Design and Evaluation of a Cell Phone Pointing Interface for Interaction with Large Projector based Displays

Saso Koceski
Faculty of Computer Science
University "Goce Delcev"
Stip, Macedonia

Natasa Koceska
Faculty of Computer Science
University "Goce Delcev"
Stip, Macedonia

Ivica Kocev
Faculty of Computer Science
University "Goce Delcev"
Stip, Macedonia

ABSTRACT

The availability of mobile phones enhanced with different sensors increased the possibility of using the mobile devices in various applications in ubiquitous computing. Mobile phones can be used as interaction device with other devices such as large projected displays. In this paper we propose a design of a cell phone pointing interface based on its embedded accelerometer sensor for interaction with large projector based displays. The main idea is to use a simple hand gestures for pointing to a certain point, for object selection, for object movement or for drawing various forms on large display screen.

Rigorous and systematic evaluation of the proposed interface has been conducted and its efficiency has been experimentally evaluated on various tasks. The results have shown that embedded accelerometer sensor in mobile phones in combination with a simple hand gestures are a promising approach for interaction with large display screen and that potential users see ease of use, intuitiveness and enjoyment as advantages of this interaction technique.

General Terms

Human-computer interaction, Human-computer interface

Keywords

Cell phone, Pointing interface, Projector based display, Evaluation methodology.

1. INTRODUCTION

Technological advances in the last decade have contributed to remarkable improvements in large display technologies, in particular: increased size, higher resolution and dropping prices. It is likely that this trend will accelerate, and that people will become more familiar with their usage. Because of their characteristics they became widely popular and are used extensively nowadays for number of purposes, ranging from content-output only to interactive workgroup applications. Their physical dimensions make them visible from greater distances than conventional desktop monitors, thus making them highly appealing for displaying shared content to a group of users. Moreover, these characteristics make them ideal for co-located, synchronous multi-user interactions.

Over the last decade, the number of large displays has been growing considerably. Although the main application area of these screens is for advertising purposes, other numerous fields of possible applications are reported. They are including information, art, culture and aesthetics, entertainment, education, improving customer experience, communication, participation, medicine, industry, etc.

Today projectors are affordable, which probably makes them currently the best technology for creating very large displays that are flexible, easy to configure and deploy. This kind of displays have numerous applications in ubiquitous computing, they can enable immersive environment in virtual reality applications, may facilitate visualization of large and complex datasets by maintaining both overview and detail views simultaneously [1].

Various scientific research projects regarding projector-based displays have been conducted. They are referring various aspects of projector based displays, like: creating a foldable interactive display [2], distributed paradigm for multiple users to interact simultaneously with large tiled rear projection display walls [3], Continuous Multi-Projector Distributed Cooperative Framework [4].

Despite their popularity, in many cases the main drawback of using large projector based displays is the stationary installation of such devices and the lack of natural interaction interface.

Various techniques that enable interaction with large displays have been developed. Among the others they include user's body silhouette interaction with the on-screen virtual 3D during his movement [5], new computer interface system based on laser points pattern recognition in beam projection, which can generate five interfacing commands has been introduced in [6], intuitiveness and effectiveness of gestures for free spatial interaction with large displays has been investigated in [7], comparison between laser and manual pointing as remote input in large display interactions has been reported in [8].

The penetration of mobile phones in everyday life and their characteristic increased the possibility of using the mobile devices for interaction with large displays.

Namely, according to ITU statistics [9], cell phone use has increased to over forty-five hundred million users worldwide, and the popularity rate of mobile phones has exceeded more than 90 percent in developed countries. The ratio of converged mobile phones in the worldwide mobile phone market has risen to 16 percent, and is estimated to reach 37 percent in 2014 [10]. Consequently, mobile services have become more and more important. Moreover, the current generations of mobile phones are developed in the same manner as personal computers. With integrated cameras, fast processors, full color displays, possibilities for wireless communication and embedded accelerometers the mobile phones have been transformed into ideal interfaces and command devices.

Various types of mobile phone interfaces for large displays have been proposed. Most of the current solutions require using finger drag and drop actions over the cell phone's screen to draw on the computer screen [11].

An intuitive basic set of tilt gestures is introduced for a stepwise or continuous interaction with both mobile applications and distant user interfaces by utilizing the handheld as a remote control [12]. An interface for interacting with large displays via cell phone with video camera is proposed in [13]. The location of the camera relative to the display using a matrix of transparent markers embedded on display. Pointing, rotating, dragging and dropping are supported.

In this paper we are discussing a detailed design of the proposed cell phone pointing interface based on its embedded accelerometer sensor for interaction with large projector based displays. The usage of natural hand gestures with the cell phone horizontally placed in the hand is used for performing the pointing and drawing tasks. The obtained acceleration signal is used to determine the direction and magnitude of the relative movement, which on the other hand is used for calculating the relative virtual cursor displacement. Our work is closely related with the one presented in [14] which uses the cell phone as a tilt device for car-game control. But, it extends this work, since in our case pointing and selection tasks, as well as drawing and movement tasks, are supported. Moreover, a rigorous and systematic evaluation of the proposed interface has been conducted in order to determine the efficiency, accuracy and usability of the proposed interface. To the best of the authors' knowledge, no similar evaluation of this type of interface has been conducted so far.

2. CELL PHONE POINTING INTERFACE DESCRIPTION

System architecture of the proposed interaction interface is presented in Figure 1. As it can be observed, the server represents the main part of the system, which is responsible for communication and data transmission between the two ends: cell phone and projector based display.
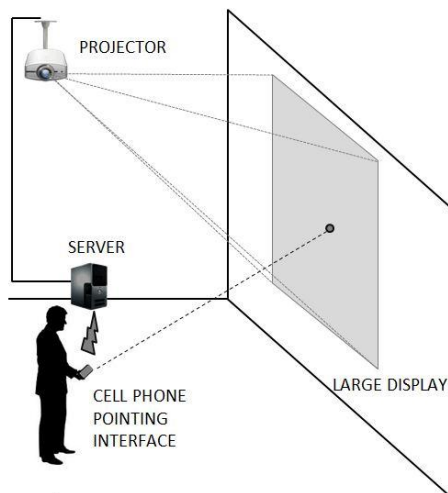

Fig 1: System architecture

From software point of view, the interface is composed of two Java-based applications: server application, working on the server computer that reads the mobile device's movements over a common WiFi network, transmitted by the client application working on the cell phone, responsible for gesture processing and their real-time translation into pointer actions. The phase of gesture processing is composed of three main steps:

- *Data acquisition* – mobile phone accelerometer acquire data in the form of 3 numbers representing the acceleration of the device along three axis (X, Y, Z), subtracted by the gravity G. Android API is not performing hardware based data sampling i.e. it is sampling the data on an event, which fires whenever the Android OS determines that the accelerometer device has determined changes in the input signal. Due to the varying activity loads of the OS, this event is not fired regularly. To solve this problem, first of all a most suitable regular sampling rate was experimentally obtained taking into account the Nyquist criteria. After analysis of multiple data series it was determined that 100 samples per second was a satisfactory sampling rate. Moreover, this is half of the standard mouse delay. To determine the value at the desired sample time a linear interpolation between two neighboring (one before and one after) samples obtained by the device was used.

- *Accelerometer signal processing* – In this phase the signal was to eliminate some parts of the data stream that do not contribute to the gesture. To remove the noise from the signal it was first filtered using the low-pass Savitzky–Golay smoothing filter. Rather than having its properties defined in the Fourier domain, and then translated to the time domain, Savitzky-Golay filter derives directly from a particular formulation of the data smoothing problem in the time domain. A digital filter is applied to a series of equally spaced data values fi ≡ f(ti), where ti ≡ t0 + iΔ for some constant sample spacing Δ and i = . . . − 2,−1, 0, 1, 2, . . . This digital filter simply replaces each data value fi by a linear combination gi of itself and some number of nearby neighbors, as given in Eq. 1.

$$g_i = \sum_{n=-n_L}^{n_R} c_n f_{i+n} \qquad (1)$$

Here nL is the number of points used "to the left" of a data point i, i.e., earlier than it, while nR is the number used to the right, i.e., later.

- *Planar position calculation* – the obtained 3D acceleration data will be used to calculate the current planar position of the pointer on the large display. By convention we have decided to use the cell phone as a pointing device horizontally placed in the operators hand and therefore, we have simply projected the acceleration vector to the XZ plane and transformed the cell phone XZ coordinate system to the projectors coordinate system.

To integrate the measured acceleration into a change in position along each of the axis the following equation was used:

$$x(n) = x(n-1) + K_x \cdot (v_x(n-1)\Delta t + \frac{a_x \Delta t^2}{2}) \qquad (2)$$

where the constant Kx is normalization constant related to the desired interface resolution. Using Verlet integration, we

would approximate the integral of the acceleration to the second degree. The above equation will be applied just in case when the acceleration signal is larger than a predefined threshold value TH. TH is the minimum signal in absolute value for recognizing a valid movement along each of the axes. The obtained planar data are finally send to the projector and projected to the environment.

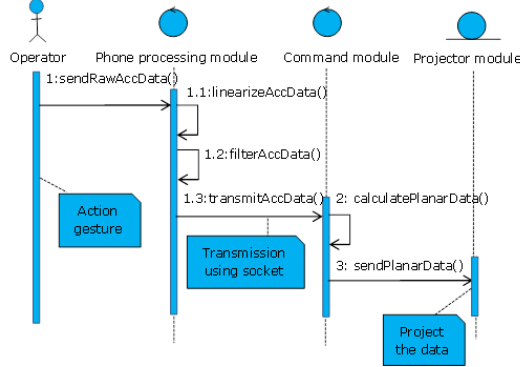The sequence diagram of the proposed interface is presented in Figure 2.

Fig 2: Sequence diagram describing the cell phone pointing interface.

3. EVALUATION OF THE PROPOSED CELL PHONE POINTING INTERFACE

3.1 Experimental setup

For experimental verification of the proposed interface a simple laboratory experimental setup has been created. It was composed of one MX613ST Benq Projector, and one PC acting as a server computer with Gigabyte GA-M770-UD3 motherboard, CPU AMD Athlon x4 620 BOX, VGA Gigabyte nVidia GTS250 1GB, DDR 2GB DDR2, with Windows 7 OS. The speakers were used to provide audio feedback to the user. LG P500 Optimus mobile phone with 600MHz CPU, 2GB RAM and Android OS (Froyo) was used as a pointing device. The photo of the experimental setup environment is shown in Figure 3.
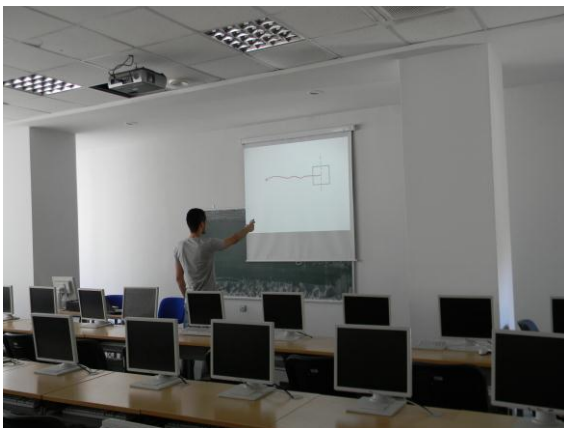

Fig 3: Experimental environment setup

3.2 Evaluation methodology

To evaluate the efficiency of the proposed interface in pointing and trajectory based tasks three kind of experiments were conducted. The first one was dedicated to the "target

pointing" task and aimed to derive a quantitative model for predicting its difficulty. The second experiment, a "tunnel steering" task aimed to describe the dependency between movement time and continuous constraint in tasks of steering along a given trajectory. The third experiment aimed at evaluation of the repeatability in reference path following tasks.

The experiment included two sessions for each task: a practice session and a measurement session. The practice session lasted until participants reached satisfactory level of working experience with the interface.

Overall 50 subjects, 25 male and 25 female were involved in the experiments. All of them have performed both tasks. The participants ranged in age from 19 to 27. All participants were students at the Faculty of Computer Science, University Goce Delcev – Stip. Each of the participants used his/her preferred hand. None of them was disabled.

In the first experiment we have used the experimental setup shown in Figure 4 to calculate the time necessary to move the pointer (MT) and select a target of width (W) which is placed at a distance (A) according to the Fitts' law [15] using the modification proposed by MacKenzie [16] and expressed by Equation (3).

$$MT = a + b \cdot \log_2 (A/W + 1) \qquad (3)$$

Where MT is the movement time; A – the distance to move; W – the width of the target region within which the move terminates; a and b are empirically determined regression coefficients (a represents the intercept and b the slope of the line - the inherent speed of the device). The factor $\log_2(A/W+1)$, called the index of difficulty (ID), describes the difficulty to achieve the task: the greater ID, the more difficult the task.
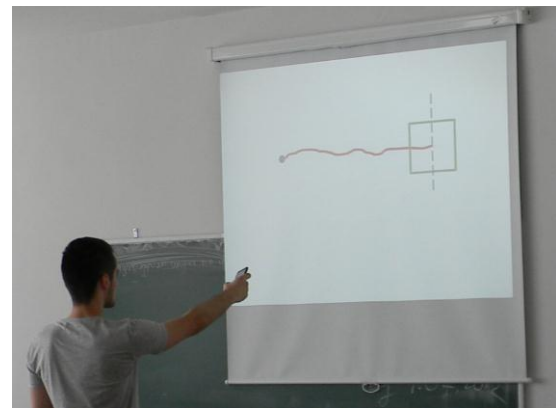

Fig 4: Measurement setup to test the pointing capabilities

In the second experiment, the participants were asked to steer the pointer using the mobile phone, inside a two-dimensional straight tunnel with constraints only at the ends of the movement was performed, as illustrated in Figure 5.

They were asked to pass the left end of the tunnel and then to reach the right one as quickly as possible.

At the beginning of each trial, a rectangle with green line presented on the large display. After issuing a command for steering, the subject began to draw a blue line on a screen, showing the pointer trajectory. When the pointer crossed the left end of the tunnel, the line turned red and the time was being recorded. When the pointer crossed the right end time

measurement was being stopped. If the tunnel sideway lines were crossed by the pointer a sound signal was emitted and the trial was classified as erroneous. Participants were asked to minimize errors.

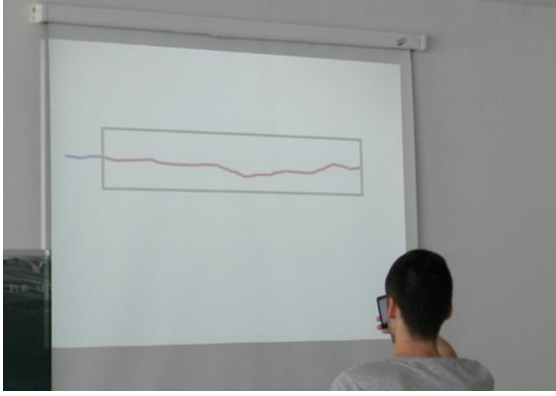The movement time between two tunnel ends was recorded and analyzed.

Fig 5: Measurement setup to test the steering capabilities

In this case the Accot-Zhai [17] steering law was applied as a predictive model for human movement (Equation 4):

$$MT = a + b \cdot \int_C \frac{ds}{W(s)} \qquad (4)$$

where MT is the average time to navigate through the path, C is the path parameterized by s, W(s) is the width of the path at s, and a and b are experimentally fitted constants. In the case of a straight tunnel the formula gets simplified:

$$MT = a + b \cdot \frac{A}{W} \qquad (5)$$

In this case the index of difficulty is expressed as ID=A/W.

In order to test the path repeatability, ISO 9283 standard was used. According to this standard path repeatability expresses the closeness of the agreement between the attained paths for the same command path followed n times in the same direction. Path repeatability is expressed by RTp-the maximum RTpi which is equal to the radius of a circle in the normal plane and with its center on the barycenter line (Figure 6).

The path repeatability is calculated as following:

$$RT_p = \max RT_{pi} = \max[\bar{l}_i + 3S_{li}] \,; i = 1...m \qquad (6)$$

where:

$$\bar{l}_i = \frac{1}{n} \sum_{j=1}^{n} l_{ij} \qquad (7)$$

$$S_{li} = \sqrt{\frac{\sum_{j=1}^{n} (l_{ij} - \bar{l}_i)^2}{n-1}} \qquad (8)$$
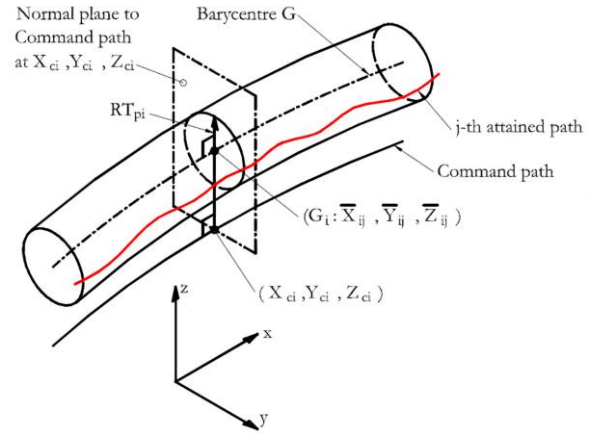
Fig 6: Target and experimentally obtained trajectory during the experiment (RT represent repeatability; G represents the barycenter of a cluster of attained poses; Xci, Yci and Zci are the coordinates of the i-th point of the command path; Xij, Yij and Zij are the coordinates of the intersection j-th attained path and the i-th normal plane)

$$l_{ij} = \sqrt{(x_{ij} - \bar{x}_i)^2 + (y_{ij} - \bar{y}_i)^2 + (z_{ij} - \bar{z}_i)^2} \qquad (9)$$

$$\bar{x}_i = \frac{1}{n} \sum_{j=1}^{n} x_{ij} \,; \bar{y}_i = \frac{1}{n} \sum_{j=1}^{n} y_{ij} \,; \bar{z}_i = \frac{1}{n} \sum_{j=1}^{n} z_{ij} \qquad (10)$$

m is number of calculated points along the path, and n is number of measurement cycles.

4. RESULTS

Regression was used to generate the Fitts' Law coefficients a and b for pointing interface. For this purpose two independent variables were considered: the distance from the target (A = 550, 650 and 750 mm) and objects' width (W = 100, 150, 200, 250, 300, 350, 400 and 450 mm). For each pair of values each of the participants has performed 5 tests and the results were recorded and aggregated. The data and the line obtained with linear regression are presented in Figure 7.
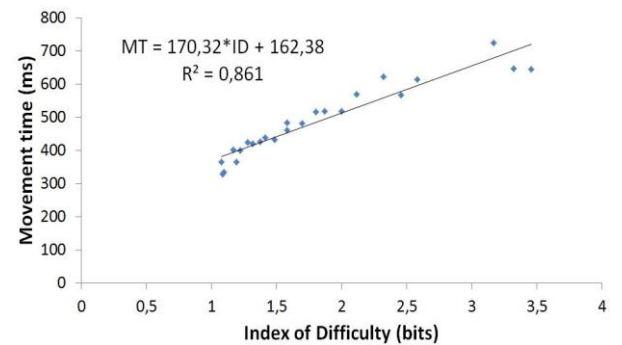
Fig 7: Scatter-plot of the MT-ID relationship for pointing task

In the second experiment two independent variables were considered: the tunnel length (A = 180, 360, 720 and 1440 mm) and path width (W = 100, 150 and 200 mm). The twelve conditions were presented in a random order and for each pair of values each of the participants have performed 5 tests. For each of the twelve conditions, the results were recorded and

aggregated. The data are plotted on a scatterplot and linear regression was performed (Figure 8).

As it can be observed the hypothesized model can be successfully applied in describing the difficulty of the task. Considering successfully completed trials only, the regression analyzes, gave:

$$y = 179.24x + 518.03; R^2 = 0.9661 \qquad (11)$$

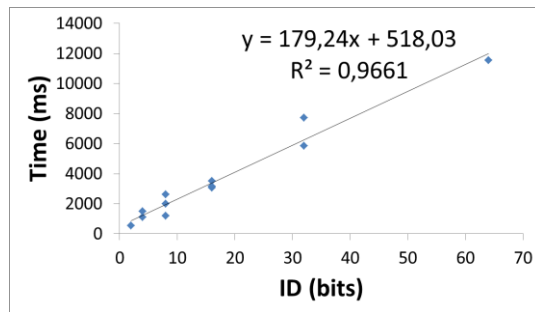The average error rate is 17% and it increases as the task become more difficult.

Fig 8: Scatter-plot of the MT-ID relationship for the steering task

An extremely strong correlation was found to exist between movement time and index of difficulty in both tasks. The R^2 values were 0.861 and 0.9661 in both cases respectively. Therefore, these findings indicate that Fitts' Law and Accot-Zhai steering law are highly accurate predictor of movement time in both interfaces. On the other hand, when performing a task based on practice trials, people improve in speed. Learning usually follows a power function, which means that people learn at a higher rate in the beginning while the curve bottoms out over time [18].

$$T_N = T_1 \cdot N^{-\alpha} \qquad (12)$$

where T1 stands for the time necessary to complete the task in the first trial, TN the time necessary to complete the task in the Nth trial, N is the trial number, and α is an empirically determined coefficient. To verify the user skill acquisition using the proposed cell phone interface in pointing tasks, five random participants without previous experience with the proposed interface were chosen. They were asked to point inside a target object with the width of 200mm starting from a point which is 450mm distant from the object. The pointing task was performed 20times by each of the participants and the averaged results for each trial are shown in Figure 9.
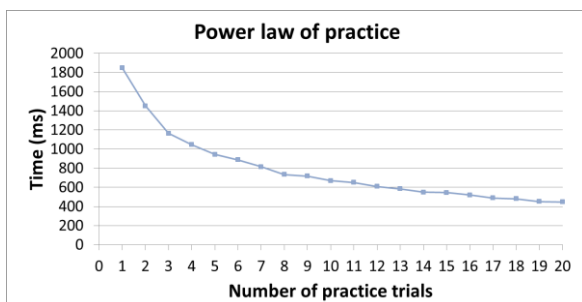

Fig 9: Power Law of Practice for pointing task

The results show almost perfect fit with the power law of practice formula, with α=0.37. One can easily observe that the

largest improvements in speed are made during the very first trials. Therefore, we should be careful with generalizing timing results from first-time users. The results are also useful to gain a conclusion regarding the number of trials after which there is no more significant learning. In our case for the given task it was determined that it corresponds to trial number 13

For this task, all the participants were asked to follow a simple circle path on the projected large display and their trajectories were recorded and analyzed offline. Each participant has made 10 trials and then their results were averaged.

The obtained results for path repeatability test were calculated according to the equations 6-10 and the corresponding results for all tests are shown in Table 1.

Table 1. Path repeatability results

Test number	Repeatability (mm)
1	1.1
2	1.3
3	1.4
4	2.3
5	1.8
6	2.3
7	2.1
8	1.7
10	1.4

Moreover, at the end of the experiment all users were asked to fill in a questionnaire with 3 questions regarding their experience with the proposed cell phone pointing interface for robot control task. The opinions of users were rated through a 5-point Likert scale (5 for strongly agree; 1 for strongly disagree) for evaluation. The results are presented in the table below.

Table 2 User opinions regarding the developed cell phone pointing interface for interaction with large projector based display

Question	Mean	Standard Dev.
Did you find the proposed interface easy for use?	4.15	0.92
Are you satisfied with the obtained results in task fulfillment?	3.96	1.12
Do you think that this interface is acceptable for non-experienced users?	4.35	0.85

5. CONCLUSION

In this paper we propose a design of a cell phone pointing interface based on its embedded accelerometer sensor for interaction with large projector based displays. The efficiency of proposed design was experimentally tested on various tasks. The results have revealed its accuracy, low error rates and easiness for usage. So, we can state that the proposed

interface seems appropriate for usage in various large projector based display applications.

In future, we want to apply this interface for enrichment of interactive teaching at the university.

6. REFERENCES

[1] Deller M., Thelen S., Steffen D., Olech P.S., Ebert A., Malburg J., Meyer J. A Highly Scalable Rendering Framework for Arbitrary Display and Display-in-Display Configurations. In Hamid R. Arabnia and Leonidas Deligiannidis, editors, CGVR, pages 164–170. CSREA Press, 2009.

[2] Lee J.C., Hudson S.E., Tse E., Foldable Interactive Displays, Proceedings of the 21st annual ACM symposium on User interface software and technology, 2008

[3] Roman P., Lazarov M., Majumder A., A Scalable Distributed Paradigm for Multi-User Interaction with Tiled Rear Projection Display Walls, IEEE Transactions on Visualization and Computer Graphics, Volume: 16 , Issue: 6, Page(s): 1623 - 1632 2010

[4] Johnson T., Welch G., Fuchs H., La Force E., Towles H., "A Distributed Cooperative Framework for Continuous Multi-Projector Pose Estimation" Proceedings IEEE VR 2009, Lafayette, LA, March, 2009

[5] Ng J., Sim T.J., Foo Y.S., Yeo V., Gesture-based interaction with virtual 3D objects on large display: what makes it fun?, Proceedings of the 27th international conference extended abstracts on Human factors in computing systems, 2009

[6] Jung C., Lee M., Laser spot pattern recognition based computer interface using I/O mapping sensitive neural networks, Proc. of 2011 IEEE International Conference on Consumer Electronics (ICCE), p. 847 - 848, 2011

[7] L. Hespanhol, M. Tomitsch, K. Grace, A. Collins, and J. Kay. Investigating intuitiveness and effectiveness of gestures for free spatial interaction with large displays. In E. Huang, editor, 2012 International Symposium on Pervasive Displays, 2012

[8] Banerjee A., Burstyn J., Girouard A., Vertegaal R., MultiPoint: Comparing laser and manual pointing as remote input in large display interactions, In Press, Corrected Proof, Available online 31 May 2012

[9] ITU Corporate Annual Report. (2009). http://www.itu.int/dms_pub/itu-s/opb/conf/S-CONF-AREP-2008-E06-PDF-E.pdf (Online, visited 10.05.2012).

[10] Smartphone forecast by Pyramid Research, http://www.pyr.com/pr_prlist/PR121009_SMART.htm (Online, visited 10.05.2012)

[11] Hardy R., Rukzio E., Touch & interact: touch-based interaction of mobile phones with displays, MobileHCI '08 Proceedings of the 10th international conference on Human computer interaction with mobile devices and services, P.245-254, 2008

[12] Dachselt R., Buchholz R., Natural throw and tilt interaction between mobile phones and distant displays, CHI EA '09 Proceedings of the 27th international conference extended abstracts on Human factors in computing systems, P. 3253-3258, 2009

[13] Hyakutake A., Ozaki K., Kitani K.M., Koike H., 3-D interaction with a large wall display using transparent markers, Proceedings of the International Conference on Advanced Visual Interfaces (AVI '10), P. 97-100, 2010

[14] Tamas Vajk, Paul Coulton, Will Bamford, and Reuben Edwards, "Using a Mobile Phone as a "Wii-like" Controller for Playing Games on a Large Public Display," International Journal of Computer Games Technology, vol. 2008, Article ID 539078, 6 pages, 2008. doi:10.1155/2008/539078

[15] Paul M. Fitts. The information capacity of the human motor system in controlling the amplitude of movement. Journal of Experimental Psychology, volume 47, number 6, June 1954, pp. 381–391 (1954).

[16] MacKenzie I. S. and Buxton W. A. S. Extending Fitts' law to two-dimensional tasks. Proceedings of ACM CHI 1992 Conference on Human Factors in Computing Systems, pp. 219–226 (1992).

[17] Accot J. and Zhai S. Refining Fitts' law models for bivariate pointing. Proceedings of ACM CHI 2003 Conference on Human Factors in Computing Systems, pp. 193–200 (2003).

[18] Stafford T, Thirkettle M, Walton T, Vautrelle N, Hetherington L, et al. (2012) A Novel Task for the Investigation of Action Acquisition. PLoS ONE 7(6): e37749. doi:10.1371/journal.pone.0037749