoor workspace. AR-HUD system that we propose in our work is closely related with the view of a tower crane driver. So we have chosen kinect sensor based head tracking method [11]. Also HUD system requires to precisely represent dynamic movements in the construction site.

Figure 2. See-through window display with voltage regulation

In this paper, we propose AR HUD as a new display component for see-through window displays(Figure 2), which detaches the display technology from the user and integrate it with the background real scene [12,13]. Compared with head or body attached displays, see-through window displays improve visual effects like resolution, field of view and focus but they are limited to non-mobile applications. The availability of projection technology, personal computers and graphics hardware makes See-through Window displays more popular for non-mobile applications.

3. Real Time Registration and Tracking

The proposed system consists of: User Tracking module, Environment Tracking module and the Visualization module as shown in Figure 3.

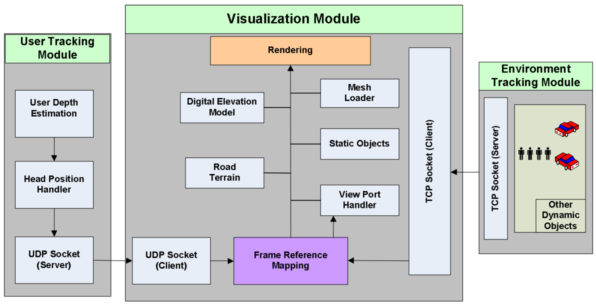


Figure 3. AR-HUD system design

In this paper, we use the Microsoft's Kinect[11] sensor data to explore the feasibility of head tracking on a smaller scale(Figure 4). The tracking of the head through a sequence of images consists of estimating the parametric representation of the head. For each time period, the previous time parametric representation of the head is known and used as prior knowledge (Figure 5). This information has to be optimized i.e. tracking can be defined as finding an optimal state configuration X_{opt} verifying:

$$X_{opt} = arg\ min_X\ D(h_{obj}, h_X)$$

where h_{obj} is a model of the head object in the observation space, h_X is the number of observations extracted from the candidate state configuration X, and D is the distance in the observation space.

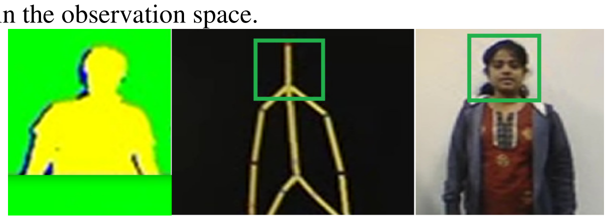


Figure 4. Head Tracking using depth cues from kinect sensor

The optimal state configuration can be obtained using gradient descent method. Here the initial state of the optimization process is taken as the optimal state of the previous time (t-1) or a prediction at time (t+1). As mentioned in the mean shift algorithm for tracking, proposed in [14], the model of the head object h_{obj} is represented by a kernel density distribution. At the candidate state configuration X, the head state observations h_X is also represented as a kernel density distribution. The distance in the observation space is obtained using Bhattacharya distance [15].

$$D(h_{obj}, h_X) = \sqrt{1-\int_u \sqrt{h_{obj}\,u)h_X\,(u)du}}$$

Using this method, we find the best tracked position of the head. Advantage of using Bhattacharya distance are that it is computationally simple and it provides a smoothed distance between the two above mentioned distributions, which is more appropriate for head tracking.

The Environment tracking module consists of GPS tracking unit that uses the global positioning system to determine the precise location of vehicles or workers in the construction site. The recorded location data is stored within the tracking unit and transmitted to a central server data base using cellular modem embedded in the unit. The server transfers the data to the visualization module using TCP protocol for further mapping the GPS co-ordinates to screen co-ordinates.

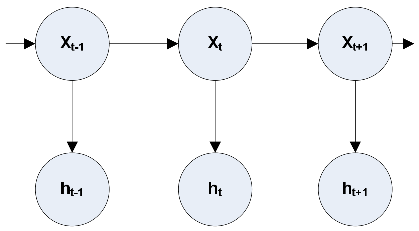


Figure 5. Standard graphical model for tracking

4. Visualization Module

The visualization module presents the graphical information to the user more precisely and exactly. The visualization module consists of information received from the GPS sensor, the mesh loader and the information about view point mapping. Figure 6, represents the conceptual model of the visualization module. Here the scenario under consideration is a construction site. Location information of equipments, workers and materials are to be represented by the system. This information is send to the module at real time by the server. The construction area is divided into different

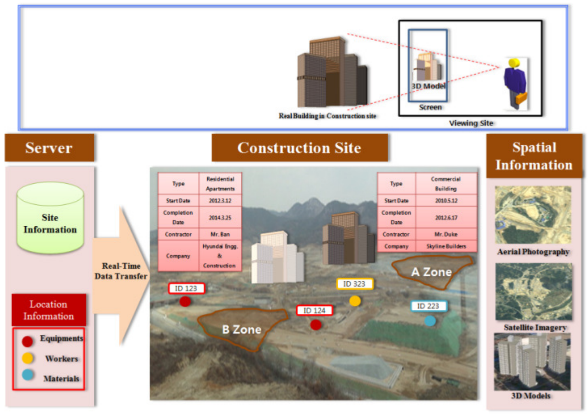


Figure 6. Conceptual model of Visualization module

zones for area distinction. The 3D models of the ongoing construction are presented to the user at real time when the user's head points to the location in geographical space. The buildings are provided with labels showing the type of building, start and completion dates, contractor details and the construction company involved in the work process. The spatial information about the construction site which is obtained from satellite imagery are provided in the server database. The visualization module also processes the information received from the head tracking module and environment tracking module and does a view point mapping for all frames of references in order to project the virtual objects in the correct location on the transparent screen. The visualization module receives the dynamic information from the server using socket programming and we define data in the form of opcode for transferring information as shown in Table1.

Table1. Definition of op-codes

Type of connection	Data	OPCode
Server->HUD	Connection Response	0x80
	FTP access information response	0x81
	Material information request/response	0x90
	Zone information request/response	0x91
	Moving object information	0x94
HUD->Server	Request for connection	0x60
	FTP access information request	0x70
Server<->HUD	(N) ACK	0xA0

To create an illusion that virtual objects are registered to real background scene for a user with head movements, we need to know the position of user head, projection parameters of the display devices and the position of the real objects in the physical environment. Thus a mapping

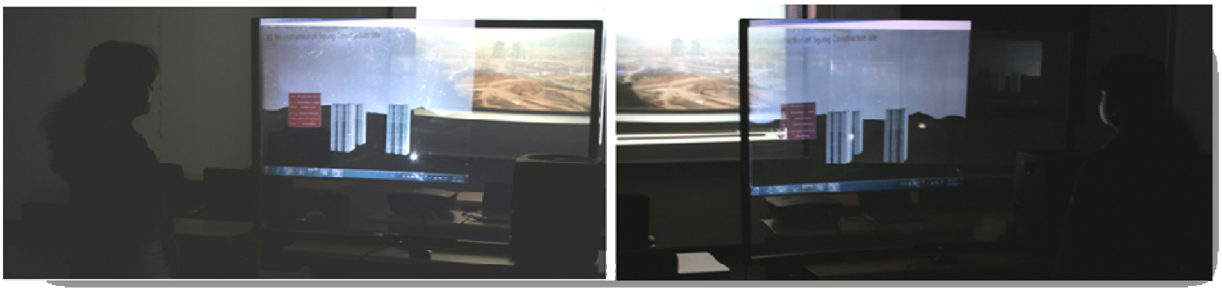


Figure 7. The experimental results showing collimation of real apartments under construction in the site and completed 3D model of the apartments viewed through a transparent projection screen

procedure is required to find the perspective projection matrix M that represents the relationship between real world co-ordinates, HUD screen co-ordinates and user co-ordinates. Kinect user positions are expressed in right handed cartesian coordinate system. The x, y, and z axes(expressed in meters) are the body axes of the depth sensor. Mapping between the world co-ordinates, kinect camera co-ordinates and HUD screen co-ordinates are represented by rotation and translation transformations which is represented by the transformation matrix .

After the mapping procedure is completed, the graphical information can be rendered using the transformation matrix. We now present a process for rendering and viewing method. Let 'I' be the image generated from 3D colored model 'G' using perspective projection matrix 'M'. Here the notation $M^{-1}*I(r,g,b)$ represents a set of colored rays and the notation $M^{-1}*I(r,g,b,z)$ represents colored surface with the corresponding depth. Compute the image parameters for rendering G

$$I(r,g,b) = E * [G]$$

where E is the projection matrix which shares the center of projection with the user's head and G is the 3D colored graphical model. Next update the depth buffer

$$I(z) = E * [S]$$

where S is the display surface model. Finally apply the view transformation from projection system defined by E to the projection system defined by the projector's perspective projection matrix P

$$I'(r,g,b,z) = P * E^{-1} * I(r,g,b,z)$$

5. Results

5.1 Experimental Setup

We tested the AR-HUD system using two display methods:

Method 1: Video see-through approach
Method 2: Optical see-through approach

In method 1, virtual images are superimposed on a recorded video stream of the real world captured by a handheld camera. In method 2, the real world is viewed using a window and the transparent screen is attached to the window. We performed the test using i5-760 CPU, 4G RAM, NVIDIA Geforce 450GTS VGA card. A transparent projection screen was attached to the window in case of optical see through approach and the screen was clamped in front of a traditional opaque projection screen in case of video see through approach. The transparent screen size is (135x90x145) cms. The projector used for projecting virtual images to the transparent screen is of 5,000 luminance. The voltage applied to the screen for adjusting its transparency ranges from 25V-80V with a frequency of 60HZ. Kinect sensor was placed in front of the user at a distance of 2.5m to detect the user's head movement. We tested in real construction site by viewing the site through a window in the Milmaru observation tower in Sejong city. The Test bed uses the residential apartments that are under construction in Sejong city. Figure 7. shows the experimental setup and the implementation result of the collimation of real apartment under construction and the 3D model of the finished building in the transparent screen.

5.2 Performance Evaluation

5.2.1 Accuracy of projection

As mentioned in the experimental setup, the performance evaluation of the AR-HUD was performed based on two display methods.

Method 1: Video see-through approach
Method 2: Optical see-through approach

The field test was performed with the visitors of Milmaru observation tower to view the new Sejong city which is under construction. In order to evaluate the accuracy of the projected virtual objects with the real scene, the user was asked to move in depth direction from the transparent screen backwards. We measured the accuracy of projection for different positions in space (distance from screen to user). The results are shown in the Figure 8. Best accuracy of about 98% lies at Area 1 which is 2.5m from the screen. Medium accuracy is observed at a distance between 2.5m-7m. After 7m the accuracy seems to decrease drastically.

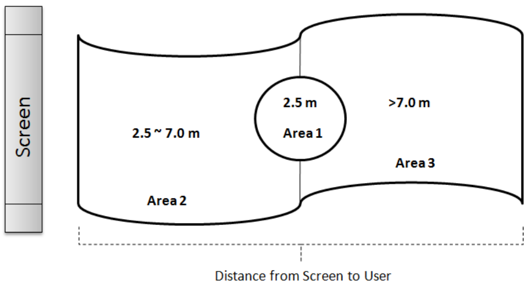


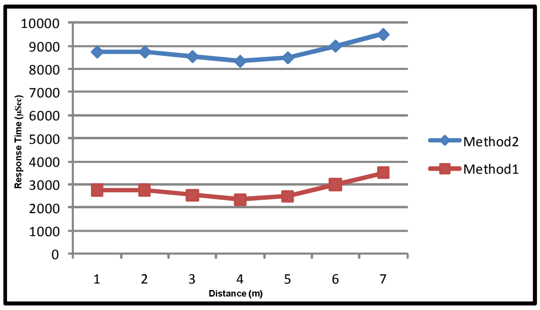
Figure 8. Accuracy of projection for three areas in space (distance from screen to user).

We also evaluated our HUD system based on the response time it takes to spatially align the virtual objects with respect to the user's head movements. Based on the two display methods described earlier, we tested our system during daylight and night. The result of the evaluation is shown in Figure 9. Method 1 showed stable results for both daylight and night. Method 2 takes more response time during daylight than night time. Method 2 takes more response time than Method 1 because of the built-in mismatch in the optical focal planes of the display and real background scene. This can also be the impact of display resolution and complexity of the overlaid graphics.

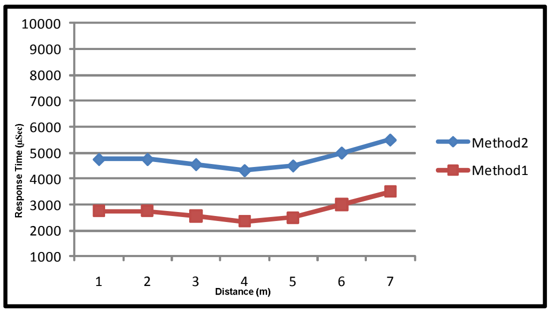
Another observation about the HUD is that better results are achieved in a dimmed room because Microsoft kinect is unable to detect the IR pattern on very bright environment. Also, it is hard for the transparent projection screen to compete with the ambient light during daytime even with the strongest projectors.

5.2.2 Head tracking accuracy

The performance of the head tracking was analyzed using the tape experiment placing it orthogonal to the field of view of Kinect. Measurements were taken along the piece of tape from 1m-7m from the sensor, standing in the centre of the Kinect's field of view, from 1m all the way back to 7m. For each measurement, extract the head position's x, y, and z values and compared it with the actual distance (ground truth) from the kinect sensor. The results are shown in Figure 10. The graphs show that Kinect's depth measurement within the region of interest were all very similar, users as close as 20cm can be imaged by this sensor, with relative error in the depth (Z) axis of approximately 5mm. However for our application, noise and accuracy loss was minimal. Also the findings show that Kinect's depth measurement as a function of distance is not linear. It appears to follow a logarithmic scaling. Therefore as the distance increases, less accurate information is obtained. To improve the performance of user tracking, depth information of the user can be obtained from multiple kinect sensors and the relative depth of user from each sensor could be considered to reduce noise in the measurements.



(a)

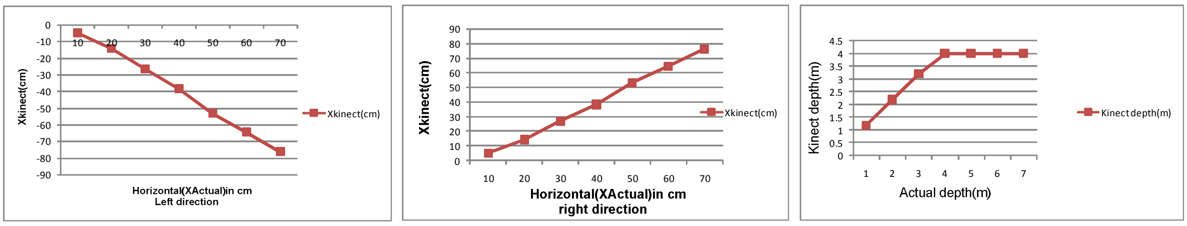


(b)

Figure 9. HUD response time vs distance (a) During daylight (b) During night

6. Conclusion

In conventional 3D displays, the user must wear glasses or special markers or polarizing filters to keep track of the user's movements. Our system overcomes these inconveniences and also provides wide field of view and possibly high-resolution images of virtual objects directly integrated with the background scene. The



(a) (b) (c)

Figure 10. Evaluation of the performance of head tracking with respect to (a) Actual horizontal distance in the left direction (b) Actual horizontal distance in the right direction (c) Actual distance from the sensor

experimental result and performance evaluation with user testing shows stable results for the system. This system can be used for a wide variety of applications like construction sector, virtual learning, tourism and so on. In medical field, medical students could use this system to visualize and discuss virtual information of the patient's body while conducting complicated surgeries

Limitation of this system is that at present it can be used only by one user at a time. As future work, we would like to use this system for multiple users by modifying the head tracking algorithm to use multiple polarization angles or color spectrum separation for multiple users.

ACKNOWLEDGEMENT

This research was partially supported by grant from Cutting-edge Urban Development Project (06KLSGB01) funded by Ministry of Construction and Transportation, Korea. A part of this research was supported by Ministry of Knowledge and Economy, Korea, under CITRC support program (NIPA-2012-H0401-12-1005) supervised by NIPA.

References

[1] J. Schmidt-Ott, S. R. Ellis, J. Krozel, R. Reisman, and J. Gips, "Augmented Reality in a Simulated Tower Environment: Effect of Field of View on Aircraft Detection", NASA TM-2002-211853, October 2002.

[2] Bourgois, M., "Augmented and Virtual Reality Research for Tower Control at Airport", In Proceeding of the NASA ICNS Conference and Workshop, Baltimore, MD, May, 2006.

[3] Caudell, T.P., Mizell, D.W., "Augmented reality: An Application of Head-up Display Technology to Manual Manufacturing Processes", In Proceedings of Hawaii International Conference on System Sciences, pages 659 - 669, USA, January 1992.

[4] L.Peterson, L.Fletcher and A.Zelinsky, "A framework for driver-in-the-loop driver assistance systems", in Proc. IEEE Intell. Transp. Syst., Sept.2005,pp. 771-776.

[5]Y.C Liu et al., "Comparison of head-up-display(HUD) vs head-down-display(HDD): Driving performance of commercial vehicle operators in Taiwan", Int. J Human-Comput. Studies, vol. 61, no.5, pp.679-697, Nov.2004.

[6] Anup Doshi, Shinko Yuanhsien Cheng and Mohan Manubhai Trivedi. "A novel active heads-up display for driver assistance", IEEE Transactions On Systems, Man, And Cybernetics-Part B: Cybernetics. Vol.39, No.1, February2009.

[7] Wen Zhan Song, Shirazi, B, Kedar, S, Chien, S, Webb, F, Tran, D, Davis, A, Pieri, D, LaHusen, R, Pallister, J, Dzusin, D, Moran, S, Lisowski, M, "Optimized Autonomous Space In-Situ Sensor-Web for Volcano Monitoring", Aerospace Conference, 2008 IEEE, DOI = 10.1109/AERO.2008.4526457

[8] Myoungjin Kim, Hanku Lee, Dongkeun Lee, Wonsa Lee, "A Distributed Real-Time System for u-GIS Informative Construction", Korea Computer Congress 2008, Vol.35, No.2(B)

[9] Kato, H. Billinghurst, M, "Marker tracking and HMD calibration for a video-based augmented reality conferencing system", Augmented Reality, 1999.(IWAR '99) Proceedings. 2nd IEEE and ACM International Workshop, October, 1999

[10] Georg Klein and David Murray, "Parallel tracking and mapping for small AR workspace", Mixed and Augmented Reality, 2007, ISMAR 2007, 6th IEEE and ACM International Symposium, November 2007

[11] http://www.microsoft.com/en-us/kinectforwindows/

[12] O. Bimber, "Spatial Augmented Reality: Merging Real and Virtual Worlds – Introduction to Current Approaches", Tutorial at the International Conference on Computer Graphics and Interactive Techniques (SIGGRAPH), 2005.

[13] D. F. McAllister, "Display Technology: Stereo & 3D Display Technologies", Wiley Encyclopedia on Imaging, Jan volume, 1327- 1344, 2002.

[14] D. Comaniciu and P. Meer, "Mean Shift: A robust approach toward feature space analysis", IEEE Trans. On Pattern Analysis and Machine Intelligence, vol.24, No. 5, 2002.

[15] F. Aherne, N. Thacker, and P. Rockett, 'The Bhattacharyya metric as an absolute similarity measure for frequentcy coded data", Kybernetica, 32(4):1-7, 1997.