Adaptive Context Aware Attentive Interaction in Large Tiled Display

Chee-Onn Wong, Dongwuk Kyoung, and Keechul Jung

Human-Centered Interface (HCI) Laboratory, Department of Media, Soongsil University, Seoul, South Korea
{cheeonn, kiki227, kcjung}@ssu.ac.kr

Abstract. We propose a conceptual model towards a context-based attentive interaction. Our focus is to improve the interaction between the user and the application of large tiled display by introducing user context and user attention. Interaction in our proposed system adapts user's visual attention region with the user's changing context based on head movement to achieve an immersive interaction with the tile display. Our experiment uses computer vision to track the user's presence and projects the most attentive regions in a tiled display in high resolution. User will be able to see other regions in higher resolution according to the head movement. At the same time, user attention is captured and modeled to learn the attentive regions to be displayed for other users. This paper will show the experimental result of the effectiveness of the perceptual interaction in a large tiled display environment.

Keywords: Tiled display, attentive user interface, context aware, perceptual interaction.

1 Introduction

The development of information and digital media makes it possible to explore into an immersive interaction experience. Perceptual interaction incorporates the perception based on the context environment that the system is situated, and enabling a desired interaction to take place accordingly. According to [1], once a computer has the perceptual ability to know who, what, when, where, and why, then probabilistic rules derived by statistical learning methods are normally sufficient for the computer to determine a good course of action.

Knowing the user's context and attention will ease the interaction. Allowing user to navigate freely in the system with pre-identified preferences and cue will help the user to get the selected information and guide the user to the user's preferred tasks. For reality and naturalness of a visual interface, it is necessary to provide immersion feeling and natural, efficient interaction with displayed objects [2].

An optimal system is a system that is smart to know what the user's current context and at the same time, prioritize the information according to the user's task. This improves the user's interaction and thus gradually making the interface invisible and focuses on user's task and solves the problem of information overload.

C. Stephanidis (Ed.): Universal Access in HCI, Part II, HCII 2007, LNCS 4555, pp. 1016–1025, 2007.

Tiled display offers greater number of pixels and has the potential to allow more applications with higher resolution display to operate. Image visualization on large screen increases the perceptual load of the user [3]. Our motivation for this paper is to provide attentive regions to be displayed in high resolution that will direct the user's focus to the specific object in a large tiled display. Our goal is also to provide a focused attentive display region for users by a bi-resolution to be projected for users.

The rest of this paper is organized as follow. Section 2 provides some related research and the achievement. Section 3 presents our proposed system overview, the framework involved and the technology used. Section 4 describes detailed implementation followed by Section 5 outlines the experimental result and finally in section 6 we present our conclusion and future work.

2 Related Research

Research on tiled display and its interaction has shown an inclination towards two directions. One main direction is on large display in virtual environment and another direction is interaction in a larger working space environment. Large display has an advantage of projecting high resolution display and interaction within the environment creates an immersive feel for the user.

2.1 System with Large Display in Virtual Environment

Research work had been carried out extensively on large display in virtual environment. Goo et al. in [4] developed an interactive virtual learning environment to evaluate the effects of guided or unguided style virtual reality (VR) learning on user attention. Large display with learning content of underwater world displayed to user while eye tracking system and various physiological sensors are used to detect user's attention state.

Li et al. presented a low cost off-the-shelf display using two projectors and render real-time 3D geometric alignment with hardware acceleration for each graphics pipeline [5]. Tao et al. in [6] demonstrated an isolation of display size and resolution as independent variables by using three different display technologies. They further show how users working with large display become less reliant on wayfinding aids to acquire spatial knowledge and construct cognitive map of virtual environments.

2.2 Tiled Display and Interaction

Human Computer Interaction (HCI) is becoming one of the most important research issues in the field of computing. It has received much attention as new paradigms of interaction evolve in different kinds of applications. Tiled display interaction as proposed by [10] utilizes avatar and gesture in a large projected display and user appears on screen to interact with icons and other items. User behavior in tiled display is observed with nine monitors to form a tiled display in [11]. User's interaction is observed for different given tasks in the regions of the tiled display. Other interaction within the large display using real time computer vision algorithm that track all ten fingers of user's hand for direct manipulation experience using finger manipulations and gestures [12].

2.3 Perceptual Interface and Attentive Interaction

Perceptual User Interfaces (PUIs) use alternate sensing modalities to replace or complement traditional mouse, keyboard, trackball and joystick input: specialized devices are exploited to provide the user with alternate interaction channels, such as speech, hand gesture, etc. [13]. With advances in computer vision, a lot of research effort has been focused on designing vision-based PUI, which are the systems that use a video camera to detect the visual cues of the user, such as motion of the face, to control a program [14, 15].

Lee et.al in [16] develop the attentive interaction design toolkit, a vision-based input toolkit that gives users an analysis of faces found in a given image stream, including facial expression, body motion and attentive activities. The attention meter will measure the level of attention by users towards the system.

3 System Overview

Our proposed system intended to provide an immersive perceptual interaction that augments user's interest based on the intention and attention that the user is currently situated, the related content of interest and hassle-free perceptual interaction. The system is based on the conceptual model of context aware attentive interaction as shown in Fig. 1.

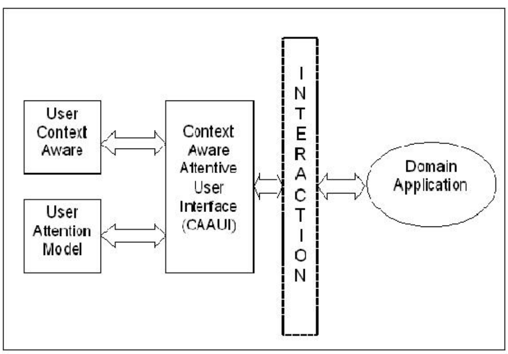


Fig. 1. Conceptual model of context aware attentive interaction

The user context aware recognizes input that is available to the context application. The information will be diverted to the application directly through the Context Aware Attentive User Interface (CAAUI). The attention model deals with user adaptation, learning, reasoning and planning for selected data to be displayed based on user's past, present and expected preference. An intermediate context aware attentive user interface (CAAUI) will act as an intelligent perceptual interface between the user and the application and optimally distribute tasks between user and system.

3.1 Interaction in Tiled Display

Tiled display system has an advantage of larger display with high resolution. Our proposed model is simulated in a large tiled display application prototype developed in our laboratory. We use two projectors to display a large display and render a low resolution data for maximizing the computational efficiency and minimizing user's perceptual load. User context aware detects the presence of users and head motion tracking of the user using a AR marker. Upon detection of the user's presence, image in the tiled display will be displaying the essential attentive regions in high resolution based on the outcome of the learning model from user's viewing behavior in the past (user attentive model). Other regions remain in low resolution.

Tiled display provides an environment for perceptual interaction. The system configuration for the tiled display uses two projectors whereby each capable to project high and low resolution. This is to minimize the computational power by only showing the attentive areas with high resolution to users. The display is divided into eight equal sections.

User will interact with the display directly with the head orientation detection with a marker. User wears a cap with AR marker and a camera detects the user's head movement to display the image cluster on high resolution. The remaining clusters will be projected in low resolution. The rationale of our implementation is we want to direct the user's attention to the focused area while minimizing the user's perceptual load. At the same time, computational power will be minimized in terms of displaying low resolution in unattended areas.

Our system in geometric calibration solves non-linear distortion using the method of dividing a small rectangular area. It changes non-linear distortion into linear distortion. It computes warping functions for each small rectangular. Warping function is determined using keystone correction method and four corner points. Fig. 2 shows the relationship between the frames of reference that are relevant to the problem of automatic keystone correction. The relationship between the three frames of reference corresponding to the source image frame, camera image frame and projected image frame is shown in Fig. 2(a). Fig. 2(b) shows the application image can be appropriately pre-warped, using the mapping W so that it appears rectilinear after projection through a misaligned projector.

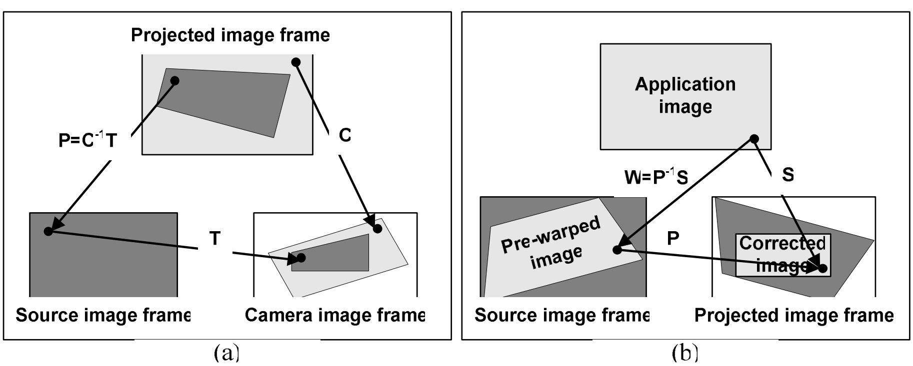


Fig. 2. Relationship between the frames

3.2 Head Direction Orientation Using AR Marker

Our prototype deploys using AR ToolKit for head direction orientation during perceptual selection by the user. User wears a normal cap with AR Marker following the user's head direction when user views the large image in the display.

AR ToolKit [17] identifies the position of the marker in the camera coordinate system. The marker following the user's head position is computed and tracked. Fig. 3 shows the coordinate system used in AR ToolKit [17]. We use this concept to determine the head orientation and position using this coordinate system [18,19,20].

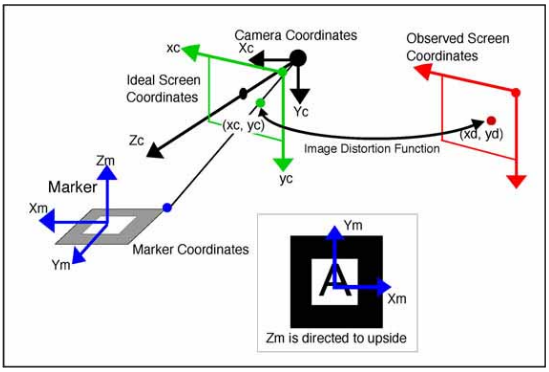


Fig. 3. Coordinate system used in AR ToolKit

4 System Description

The first component of the prototype consists of detecting the user's presence. We use a web camera as an initial set up to detect the user's present and projects the attentive regions with high resolution to capture user's attention. This is based on the second component; the attentive model that is built with an unsupervised learning method. The user's preference of the attentive region in the display will be projected as high resolution so that user's perceptual load is minimized. The attentive model uses time and frequency of re-visiting the same region to determine the most attentive regions for that particular image. The data will be used as a predictive input for user modeling in the user attention model for displaying as a default high resolution attentive region in future.

The attention model learns the attentive viewing regions for user modeling and adaptation. Learning method based on accumulative time spent and frequency of revisiting the attentive regions across a sample of users. The display is divided into eight sections and computation of the time spent in each focused section together with the frequency of revisit will be stored in the database. Our system will obtain the comparison result of the data and based on these parameters, it will determine the attentive region for that particular display.

Rendering of images onto the tiled display uses two computers with two different projectors which projects both high and low resolution areas that are depending on the region of focus. User's selections of region to be displayed as a focused region (high resolution) rely on the marker detection of the user's head direction orientation. The description is depicted in Fig. 4.

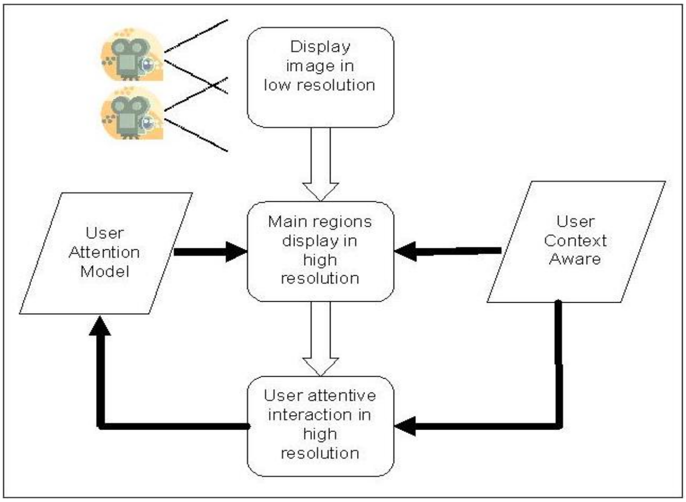
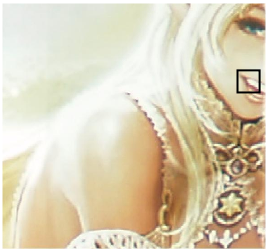


Fig. 4. Overview of the interaction in tiled display

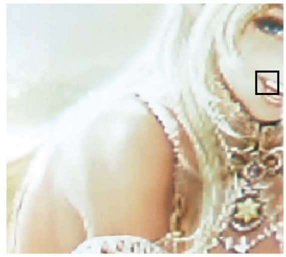
System configuration for the prototype utilizes simple off the shelf equipments. The screen size is 3m X 2m (width X height in meter). Two projectors and one web camera are used to set up the environment. AR marker is used to track the orientation direction of the user's head. The tracking is built with AR Toolkit.

5 Experimental Result

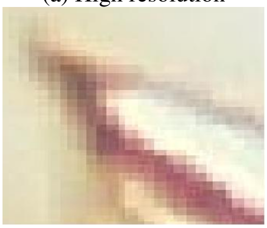
Our experiment shows the tiled display operated by the user's presence detection and by default showing the attentive region when the user is near to the display. User will be detected by a camera and display of high and low resolution takes place. Fig. 5 shows a comparison of an image projected by (a) high resolution, (b) low resolution, (c) inset from high resolution (zoom) and (d) inset from low resolution (zoom). The prototype will display a high resolution as in 5(a) when the region is selected. Otherwise, it will display a low resolution as in 5(b). The high and low resolution display has a clear distinct difference with the quality of projection in the overall display of an image and thus contributes to the minimizing of the computation power as a whole. Our tiled display projects a total of 8 sections (4 sections for each projector) and display accordingly when user's selection is detected via the AR marker. Low resolution is four times lower pixel than the high resolution.



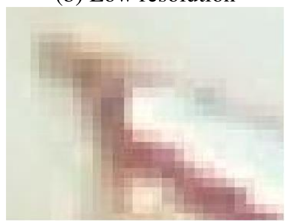
(a) High resolution



(b) Low resolution



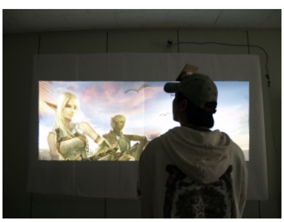
(c) Inset from high resolution (zoom)



(d) Inset from low resolution (zoom)

Fig. 5. (a) High resolution display, (b) low resolution display, (c) inset from high resolution (zoom) and (d) inset from low resolution (zoom)

Fig. 6 shows user perceptual selection in the tiled display. We use two projectors projecting high and low resolution and user's head orientation will be captured according to the direction that the user prefers. This is also to minimize the perceptual load of the user when the user is focusing on a small part of a large display. This concept is based on the attentive display [17] that refers to the demand for rendering power and display resolution, but it is different from parallel display approach. Instead of requiring additional hardware, attentive display make more out of available hardware by directing display and computation resources to where they count most. AR marker is used to trace the head orientation.



(a) Upper left direction

(b) Lower right direction

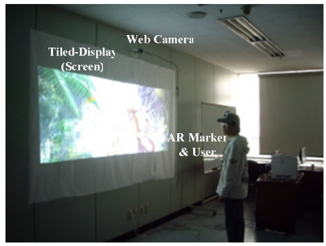
(c) AR marker

Fig. 6. (a) and (b) User direction orientation with high resolution display on selected area and (c) AR marker on user's head for direction orientation

An attentive model based on calculation of time and revisit of each user will determine the most attentive region in that particular display. The attentive model acts as an intelligent component to pre-define the attentive area for each display based on the accumulative user's data. It is modeled with a learning capability that learns the attentive region based on viewing parameters.

The overall system configuration of our prototype is shown in Fig. 7. Our experimental result shows a reduction of processing time and minimizes the computational power by projecting images to large display with combination of high and low resolution. We use the method of focusing only the essential based on the attention from the user and thus reduce the overall computational power. We averaged out the iterations of testing and the computation time is summarized in Table 1.

(a) Hardware setup



(b) User interaction

Fig. 7. Overall system configuration

Table 1. Comparison of computational power and accuracy of head orientation

Number of regions displayed (per projector)	Computational time (seconds)
0	0.2105
1	0.2890
2	0.3435
3	0.4220
4	0.4925

6 Conclusion and Future Work

We proposed a perceptual interaction in a tiled display setting. Our prototype follows the conceptual model of context aware attentive interaction. User interacts with the tiled display based on their head orientation with the direction of user's view. This will lighten the user's perceptual load and minimize the computational power. At the same time, user's interaction data will be constantly feed as an input to the attention model in order to display the most attentive region in a particular display to the new users.

There are several advantages in our proposed model. By redirecting the user's attention to their focus of interest, users' experiencing an immersive perceptual interaction with lighter perceptual burden in a large tiled display. Context aware that reasons the user's cues together with attention model that prioritize and learn user's behavior will decrease user's perceptual load thus interacting directly with the desired attentive information. Minimal computational resources needed to render high resolutions only on selected attentive regions.

Future work will be the possibility to include other context aware input and improving calibration of visual attention based on both head tracking and eye tracking without using a marker. Real time tracking of the human head movement with direction orientation detection will enhance the perceptual and immersive feel when interacting in a tiled display.

Acknowledgments. This work was supported by grant No.(R01-2006-000-11214-0) from the Basic Research Program of the Korea Science & Engineering Foundation.

References

1. Pentland, A.: Perceptual Intelligence. Communications of the ACM, vol. 43(3) (2000)
2. Ahn, S.C., Kim, I.J., Kim, H.G.: Interactive Immersive Display. The 14th International Conference on Artificial Reality and Telexistence (ICAT, TSI workshop (2004) pp. 63–68 (2004)
3. Baudisch, P., DeCarlo, D., Duchowski, A.T., Geisler, W.S.: Focusing on the essential: Considering attention in dislay design, Comunications of the ACM, vol. 46(3) (2003)
4. Goo, J.J., Park, K.S., Lee, M.H., Park, J., Hahn, M.S., Ahn, H.G., Picard, R.W.: Effects of Guided and Unguided Style Learning on User Attention in a Virtual Environment. International Conference on E-Learning and Games (Edutainment 2006) pp. 1208–1222 (2006)
5. Li, Z.Y., Varshney, A.: A Real-Time Seamless Tiled Display System for 3D Graphics. In: Proceedings of Seventh Annual Symposium on Immersive Projection Technology (IPT (2002)
6. Tao, Ni., Bowman, D.A., Jian, Chen.: Displays : Increased display size and resolution improve task performance in information-rich virtual environment. In: Proceedings of the conference on graphics interface (GI'06) (2006)
7. Jo, D., Kang, H., Lee, G.A., Son, W.: Xphere: A PC Cluster based Hemispherical Display System. IEEE Virtual Reality Conference VR, 312 (2006)
8. Vertegaal, R., Shell, J.S., Chen, D., Mamuji, A.: Designing for Augmented Attention: Towards a Framework for Attentive User Interfaces. Computers in Human Behavior 22, 771–789 (2006)
9. Tahboub, K.A.: Intelligent Human-Machine Interaction Based on Dynamic Bayesian Networks Probabilistic Intention Recognition. Journal of Intelligent and Robotic Systems 45, 31–52 (2006)
10. Ahn, S.C., Lee, T.S., Kim, I.J., Kwon, Y.M., Kim, H.G.: Large display interaction using video avatar and hand gesture recognition. In: Campilho, A., Kamel, M. (eds.) ICIAR 2004. LNCS, vol. 3211, pp. 261–268. Springer, Heidelberg (2004)
11. Ball, R., North, C.: Analysis of user behavior on high-resolution tiled displays. In: Costabile, M.F., Paternó, F. (eds.) INTERACT 2005. LNCS, vol. 3585, pp. 350–363. Springer, Heidelberg (2005)
12. Malik, S., Ranjan, A., Balakrishnan, R.: Interacting with large displays from a distance with vision-tracked multi-finger gestural input. The 7th Annual ACM Symposium on User Interface Software and Technology (UIST05) (2005)

13. Iannizzotto, G., Rosa, F.L., Costanzo, C., Lanzafame, P.: A multimodal perceptual user interface for collaborative environment. In: Roli, F., Vitulano, S. (eds.) ICIAP 2005. LNCS, vol. 3617, pp. 115–122. Springer, Heidelberg (2005)
14. Gorodnichy, D.O., Roth, G.: Nouse ‘Use your nose as a mouse’ perceptual vision technology for hands-free games and interfaces. Image and Video Computing NRC 47140 12(12), 931–942 (2004)
15. Gorodnichy, D.O.: Perceptual cursor – a solution to the broken loop problem in vision-based hands-free computer control devices. NRC/ERB 1133, NRC48472 (2006)
16. Lee, C.H.J., Wetzel, J., Selker, T.: Enhancing interface design using attentive interaction design toolkit. International conference on computer graphics and interactive techniques (ACM SIGGRAPH, Educators program (2006), http://www.hitl.washington.edu/artoolkit/
17. Kato, H., Billinghurst, M.: Marker tracking and HMD calibration for a video-based augmented reality conferencing system. In: Proceedings of the 2nd International workshop on Augmented Reality (IWAR99) (1999)
18. Billinghurst, M., Kato, H., POupyrev, I.: The MagicBook: A transitional AR Interface. Computers and Graphics, pp. 745–753 (2001)
19. Billinghurst, M., Kato, H.: Collaborative augmented reality. Communications of the ACM 45(7), 64–70 (2002)