Gesture-Based Interaction for Cultural Exhibitions

The Effect of Discrete Visual Feedback on the Usability of In-Air Gesture-Based User Interfaces

Tin-Kai Chen[1,*], Robert C.C. Chen[2], and Fong-Gong Wu[3]

[1] Department of Animation and Game Design, Shu-Te University,
No.59, Hengshan Rd., Yanchao Dist., Kaohsiung, 82445, Taiwan, ROC
tkchen@acm.org
[2] Art & Design Faculty, De Montfort University,
The Gateway, Leicester, LE1 9BH, United Kingdom
rchen1@dmu.ac.uk
[3] Department of Industrial Design, National Cheng Kung University,
No.1, Ta-Hsueh Road, Tainan, 70101, Taiwan, ROC
fonggong@mail.ncku.edu.tw

Abstract. The study aims to reveal the effect of in-air gesture interaction using the depth camera technology on complex human performance and to identify possible design failures and its implementation to the digital shadow play. Since in-air coordinate system of body kinematics shares the same directional vector with on-screen coordinate of the visual character, a systematic approach "directional vector transformation" has been proposed for transforming the in-air coordinate into the on-screen coordinate. A comprehensive literature review of human computer interaction, digital shadow play and gesture interfaces is given. Finally, identification of design failures and design guideline for further study are made for the design of in-air gesture interfaces.

Keywords: Gesture Interfaces, Depth Camera, Fitts' law, FLG, Usability.

1 Introduction

Among recent studies toward the digital shadow play, it was claimed that shadow play is one of the intangible cultural heritages in China and was popular in most areas of the country but only few people watch the play now because of the increasing popularity of new media [3]. In order to attract more people's attention, a new digitized performance method needs to be designed based on multi-touch technology which the method maps transforms of a small group of controlling points to character controlling points accordingly, with innovative mapping rules that decrease the complexity of controlling with fingers [3] . In Malaysia, shadow puppet plays is a traditional Malaysian theater art, which slowly loses its appeal to adolescents, who prefer computer games. In order to help reverse this decline, a 3D Seri Rama prototype was developed incorporating the traditional Seri Rama character into the Street Fighter video game, using modeling, texturing, and animation [4]. The prototype allows users to control

C. Stephanidis and M. Antona (Eds.): UAHCI/HCII 2013, Part I, LNCS 8009, pp. 454–463, 2013.

Seri Rama with a PlayStation game controller. This approach is mainly more towards the signs of awareness especially towards the younger generation that shadow play puppet is still an influential performing arts master piece heritage. Indirectly, this will help to open more opportunities for businesses and marketing plans including the tourism sector as more visitors will visit Malaysia in future [4].

In Taiwan, the Kaohsiung County is known as the homeland of puppet arts, digital shadow play has been developing and exhibiting to installing new energy into Taiwan's traditional puppet theatres [1][2]. (Fig. 1 and Fig. 2).


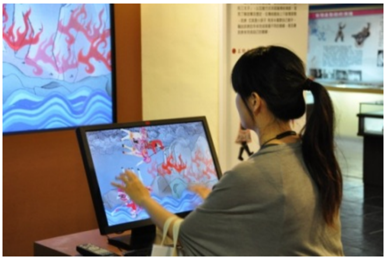

Fig. 1. and **Fig. 2.** The photo at left hand side demonstrates the first version of the digital shadow show "Wu-Song and Tiger", developed based on the 80" infrared-based multi-touch technology and the open source software library that allows the touch sensitivity of the wall to be modified. Furthermore, the photo at the right hand side shows the second version of the digital shadow show based on Multi-Touch Surface Capacitive Touch Display.

To sum up, this study aims to reveal the effect of in-air gesture interaction using depth camera technology on complex human performance and to identify possible design failures and its implementation to the digital shadow play. The remaining of the paper is organized as follows: In Section 2, the background review of related areas is surveyed. In Section 3, the proposed experimental approach is demonstrated. The result analysis is given in Section 4. Section 5 provides a detailed discusses based on the results. Finally, conclusions and future work are highlighted in Section 6.

2 Related Works

2.1 Directional Vector Transformation

Depth camera technology observes users through a video camera and recognizes gestures they make with different body parts, including hands, arms, legs, and general posture. It presents a far more advanced gesture-based user experience than any previous input devices.

In order to utilize the gesture data in the air for the design of 2D GUI, the gesture data needs to be transformed into on-screen coordinates. For instance, the relation between both coordinate systems is illustrated in Fig. 3(a) and 3(b). In Fig. 3(a), the right palm moves from in-air coordinate A (X1,Y1) to B (X2,Y2). Following by Figure 3(b), the on-screen coordinate known as A (x1,y1) is the right palm of the visual character on the screen, it tends to move to unknown B (x2,y2) with the same directional vector based on the in-air palm movement from A to B.

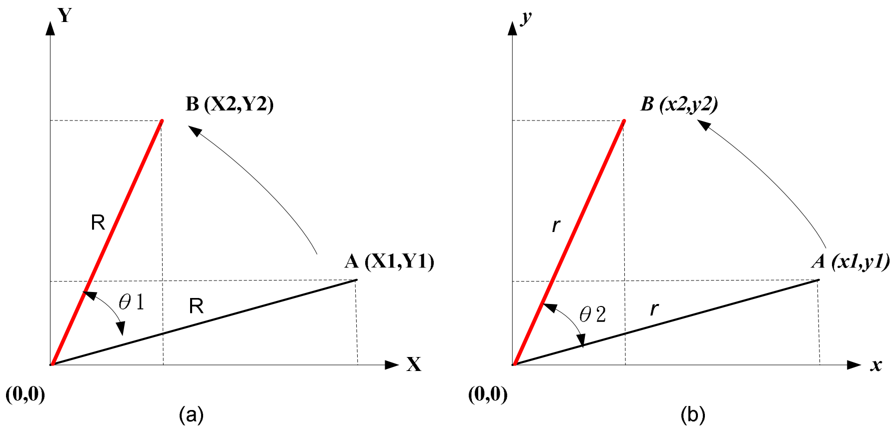

Fig. 3. (a) in-air coordinate system of the gesture in the real world and (b) on-screen coordinate system of visual gesture

This study proposes a systematic approach "directional vector transformation" for transforming the in-air coordinates into on-screen coordinate. It refers to two difference coordinates sharing the same directional vector, i.e. in-air coordinate of body kinematics, and on-screen coordinate of the visual character.

The followings are the procedure of the directional vector transformation based on the right palm movement; Firstly, the right palm moves from in-air coordinate A (X1,Y1) to B (X2,Y2), the directional vector for on-screen movement of the palm of the visual character can be expressed by a simple directional vector equation shown in Equation (1):

$$\mathrm{A\overline{B}} = (\mathrm{X1} - \mathrm{X2}, \mathrm{Y1} - \mathrm{Y2}) \tag{1}$$

where (X1-X2) and (Y1-Y2) represent the direction of the movement of the related body kinematics in terms of x and y coordinate system. Since there is a coordinate difference between of in-air and on-screen, the value of the directional vector needs to be normalized by the displacement between A (X1,Y1) and B (X2,Y2), namely $\mathrm{A\overline{B}}\, normalized$, shown in Equation (3):

$$Displacement = \sqrt{(\mathrm{X1} - \mathrm{X2})^2 + (\mathrm{Y1} - \mathrm{Y2})^2} \tag{2}$$

$$\mathrm{A\overline{B}}\, normalized = \left(\frac{(X1-X2)}{Displacement}, \frac{(Y1-Y2)}{Displacement}\right) \tag{3}$$

In this case, the on-screen coordinate known as *A (x1,y1)* indicates the right palm of the visual character, it needs to move to unknown *B (x2,y2)*, which the unknown *B (x2,y2)* can be obtained via the Equation (4):

$$B(x2, y2) = A(x1, y1) \times A\overline{B}\, normalized \times Constant \tag{4}$$

Based on the directional vector transformation, the digital shadow play based on in-air gesture interface using Kinect sensor was developed using Adobe AIR, show in Fig. 4. Essentially, there is a difference coordinate between of the coordinate mapping from the body kinematics captured by Kinect sensor, and of the inverse kinematics (IK) on the screen. Even through, the direction between of those is the same. In this study, Constant is the length of the lower limb on the screen, i.e. Constant = 200 dot. Although the *B (x2,y2)* will be never the same as B (x2,y2), but movement direction is the same since the angle between B and *B* is 0°.

Fig. 4. The picture shows the difference coordinate between of the coordinate mapping from the gesture in the real world (i.e. colour dots on the screen), and of the visual character on the screen. Even through, the direction movement between two coordinate systems is the same.

2.2 Kinect Cursor Emulator

It is difficult to measure the objective human performance of the prototype demonstrated in Fig. 4. Instead, this study designed a simulated programme based on Equation (4) to emulate the mouse cursor movement mapping from the right palm in the air, namely Kinect Mouse Emulator. The programme uses the Kinect for Windows SDK and its skeletal tracking features to allow a user to use their hands to control the Windows mouse cursor, modified from a demo application (Source: http://kinectmouse.codeplex.com/). It demonstrates how to use the Kinect skeletal tracking feature to move the mouse cursor with a user's hands based on.

2.3 Human Performance Model

Regarding measurement of human performance, there have been many practices in the field of Human-Computer Interaction (HCI). One of famous practices is Fitts' law. [5] The mathematical relationship among speed, accuracy, amplitude of movement, and target size for upper extremity tasks, which can be expressed by a simple liner regression equation shown in Equation (5):

$$MT = a + b \times ID \tag{5}$$

where ID is index of difficulty proposed by Fitts, *D* is distance between targets, *W* is target width, *MT* is movement time, and parameters a and b are calculated on the basis of simple linear regression. As expected, movement time for hard tasks is longer. Furthermore, MacKenzie [7] recommended the use of effective target width *We* instead of nominal target width *W* to measure actual performance of either devices or tasks:

$$W_e = 4.133 \times S.D \tag{6}$$

$$ID_e = \log_2(D / W_e + 1) \tag{7}$$

where *S.D.* is standard deviation of endpoint over target region, and ID_e is effective index of difficulty. Recently, the ID_e model in Equation (7) had been standardized in ISO 9241 [8] as a design and testing guideline, and specification of non-keyboard input devices (NKIDs). In this study, Fitts' law was expanded into two-dimensional description in a polar coordinate system.

3 Method

3.1 Subject Selection

A total of ten Taiwanese undergraduate students in the Department of Animation and Game Design, Shu-Te University volunteered. The participants consisted of five males, i.e. age range from 20 to 21 years, and five females, i.e. age range from 20 to 21 years. All participants used their preferred right hand to perform the tasks, and reported over 6 years' experience with PCs.

3.2 Testing Apparatus

The laboratory used for the experiment is a computer laboratory in Room DB105-3 in Department of Animation and Game Design, Shu-Te University. The experiment was conducted based on following equipment:

- PC with a Intel i5 2.67 GHz CPU, 4 GB of RAM;
- 60" projector screen showing targets.
- A standard two-button optic mouse with 800 dpi, manufactured by Logitech®.
- A 'Kinect Sensor', manufactured by Microsoft ®.
- Fitts' Law Generator (FLG) [1][2][13][16]. It can be used for multidirectional human performance measurement of Non-Keyboard Input Devices (NKIDs)
- Subjective assessment based on five-scale subjective questionnaire.
- A digital camera used to capture awkward postures of participants' performance during the experiment.
- A standard operation procedure (SOP).
- The data analysis is performed using SPSS version 17.

Fig. 5. Experiment condition in the computer laboratory in Room DB105-3

3.3 Experimental Procedure

The triangulation mixed method was implemented, it is a 2 × 2 × 2 × 8 within-subjects repeated measurement laboratory-based experiment, as shown in Table 1.

Totally, there were n = 10 subjects × 2 devices × 3 blocks × 32 target conditions = 1,920 pairs of dependent variables being observed by a measurement platform Fitts' Law Generator (FLG). Therefore, quantitative analysis method can be applied. Those dependent variables and related methods are shown in Table 2.

Table 1. 2 × 2 × 2 × 8 factorial conditions

Factors	Levels
Devices (2)	Mouse, Kinect Sensor
Target distance (mm) (2)	100, 300
Width/Height (mm) (2)	300, 500
Angle of Approach (degree) (8)	0°, 45°, 90°, 135°, 180°, 225°, 270°, 315°

Table 2. Objective Dependent Variables

Factors	Description
De	Cursor movement distance (mm)
AT	Time of aiming target (ms)
PT	Time of decision making (ms)
MT	Time of movement (ms)
Error	Error (%)
TRE	Target-re-entrance (%)
W_{sd}	Standard deviation of pointing x by conditions (mm)
H_{sd}	Standard deviation of pointing y by conditions (mm)
W_e	$4.133 \times W_{sd}$ (mm)
H_e	$4.133 \times H_{sd}$ (mm)

3.4 Research Limitation

Kinect Cursor Emulator and the digital shadow play are both the in-air gesture interface using Kinect Sensor, sharing the same concept expressed in Equation (4). Since the digital shadow play does not require button activation, thus the result analysis regarding the pointing time is only recommended for further study.

4 Result Analysis

4.1 Data Adjustment

An error occurred when a participant registered a target acquisition while the cursor was outside the target.[13] Therefore, error cases are analyzed separately. Since there are two pointing device being tested, a total of 15 errors occurred out of 960 total trials with the mouse (1.6% error rate) and a total of 65 errors occurred out of 960 total trials with the Kinect sensor (6.8% error rate).

As for the mouse, the mean *MT* for all trials is 625 ms, and the removal of the error trials reduces the mean *MT* to 604 ms. With regards to the Kinect sensor, the mean *MT* for all trials is 2,689 ms, and the removal of the error trials reduces the mean *MT* to 2,528 ms.

4.2 Fitness-of-Models

As can be seen in Table 3, comparing with the result done by Chen and Chen [13] who reported an adjusted $R^2 = 0.44$ with a mouse, this study obtain higher adjusted R^2 values across the *ID* model. Therefore, our study is consistent with current study.

Table 3. The prediction of the total movement time (*MT*) (ms) across models (ID, ID_e, ID_{e2})

Device	N*	Model	
		ID	ID_e
Mouse	945	0.479	0.396
Kinect Sensor	895	0.613	0.436

* The error trials were excluded for the analysis.

** The linear regression analysis was applied on the adjusted data for the prediction of the movement time *MT* across models (ID and ID_e). The adjusted R^2 value was used since the sample size was difference among these studies.

4.3 Device Difference

As shown in Table 2, Independent T test indicates that the mean of movement time, i.e. 1,941 mm, is significantly longer than for the mouse, i.e. 456 mm, $p < 0.000$. Moreover, the approaching time, the pointing time and the total movement time for Kinect sensor are all significant longer than the mouse, i.e. $p<0.05$. However, the result analysis reveals that there is no difference of cursor movement distance between of the mouse and of the Kinect sensor, $p=0.212$, Hence, it is necessary to explore the design failures that cause similar cursor movement distance and longer movement time via the posture analysis and the subjective assessment.

Table 4. Objective data

Dependent Variable	Device	N	Average	S.D.
Error Rate*	mouse	960	1.6%	.065
	Kinect	960	6.8%	.111
Target-Re-Entrance*	mouse	945	3%	.161
	Kinect	895	9%	.323
Movement Time (ms)**	mouse	945	456.14	1449
	Kinect	895	1940.58	830
Pointing Time (ms)**	mouse	945	148.56	152
	Kinect	895	587.21	218
Approaching Time (ms)**	mouse	945	604.71	227
	Kinect	895	2527.78	868
Cursor Movement Distance (mm)	mouse	945	437	119
	Kinect	895	419	153

Remark: *: $p<0.05$. **: $p<0.001$. Except of the error rate, error cases are excluded for analysis.

4.4 Direct Observation

When using the Kinect Cursor Emulator, various arm and body postures of those with a preference for right handed working can be categorized in terms of the shoulder flexion angle $\theta 1$ and the back flexion angle $\theta 2$:

- Angle $\theta 1$: shoulder flexion
- Angle $\theta 2$: back (spine) flexion

Based on the reviewed literature, the awkward working posture is defined as the posture having the joint range apart from the neutral posture.[15]. The observation reveals that there were two operational postures being defined, shown as follows:

- Type I: It is the neutral position where *$\theta 1$ and $\theta 2$ are approaching to 0^{o}.*
- Type II: It is the awkward position where *$\theta 1$ and $\theta 2$ apart from 0^{o}.*

As can be seen in Fig. 6, it is a selected pictures taken from the experiment, where the awkward postures are identified among the pointing task with in-air gesture interface using Kinect sensor.

Fig. 6. Type II operational posture, where the subject's back flexion is 25 ^{o} approximately

5 Discussion

By the implementation of the triangulation mixed method, two possible issues that caused awkward working posture whist using the in-air gesture interface, were highlighted: Firstly, based on the human performance study, it reveals that end-users tend to maintain the same joint range of particular body regions in order to take control of the cursor movement, especially where the angle of approach appears from the middle to bottom of the screen, i.e. 135°, 180 °, 225 °, 270 ° and 315 °:

- Kinect sensor is not suitable for rapid and accuracy pointing activity. It was highlighted that the scenario related with Kinect application is not focused on productivity, speed, and precision, but enables an interaction that other input devices cannot [17].
- The position of the visual target must be designed over the middle of the screen.
- The cursor movement speed should be adjusted based on a repetitive design process with user test to achieve better quality-in-use.

6 Conclusion

The study has achieved it aims by revealing the effect of in-air gesture interaction using the depth camera technology on complex human performance and to identify possible design failures and its implementation to the digital shadow play.

References

1. Chen, T.K., Chen, C.C., Wu, F.G., Yang, H., Nivala, W.: The Social Interfaces Development for Cultural Creativity Industry: A User-centred Design Case Study on Chinese Shadow Puppetry. In: Proceedings of International Association of Societies of Design Research (IASDR 2011), TU Delft (2011)
2. Chen, C.C., Chu, C.C., Chen, T.K., Yang, H.J.: An Investigation into the Effect of Gesture Interaction in Relation to Human Performance for Identifying Possible Design Failures. International Transactions on Systems Science and Applications (ITSSA) 6(1), 1–12 (2010); Systemic and Informatics World Network (SWIN), Glasgow (2010)

3. Yao, C., Xiao, J., Guo, T.G., Xia, L.: Real-time DigitisedShadow Play Performance Method based on Multi-point Interactive Controlling Method. International Journal of Computer Applications in Technology (IJCAT) 38(1/2/3), 86–92 (2010)
4. Bin, D., Ghani, A.: Dissemination of "Seri Rama" Shadow Play Puppet as a Cultural Heritage through Capcom's Street Fighter IV. In: Proceedings of the 6th International Conference on Ubiquitous Information Management and Communication (ICUIMC 2012), article no. 106 (2012)
5. Fitts, P.M.: The Information Capacity of the Human Motor System in Controlling the Amplitude of Movement. Journal of Experimental Psychology 47(6), 381–391 (1954); Reprinted in Journal of Experimental Psychology: General 121(3), 262-269 (1992)
6. Yang, N., Zhang, M., Huang, C., Jin, D.: Motion Quality Evaluation of Upper Limb Target-researching Movements. Medical Engineering & Physics 24, 115–120 (2002)
7. MacKenzie, I.S.: Movement Time Prediction in Human-Computer Interfaces, 2nd edn. Readings in Human-Computer Interaction, pp. 483–494. Kaufman (1995)
8. [8] ISO 9241: Ergonomic Requirements for Office Work with Visual Display Terminals (VDTs) - Part 9: Requirements for Non-keyboard Input Devices. BSI press, UK (1999)
9. Akamatsu, M., MacKenzie, I.S.: Movement Characteristics Using a Mouse with Tactile and Force Feedback. International Journal of Human-Computer Studies 45, 483–493 (1996)
10. Hsu, S.H., Huang, C.C., Tsuang, Y.H., Sun, J.S.: Effects of Age and Gender on Remote Pointing Performance and Their Design Implications. International Journal of Human-Computer Interaction 23, 461–471 (1999)
11. Sato, S., Kitajima, M., Fukui, Y.: A Proposal of Grasping Information Interface for Remote Pointing Task. In: Proceedings of the 5th IEEE International Symposium on Assembly and Task Planning, Besancon, pp. 175–180 (2003)
12. MacKenzie, I.S., Kaippinen, T., Silfververq, M.: Accuracy Measures for Evaluating Computer Multimodal Interfaces. In: Proceedings of the SIGCHI Conference on Human Factors in Computing System, Washington, pp. 9–16 (2001)
13. Chen, C.C., Chen, T.K.: The Effect of Gender-Related Difference on Human-Centred Performance Using a Mass Assessment Method. International Journal of Computer Applications in Technology (IJCAT) 32(4), 322–333 (2008)
14. Law, C.M., Vanderheiden, G.C.: Tests for Screening Product Designs Prior to User Testing by People with Functional Limitations. Proceedings of the 43rd Triennial Congress of the International Ergonomics Association and 44th Annual Meeting of the Human Factors and Ergonomics Society 5, 868–872 (2000)
15. Pheasant, S.: Body space – Anthropometry, Ergonomics and the Design of Work, 2nd edn. Taylor & Francis Press (1998)
16. Chen, C.C., Wu, F.G., Chen, T.K.: Extension of Fitts' Law for the Design of the Gesture Pointing Interaction: The Effect of the Phenomenon of Discrete Cursor Movement on the Usability of Gesture-based Pointing Devices. In: International Association of Societies of Design Research (IASDR 2011), Seoul (2009)
17. Microsoft: Human Interface Guidelines-Kinect for Windows v1.5., `http://www.microsoft.com`