

# Kinect© as Interaction Device with a Tiled Display

Amilcar Meneses Viveros<sup>1</sup> and Erika Hernández Rubio<sup>2</sup>

<sup>1</sup> Departamento de Computación, CINVESTAV-IPN, México D.F.

`ameneses@cs.cinvestav.mx`

<sup>2</sup> SEPI-ESCOM, Instituto Politécnico Nacional, México D.F.

`ehernandezru@ipn.mx`

**Abstract.** The use of high resolution tiled display has become popular in the scientific community. User interaction with these devices depends on the hardware configuration and the software in use. The variety of hardware configurations and software generates various types of execution modes and interaction in the tiled display, this diversity has resulted in not having a standard for human computer interaction. This paper shows the results of the interaction between users and the tiled display using the Kinect©. The results help us find improvements in hardware configurations of this arrays of displays, applications design and try to find standards in defining user-defined motion gestures.

## 1 Introduction

The use of high resolution tiled display has become popular in the scientific community [TS][SC][MC] [NH][NV][CC][ET]. They are used for collaborative work and for the deployment of information and navigation of large data volumes [SC][MC][NV][BG]. Current research on human computer interaction in the tiled display is focused on two main points: the perception of information displayed [ET][YN][BN1] [BN2][TGSP] and control applications [BN2][AT][AG][BG][NJ]. User interaction with these devices depends on the configuration of hardware for controlling video wall (visualization server or visualization cluster) and the applications running on it [TS][NH][CC][BN1] [RCMM]. Applications running on these devices can be of two types: distributed applications or desktop applications [TS][MH][NJ].

Distributed applications are developed to take advantage of the hardware features of the visualization cluster [PR][HH][GdS]. This applications are highly scalable. When the application is running on the cluster, it is very common to use the master node (or front-end node) as responsible for the interaction with the application running on the tiled display. These applications are for very specific purpose and user interaction with the device is made from the master node. In this case, the user interaction with the tiled display is limited to a set of basic operations that fulfill the functionality of the application [TS][NH][CC][AG][R][BL][NJ].

Desktop applications take advantage of the operating system's ability to export graphical interface to tiled display either the case of a visualization cluster or visualization server [TS][NJ]. The scalability of desktop applications depend on the capabilities of operating system and graphics cards. When wing applications running in an extended desktop, the user can work through the master node (for example, in case of using the VNC protocol) or directly on the arrays of displays (for example, when using XDMX). User interaction is with the window manager or the graphical user interface of the operating system. The user interacts with the window manager using mouse, keyboard or combination of both.

The variety of hardware configurations and software generates various types of execution modes and interaction in the tiled display, this diversity has resulted in not having a standard for human computer interaction[YN]. On the one hand the problem is show de information for the users and the other hand the application control. Furthermore, when adding new specific use applications, they are restricted to only user interaction that is in the master node [TS][CC][HH][NJ]. User interaction is limited to the capabilities of the master node. When applications working at a desktop, some interactions mechanism can have a critical impact, for example, mouse manipulation, can cause functionality problems if not cared Fitt's law [BN1][BN2]. It is possible to take advantage of MOCAP technology, specifically the Kinect©, to provide the user with an intuitive means that permits them to interact with tiled display applications. This approach allows user mobility along the tiled display freeing interaction with the master node and to define a set of gestures to control applications running on the tiled display.

In this paper, we present the results and proposals to use the Kinect©as interaction device with a tiled display. This tiled display is controlled by a cluster of Apple Mac Mini [MC]. Interaction experiments reported in this paper are related to the control of the tiled display desktop through VNC protocol and several special-purpose distributed applications, such as an image viewer, a web browser and ParaView. The applications have been selected to cover the widest possible use case. For interaction with applications, we tested a set of gestures (proposed in other works and added some) [AG][KB], that through Kinect©, are interpreted to perform specific actions on the tiled display or in the desktop of the master node. Tests were conducted to emulate the behavior of the mouse and keyboard. In particular, for the case of interaction with widgets that respond to the keyboard events, such as text fields,we use the technique of virtual keyboards (common in environments of smartphones with touch screens), and we did a comparative metrics based on usability, functionality and effectiveness between the keyboard virtual (stylus based) and 8pen Android keyboard.

## 2 MOCAP Technologies

The motion capture, motion tracking or MOCAP are terms used to describe the recording process and translating of that motion into a digital model. This

Technology has been used for capture and analysis of the human movement. Control applications where estimated motions can control something benefit about these technologies as interfaces with gaming, virtual reality and human computer interfaces. There are many MOCAP technologies, the main categories of these are using markers or without the markers to recognize parts of the human body. The kinect is a technology that does not use markers for this purpose.

## 2.1 Kinect©

The Kinect© is a free gaming device developed by Microsoft© for the XBOX 360 and PC in the future through Windows 8. Kinect© enables users to control and interact with the console without having physical contact with a traditional video game control, by a natural user interface that recognizes gestures and voice commands.

The main objective of this device is increase the use of the Xbox 360. In our case we use this device to allow a user to control a tiled display, using only the movements and efforts of his body. Kinect© incorporates different technologies as an emissary of infrared that emits an invisible light which together with a CMOS sensor looks how this emission is reflected back and passes de data to the console or in this case to the PC in a grayscale format so it can determinate the depth of the scene and the motion of the user.

The algorithm which operates under the Kinect© was developed by Microsoft Research Centre in Cambridge. The algorithm takes as an input a depth image which is analyzed pixel by pixel, where each pixel is assessed according to their characteristics such as its depth, if the pixel belongs to an upper or lower part of the body; the result of each characteristic about pixel is combined with the research in a classifier called forest of decision thus a collection of decision trees where each decision tree was trained or specialized in a set of characteristics of an image depth which different parts of the body were previously labeled. The classifier assigns certain probability to the pixel belongs to a particular body part, then the algorithm assigns the pixel belonging to a body part that have obtained most probability in the classifier. Finally which each pixel belonging into a particular body part, hinge points are assigned to the identified areas in the body in three views, frontal, side and upper[MRCC]. This procedure is made 3 times per second; this is analyzed 200 pictures or images per second, which is about 10 times faster than other techniques for recognition and motion detection.

## 3 Experiment

We are interested in to know the interaction of the user with the tiled display through the Kinect©. HCI metrics that we are using are usability, performance, effectiveness and comfort. For this purpose we have developed several questionnaires that apply to users after an action or task running in the tiled display. We applied twelve questionnaires. Which includes actions such as moving the

- 
1. How easy it was to control the mouse pointer through the Kinect?  
a) Very Diffcult    b) Difficult    c) Easy    d) Very Easy
  2. How fast did you control the mouse pointer through the Kinect?  
a) Vey Slow    b) Slow    c) Fast    d) Very Fast
  3. How comfortable did you control the mouse pointer through the Kinect?  
a) Very Confortable    b) Confortable    c) Uncomfortable    d) Very Uncomfortable
  4. How fast was the response of the mouse pointer using the Kinect?  
a) Vey Slow    b) Slow    c) Fast    d) Very Fast
  5. Did it seem appropriate at the time it took to answer the pointer?  
a) Yes    b) No
  6. Controlling mouse pointer through the Kinect, was successful?  
a) Yes    b) No
  7. The mouse pointer control, through the Kinect, is performed as you thought?  
a) Yes    b) No
- 

**Fig. 1.** Example of questionnaires used in this experiment

cursor, select a character from virtual keyboard until tasks such as opening an application, close a window and delete text, to name a few. An example of this type of questionnaire is shown in the Figure 2.1.

We applied two types of tests. The first is to study the interaction of the user with the desktop and specific applications. The second is to compare the use of two types of virtual keyboards: Stylus based and Android 8pen. The tests were performed with two different groups of students: We use a sample of twenty people for the first test and twenty one for virtual keyboard tests. Ten questionnaires were used for the first test and two questionnaires for the second test.

### 3.1 User-Defined Motion Gestures for Tiled Display Interaction

User-defined motion gestures are associated to actions and task. A task is a sequence of actions. In general, the actions that we consider in this work are move de the mouse, push the button and select a character.

The gestures considered for the first test (use of applications and desktop) are:

- Move the mouse pointer** The gesture associated with this action, is to point the palm towards the Kinect ©, each movement (up, down, left or right) will have the effect to control the mouse pointer.
- Click** The gesture associated with this action, is to point the palm towards the Kinect©, but making a move forward and return to the original position.
- Item selection** This is done using the click action. Is to point the palm towards the Kinect ©, but making a move forward and return to the original position. Allows point and select objects on the screen. The system applies to a function or process these objects.
- Open Application** This gesture is similar to the Item Selection task. Is to point the palm towards the Kinect©, but making a move forward and return to the original position. Allows point and select objects on the screen. The system applies to a function or process these objects.
- Drag Item** It's like the click action, except that the hand stays on and does not return to its original position. Allows you to move (drag) an object on the screen or rotate it.
- Open Context Menu** The gesture associated with this task is to carry out click movement, with the helping hand (one that does not control the mouse pointer).
- Rotate Object** It's like the click action, except that the hand stays on and does not return to its original position. Allows you to move (drag) an object on the screen or rotate it.
- Slide Up/Down** This gesture is to put the palm facing the Kinect© and make a move down quickly.
- Slide Left/Right** Involves placing the palm facing the Kinect© and make a move to the left or right, quickly.
- Zoom** This gesture is to put the palm facing the Kinect© and make a move down quickly.
- Close Window** This gesture is a change of direction of the palm, on the X axis, four times.



**Fig. 2.** Cinveswall and Kinect©

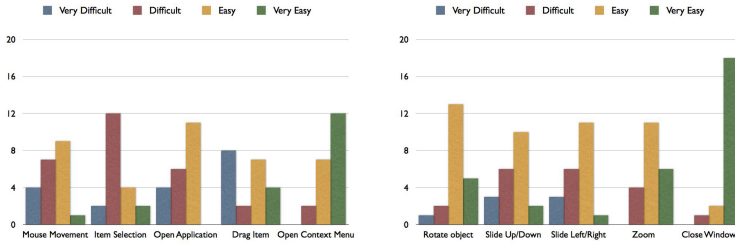


Fig. 3. Usability charts

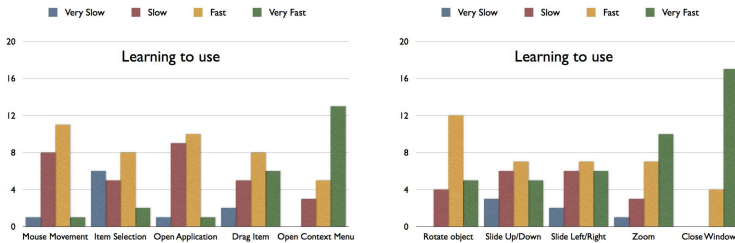


Fig. 4. Charts of user training

### 3.2 Tools and Equipment

The gestures considered for the second test (use of applications and desktop) are:

**Select character in the virtual keyboard** It involves moving the mouse pointer to the button of the character and click.

**Write text** This gesture includes selecting a text field, and then select the character set of the text.

**Write number** This gesture includes selecting a text field, and then select the numeric character set of the number.

**Delete text** Position the pointer to the left of text character to delete and press the delete key.

**CinvesWall.** The CinvesWall is an array of 12 Apple Cinema Display 24-inch, in a 3x4 configuration, controlled by a cluster of 12 Mac mini and one MacPro server as the master of visualization cluster, figure 2. The network visualization cluster interconnect is Gigabit Ethernet. The total resolution of the display device is 27 megapixels (7680x3600). The nodes in the cluster work with OSX version 10.6.5 (Snow Leopard). The administration of the nodes is done with the application Apple Remote Desktop.

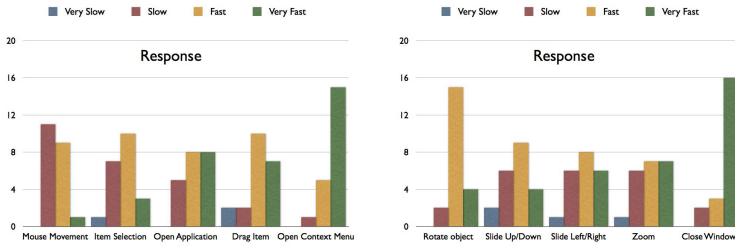


Fig. 5. Performance charts

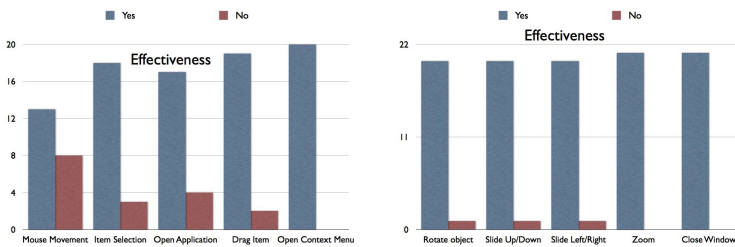


Fig. 6. Effectiveness charts

**Kinect©Setup.** The Kinect© control the user gestures. The Kinect is connected to the master node, but is placed in the middle of the video wall to give the illusion that the user interacts directly with the high-resolution display. In order for users to interact with the video wall, standing five feet away from the Kinect©, as show in Figure 2.

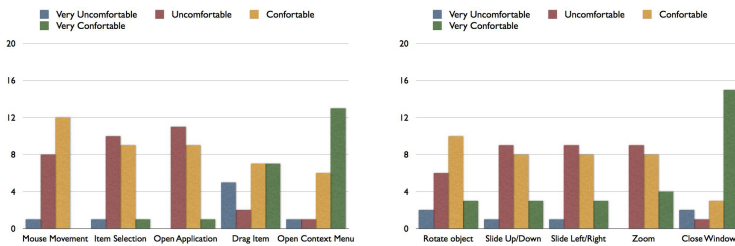


Fig. 7. Charts of user comfort

**Applications Settings.** Tests of actions and tasks, with the Kinect<sup>®</sup>, were performed using distributed applications running on tiled display and applications running on the distributed desktop. Fortunately we were able to use an application that runs in both formats: ParaView. We also use PreView application running on the OSX desktop and equivalent distributed image display application. Several tasks are able to test in the context of execution of an application. The rest of the tasks and actions were used to manage the distributed desktop.

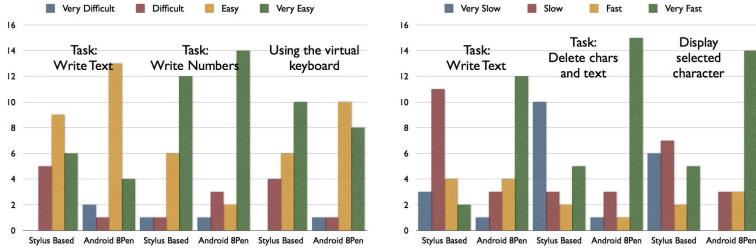


Fig. 8. Usability and performance charts of use the virtual keyboard

## 4 Results

Since we want to measure the use of Kinect<sup>®</sup> as interaction device with video wall, we intend measure usability, effectiveness, performance and comfort. Started showing the results of applying the tests to the tasks and actions. And then see the results of tests using virtual keyboards.

### 4.1 Results for Test on Actions and Tasks

The chart in the Figures 3 and 4 displays the usability and user training results. This training is related to learn and associate a gesture to an action or task.

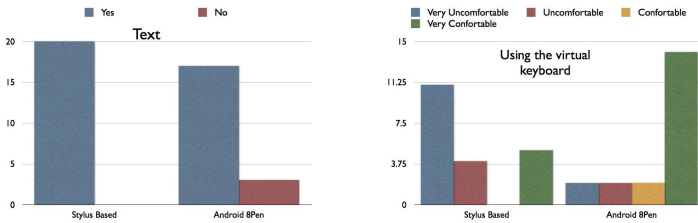


Fig. 9. Effectiveness chart and chart of user comfort using the virtual keyboard



We can measure the performance of the Kinect© using the time of respond to action or task. The Figure 5 show the user's point of view about the response. The effectiveness of actions and task performed by the users throught Kinect© is showed in Figure 6. Figure 5 show the user's point of view about the response. The ease of use of gesture for Kinect, shown in Figure 7.

## 4.2 Test Results of the Use of Virtual Keyboards

The Figure 8 show the charts of usability and performance of virtual keyboards: Stylus based and Android 8Pen. And figure 9 show the effectiveness chart and chart of user comfort using the virtual keyboard.

## 5 Discussion

The configuration of tiled displays based on visualization server use the homogeneous visualizations capabilities in every output of the graphics cards. This configuration take advantage that the response to user actions are very quick. In this work we are using a distributed control to the tiled display. When the main task is to visualize these devices with high-resolution images, the user might be tolerant synchronization waits and displays. However, when the main task is the interaction, the user waits for answers in a very short time. The Cinveswall use a Gigabit Ethernet switch that does not allow multicast operation. This defect causes users do not have such good results in the tasks associated with drag the mouse pointer and selection.

## 6 Conclusions

The interaction use of the tiled display need that the time of the graphics response is in the range accepted for the user. In the case of having a visualization cluster, you must ensure rapid communication between different nodes, so that the user has the idea that the response is adequate in the tiled display.

There is a relationship between the area of deployment of the video wall and the step size of the cursor on the remote desktop. The relationship is that the higher the display area, the step size should be smaller.

In testing, you can verify that the user's interaction with applications running in a distributed, as ParaView and image viewer, is similar to the interaction of the applications that run on the desktop distributed as ParaView and Preview. For the type of results obtained in this work. Best tasks associated gestures using the Kinect, are not related to the use of the mouse pointer.

Finally, the use of virtual keyboards for handling text fields seems appropriate. Surprisingly, 8pen Android appears to offer advantages over the Stylus based.

**Acknowledgements.** The authors wish to thank Arellano Cenizeros Israel, González Hernández Carlos Omar, Martínez Martínez Jesús Oswaldo and Julia Leticia Sánchez Sánchez for their valuable assistance in the implementation of the tests were performed.

## References

- [TS] Tao, N., Schmidt, G.S., Staadt, O.G., Livingston, M.A., Ball, R., May, R.: A Survey of Large High-Resolution Display Technologies, Techniques, and Applications. In: Proceedings of the IEEE Conference on Virtual Reality, vol. 236, pp. 223–236. IEEE Computer Society, Washington, DC (2006)
- [SC] Smarr, L.L., Chien, A.A., DeFanti, T., Leigh, J., Papadopoulos, P.M.: The OptIPuter. *Communications of the ACM* 14(11), 58–66 (2003)
- [MC] Viveros, A.M., Vergara, S.V.C.: The CinvesWall in More than Research. In: Proceedings of the 2nd International Supercomputing Conference in México, ISUM 2011, vol. 2, pp. 177–183. Universidad de Guadalajara, México (2011)
- [NH] Nirnimesh, H.P., Narayanan, P.J.: Garuda: A Scalable Tiled Display Wall Using Commodity PCs Visualization and Computer Graphics. *IEEE Transactions on* 13(5), 864–877 (2007)
- [NV] Naveen, K., Venkatram, V., Vaidya, C., Nicholas, S., Allan, S., Charles, Z., Gideon, G., Jason, L., Andrew, J.: SAGE: the Scalable Adaptive Graphics Environment. In: WACE (2004)
- [CC] Chen, Y., Chen, H., Clark, D.W., Liu, Z., Wallace, G., Li, K.: Software Environments for Cluster-based Display Systems. In: First IEEE/ACM International Symposium on Cluster Computing and the Grid (2001)
- [ET] Ebert, A., Thelen, S., Olech, P.-S., Meyer, J., Hagen, H.: Tiled++: An Enhanced Tiled Hi-Res Display Wall. *IEEE Transactions On Visualization and Computer Graphics* 16(1) (2010)
- [YN] Yost, B., North, C.: The perceptual scalability of visualization. *IEEE Transactions On Visualization And Computer Graphics* 12(5) (September/October 2006)
- [BN1] 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)
- [BN2] Ball, R., North, C.: Effects of tiled high-resolution display on basic visualization and navigation tasks. In: *CHI 2005 Extended Abstracts on Human Factors in Computing Systems (CHI EA 2005)*, pp. 1196–1199. ACM, New York (2005)
- [AT] Ahlborn, B.A., Thompson, D., Kreylos, O., Hamann, B., Staadt, O.G.: A practical system for laser pointer interaction on large displays. In: Proceedings of the ACM Symposium on Virtual Reality Software and Technology (VRST 2005), pp. 106–109. ACM, New York (2005)
- [AG] Israel, A.C., Omar, G.H.C., Oswaldo, M.M.J.: Control de un VideoWall via Kinect®. In: México, D.F. (ed.) *Trabajo Terminal, Escuela Superior de Cómputo, ESCOM-IPN, Junio 2012* (2012)
- [RCMM] Ramírez, L., Chapa, S., Meneses, A.: DVO: Model for Make a Handler for a Tiled Display. *Lecture Notes in Engineering and Computer Science*, vol. 2198(1), pp. 995–1001 (2012)
- [MH] Myers, B., Hudson, S.E., Pausch, R.: Past, present, and future of user interface software tools. *ACM Trans. Comput.-Hum. Interact.* 7(1), 3–28 (2000)
- [PR] Puder, A., Romer, K., Pilhofer, F.: *Distributed Systems Architecture, A Middleware Approach*. Elsevier Inc. (2006)
- [HH] Ng, Y., Humphreys, G., Houston, M.: Chromium: A stream processing framework for interactive rendering on clusters. *ACM TOG* 21(3) (2002)

- [GdS] Germans, D., van der Schaaf, T., Renambot, L., et al.: Retained mode parallel rendering for scalable tiled displays. In: IPT (2002)
- [R] Patricia, L., Rivera, R.: “Minería de datos visual sobre una pared de video”, Master Degree Thesis, Departamento de Computación, CINVESTAV-IPN (2008)
- [BL] Granados, G.A.B., García, A.L.L.: Manejo de video distribuido sobre un Video Wall, Trabajo Terminal, Escuela Superior de Cómputo del IPN, Junio de (2009)
- [TGSP] Tan, D.S., Gergle, D., Scupelli, P.G., Pausch, R.: Physically large displays improve path integration in 3D virtual navigation tasks. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI 2004 (2004)
- [BG] Birnholtz, J.P., Grossman, T., Mak, C., Balakrishnan, R.: An exploratory study of input configuration and group process in a negotiation task using a large display. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI 2007). ACM, New York (2007)
- [NJ] Nam, S., Jeong, B., Renambot, L., Johnson, A., Gaither, K., Leigh, J.: Remote visualization of large scale data for ultra-high resolution display environments. In: Ma, K.-L., Papka, M.E. (eds.) Proceedings of the 2009 Workshop on Ultra-scale Visualization (UltraVis 2009), pp. 42–44. ACM, New York (2009)
- [KB] Boulos, M.N.K., Blanchard, B.J., Walker, C., Montero, J., Tripathy, A., Gutierrez-Osuna, R.: Web GIS in practice X: a Microsoft Kinect natural user interface for Google Earth navigation. *International Journal of Health Geographics* 2011 10(45) (2011)
- [MRCC] Microsoft Research Cambridge & Xbox Incubation, Real-Time Human Pose Recognition in Parts from Single Depth Images