

Exploring the Effects of Group Interaction in Large Display Systems

Hao Jiang^{a,b,c,*}, Chang Gao^{b,c}, Tianlu Mao^{b,c}, Hui Li^d, Zhaoqi Wang^{b,c}

^aHunan Engineering Laboratory of Digital Preservation Technology of Traditional Settlements, Hunan, 421002, China

^bBeijing Key Laboratory of Mobile Computing and Pervasive Device, Beijing, 100190, China

^cInstitute of Computing Technology, Chinese Academy of Science, Beijing, 100190, China

^dSchool of Computer Science and Technology, Sichuan University, Sichuan, 610065, China

Abstract

Large display systems have been successfully applied in the virtual reality domain because they can provide a full immersion sense through large visual space and high display resolution. However, most of the previous studies on the interaction method of those systems focused on single or double users. In this paper, we study the effects of integrating group interaction in such systems and we propose a framework called “Groupnect”, which enables the unique experience of group interaction in a large display system. By using optical tracking and 3D gesture recognition technologies, our approach can automatically recognize gesture-based control signals for 20 users simultaneously and trigger corresponding real-time actions in a back-end system. Through a comparison experiment between standard interaction mode and group interaction mode, the results demonstrate that physical and mental participation of users could be promoted in group interaction mode. It has immense potential to design group interaction applications on entertainment, education and training areas.

Keywords: large display; group interaction; VR application and framework

(Submitted on October 18, 2017; Revised on November 19, 2017; Accepted on December 17, 2017)

© 2018 Totem Publisher, Inc. All rights reserved.

1. Introduction

Nowadays, large display systems play a significant role in different VR applications. These systems [5,6] not only provide excellent visual quality, but also can support a dozen users in a shared environment.

Users in large display systems can obtain high immersion feelings, but only a few of them may interact with large screens. Therefore, it constrains user experience and application design in VR areas. On the other hand, some researchers [23, 19] found that, in their applications (for example, sports game), better performances can be achieved when users compete against others or when they collaborate to achieve a common goal. So, we think that a large display system that supports multi-user interaction can attract users to pay more effort and perform better than standard systems. By using natural interaction devices, many existing large display systems can support interactive operations for several users (general one or two) [24,17]. But if the number of users increase (maybe a dozen users), the technical framework and application need to be carefully elaborated to provide a natural interactive experience for all users.

Aiming at this problem, this paper presents Groupnect, which enables the unique experience of group interaction (a dozen users naturally interact with a shared system simultaneously in real time) in large display systems. In this framework, we use optical technology to track the markers that the users hold in their hands (as shown in Figure 1.). By using temporal spatial cooperated labeling technology and the multi-threading programming method in 3D gesture recognition process, the system can automatically recognize a dozen users' gesture-based control signals simultaneously. The background system can receive those signals and trigger corresponding actions in real time. Moreover, an entertainment application named “Hitting monsters” is developed for testing this group interaction framework. Based on this application, we designed a contrast experiment to

* Corresponding author.

E-mail address: jianghao@ict.ac.cn

compare a standard interaction mode to our approach and evaluate the result from both subjective and objective aspects. The experiment results demonstrate that, by integrating group interaction into a large display system, this framework can promote the users' physical and mental participation. So, we believe this framework has immense potential to design novel applications in entertainment, education and training areas.

This paper mainly introduces Groupnect, a system integrating group interaction into a large display system. In the next section, we introduce the related work, and compare some related system with our framework in details. In the third section, we mainly describe hardware and software modules of our framework. We show the experiment for the work of this paper in section 4. Section 5 is the experiment result. The conclusion and discussion is in Section 6.

2. Related Work

Since the CAVE (cave automatic virtual environment) was invented at the University of Illinois Electronic Visualization Lab in 1992[5], large display systems have exploded in size, resolution, and ubiquity. Those systems have been the focus of many researchers and have been applied widely in entertainment, education and training areas for its high immersion characteristics.

Classified by hardware configurations, Ni et al. [18] divided large display systems into CAVE and derivatives, multi-monitor workstations, tiled LCD panels, reconfigurable projector arrays, stereoscopic displays and volumetric displays. Various kinds of large display systems are used in different areas. Febretti et al. [6] introduced Omegalib, a Multiview application framework, for hybrid reality display environments that support collaborative work. The Super-KAVE [9] was employed as an immersive visualization tool for neutrino physics, which used a CAVE to immerse users and provided a new visualization technique for neutrino-interaction patterns.

With decades of research, lots of devices have been invented to interact with large display systems. The traditional mouse and keyboard are somewhat difficult to use when users want to move. Randy Pauschs et al. made some study on a Disney Aladdin ride [20]. Those kind of interaction devices in this system are widely used in many large display systems. Aspin took an initial study into augmented inward-looking exploration and navigation in CAVE-like IPT systems [1]. In this study, sensors were attached to the user's head, tablet device and joystick device. Kim, J et al. [11] tried to use iPhone/iPod touch as input devices for navigation in immersive virtual environments. By using the FWIP technique motioned in this work, users can navigate large display systems with sufficient precision. Nancels et al. [16] implemented high-precision pointing on large wall displays by using small handheld devices (a tablet), and designed and evaluated techniques that use various input channels to improve pointing at very small targets across large amplitudes. Olivier et al. [3] made Smarties, which allowed wall application developers to easily add interactive support to their collaborative applications through a mobile device. Users can interact with large displays with high accuracy using devices while carrying accessory equipment (pad, tablet, wand), which may be a little cumbersome.

Compared with the above interaction method, 3D gesture interaction is more natural. By tracking markers, the hand or the whole body can interact with the system in any location inside the tracking region with small hardware in hand. Cao and Balakrishnan [2] invented a variety of techniques and widgets for interacting with large scale displays using a button-free wand tracked in 3D, which is a new input mechanism for interacting. LaViola et al. [13] developed a set of novel input devices for CAVE-based virtual environments. Hand gestures or foot gestures were used for navigation in this work. Cheng and Pulo [4] addresses the Ray-pointing problem in large displays with hot spots, regions surrounding objects of interest, and triggers an action on the display naturally and easily. Vogel and Balakrishnan [24] explored the design space of freehand pointing and clicking interaction with very large high-resolution displays from a distance. In their work, three techniques for gestural pointing and two for clicking were developed and evaluated. Kopper and Bowman et al. [12] studied four models and proposed a model of human performance for distal pointing. Jota et al. [10] tested four ray pointing variants on a wall display that covers a large part of the users' field view to find influence of control type and parallax. Nancels et al. [17] studied different families of location-independent, mid-air input techniques for pan-zoom navigation on large displays and identified the key factors for the design of such techniques. By using different tracking technologies, those systems enable 3D gesture interactions with large displays in a natural way.

But as we know, most of the 3D interaction systems only support single or a few users' (general one or two) interaction in a local shared environment (some work support users' collaboration but users are in separate places [22]). To deal with a dozen users' interaction, they need to be redesigned to solve the labeling problem. In recent years, with the increasing size of large display systems, more and more users can experience those systems together at the same time. If more users can interact

with large displays and have competitions and cooperation, the applications can be more effective and attractive. So, we made Groupnect, a framework integrating group interaction into a large display system, which enables a dozen users' real-time interactions in a shared large display system simultaneously.

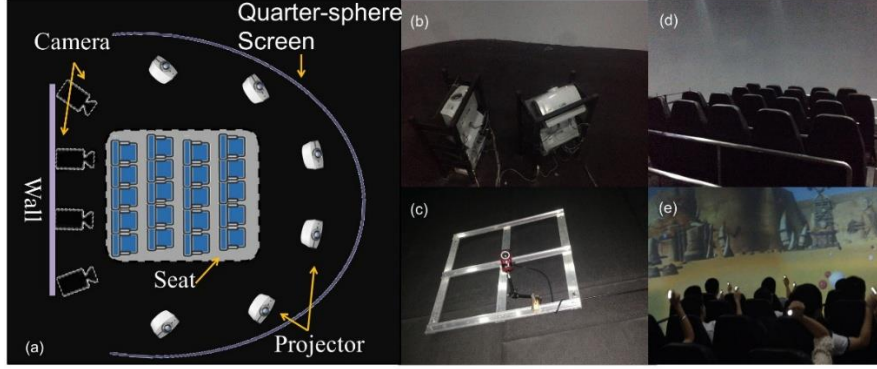


Figure 1. Overview of Groupnect. (a) Schematic diagram of the experimental environment. (b) Tracking cameras on the backwall. (c) A tracking camera. (d) The platform with 20 seats. (e) 20 people interact with the large display system.

3. The Groupnect Framework

As shown in Figure 2, the complete system consists of a group interaction and large display module. The main task of a group interaction module is the group users' hand gesture trajectory recognition. In the interaction module, the depth camera firstly collects the data based on computer vision techniques. Then, the origin data will be sent to the labeling and recognition part. When these procedures are finished, the users' commands will be sent to the back-end system in real time. The main task of the large display module provides vivid display effects. It consists of a parallel rendering module and real-time projection module.

3.1. Group Interaction

3.1.1. Tracking and Labeling

Group interaction means multi users, and to ensure the tracking accuracy, we chose the optical devices for data collection. The camera used in this framework is infrared optical camera, and the tracking region is a 3 meter long, 2.5 meters wide platform with 20 seats. To collect real-time and accurate data of the markers, the cameras are placed on specific locations with angles adjusted by the actual scene, making sure that the target recognition area is covered by two or more cameras. The four cameras share a same synchronizer to synchronous their data.

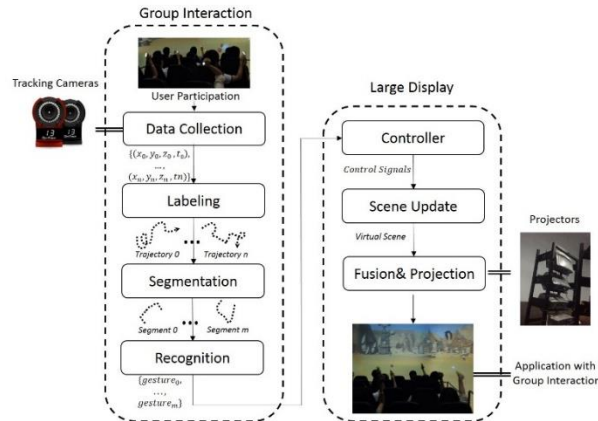


Figure 2. Overview of our Groupnect system, which consists of two stages: Group Interaction (for group users' gesture recognition) and Large Display (for providing immersive display effect).

As shown in Figure 3, the original data is $\{(x_1, y_1, z_1), \dots, (x_n, y_n, z_n)\}_t$, which represents the spatial coordinates of the different markers on frame t . To label the user data, we must classify and label markers data to a user in real time. We combine the temporal and spatial information of the origin data to build a loss function. Because of the fixed seats' position and limited users' activity area, the spatial loss can be calculated by the Euclidean distance between the markers data and seats. Because hand movement trajectory is continuous, the temporal loss can be calculated by the Euclidean distance between the markers data and the history data.

On frame t , for a single marker's data, we can label it by Equation (1).

$$C(X) = \operatorname{argmin} \left(C_s(X, S_n) + \omega C_t(X, H_n) \right) \quad (1)$$

In this equation, $C(X)$ represents the labeling result of marker data X , $C_s(X, S_n)$ represents the spatial cost of X with seat S_n , $C_t(X, H_n)$ represents the temporal cost of X with user history trajectory data H_n . and ω is the relaxation factor to balance the influence of spatial cost and temporal cost. It can be set up by prior knowledge and adjusted by the experiment results. Overall, this function combines the temporal and spatial information and performs well in practice.

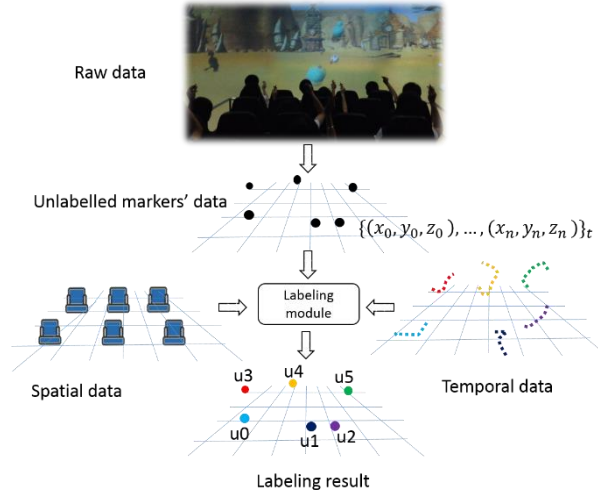


Figure 3. Main process of tracking and labelling. The raw data is a picture with markers in it. By using computer vision technology, the markers' position can be calculated but unlabelled. We can get the owner ID of them by combing spatial loss and temporal loss on the unlabelled marker data.

3.1.2. Gesture Spotting and Recognition

In practice, gestures need to be simultaneously detected and recognized from continuous movement data, which is commonly referred to as gesture spotting [21,25]. Furthermore, the recognition decision is made using data up to the current frame during online gesture spotting process. We cannot get future data in the current frame.

In our system, because the real time computing demand increased with more users, it is hard to realize real time gesture spotting and recognition. As Figure 4 shows, firstly we spot the trajectory based on the pause point. After the initial segmentation, we remove some of the track segments that are too long or too short, or completely contrary to human motion. Then, we can get 3D trajectory sequence for gesture recognition. There are many methods for 3D trajectory recognition (For example, HMMs [14] and CRF [25]). To keep recognition efficiency and accuracy, we use the Protractor recognizer [15]. Because the recognizer is only suitable for 2D trajectory sequences, we reduce the 3D trajectory dimension by the PCA method. After data reduction, we sample a points path into N evenly spaced points, rotate, scale and translate to reduce the spatial difference of the gestures. Finally, we get $(x'_1, y'_1), \dots, (x'_N, y'_N)_t$, the vector representing the gesture. When dealing with gesture template t and the unknown gesture g , we use the inverse cosine distance between their vectors.

We adopt multi-thread technology to accelerate the recognition process and the result will be sent to the back-end system in real time. In practice, our framework supports 9 types of gesture recognition, and users can train and add custom gestures by themselves.

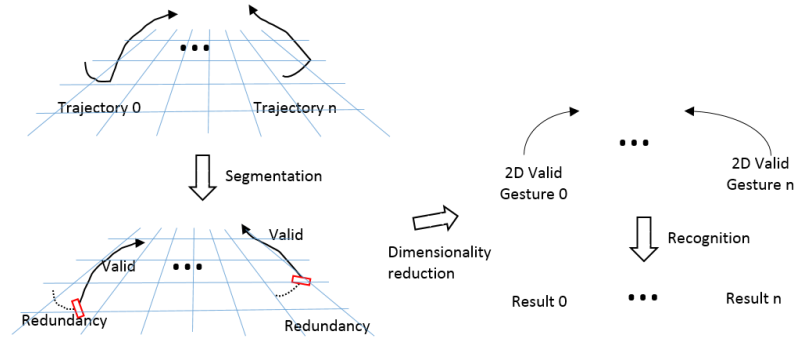


Figure 4. Main process of gesture segmentation and recognition. The original trajectories are continuous and contains redundancy gestures, we segment and reduce the dimension of it. Then use Protractor recognizer to recognize it.

3.2. The Large Display Module

The entire display system hardware includes 18 project projectors, a quarter-sphere screen and sever for rendering. The brightness of the projectors is 5,000 lumens, and the resolution is 1920X1200. The display contrast is 1000:2000. The projection screen is a quarter-sphere screen with radius of 4 meters. In addition, the rendering sever consists of 18 computers (Intel core i7 processor, NVidia GTX680 graphics display card, 8GB memory).

To achieve the super-resolution display effect and improve the sense of immersion, we use stitching and fusion technology. Firstly, the three-dimensional scene is allocated to different channels and each channel shows a part of the scene (The resolution of a single channel is a 1920 X 1200). By using the image stitching technology, we realize the seamless splicing of multiple channels. Meanwhile, to make the image display on the quarter-sphere screen display normally (does not deform, twist), we use coordinates graphics projection technology, as a two-dimensional planar image is projected on the cylindrical screen. In practice, much overlap of the projector does not match the brightness and the surrounding area. We use fusion technology to deal with this problem. We sample the overlap and calculate the contribution rate of each projector brightness and adjust the brightness automatically. By using those methods, we finally realize the seamless stitching and fusion and get an ideal display effect.

4. Pilot Study

4.1. Design Goals

To improve the physical and mental participation of users by group interaction, we design an experiment comparing standard interaction mode with group interaction mode.

4.2. Participants

Twelve able-bodied students between 20 and 25 years old participate in the experiment. Participants were asked not to consume alcohol or nicotine-based products 24h before testing and they were familiarized with the experiment procedures before testing.

4.3. Procedure

In the test “Hitting monsters”, users sit on seats and monsters walk or fly towards them in a distance. By waving hands, users can throw balls toward their waving direction in the virtual environment (users' balls are represented by distinct colors). The game has a score system, and users get points when their balls hit the monsters. All the user can see is their real-time score on the screen.

The experiment has two modes: single player mode and group player mode. Each mode has 6 participants. To avoid possible influence, the participants in these two modes are all different.

In the single player mode, six users participate in this test one by one. Before the experiment starts, the participants close their eyes and rest on seats for three minutes to collect their resting heart rate data. After that, the test starts and lasts six minutes without any other spectator. Users can see their real-time scores on the screen, and consequently the motivator is to get the high score.

In the group player experiment, six users are divided into two teams to experience the test. The team score is the sum of the individual scores. When the experiment starts, the participants also rest for three minutes, and they are forbidden to talk in this period. The test starts after the three minute rest, and users wave their hands to interact with the system. The test lasts six minutes and users can see their real-time score (including the team score and their own score) on the screen.

4.4. Collected Data

We collect objective data and subjective data in this experiment. Objective data includes motion frequency (MF, means the frequency of waving hands) and heart rate(HR), while subjective data is a questionnaire based on NASA-Task Load Index.

We can measure the user's sense of participation by using objective techniques. Objective techniques rely on quantitative measures based on performance or physiology, and generally are quite reliable. The major physiological measures are aerobic capacity, heart rate (HR), blood pressure, body temperature, electromyogram (EMG), muscle tension (strength), pupillary dilation, and speech analysis [7]. In this experiment, we chose heart rate (HR) and motion frequency (MF) as two factors.

Based on individuals' feelings and perceptions, we can also measure user's participation level by subjective techniques. The most common subjective techniques are: NASA-Task Load Index (TLX) [8], subjective workload assessment techniques (SWAT), and subjective rating of task difficulty (SRTD). The NASA-TLX is available in software form for use and users can easily accomplish it online. This index provides an overall subjective workload score based on a weighted average (WWL) of ratings on six dimensions: mental demands (MD), physical demands (PD), temporal demands (TD), own performance (OP), effort (EF), and frustration(FR).



Figure 5. The comparison of users' heart rate in single player mode and group player mode. The whole process takes nine minutes, the users rest for three minutes and play for six minutes.

5. Results

As the methods described above, we have collected all the users' heart rate data during both the resting and playing periods. Furthermore, during the playing part, the times the user hands wave were counted. After the whole experiment process, they all filled out the NASA-TLX forms. In the following sections, we will compare and evaluate the group interaction with the standard system (single player mode) by analyzing users' heart rate diagram, motion frequency diagram and data summary table.

The following Figure 5 is the users' heart rate comparison diagram in both single player mode and group player mode (we choose two representative users in the two modes respectively). The horizontal axis represents time, and the unit of it is in minutes. The process takes nine minutes, resting takes the first three minutes and the background is green, while the remaining

six minutes is game time and the background is red. The vertical axis represents heart rate, and the unit is in beats/min. The red line and the blue line represent the group player mode and the single player mode respectively.

As we can see from the diagram, in the process of resting, the players' average heart rate in single player mode is a little lower than group player mode (as calculated from the origin data, they are 79beats/min and 73beats/min). This circumstance is caused by the users' individual difference. Different users' resting heart rate is different because of their different gender, ages, physical quality et al. In the first minute after the game begins, users' heart rate in two modes have an increase tendency, but the group player mode increases sharply.

Therefore, at the beginning of the game, users' enthusiasm in group player mode is higher than single player mode. During the game period from the fourth minutes to the ninth minute, the users' average heart rate in group player mode is much higher than single player mode, and it stays high in group player mode with no decline tendency. But, the single player mode has the decline tendency after the eighth minute. Therefore, the users' participating enthusiasm in group player mode is obviously higher than in the single player mode. And through the entire process, the users' participating enthusiasm of the group player mode does not decline, which it is opposite in the single player mode.

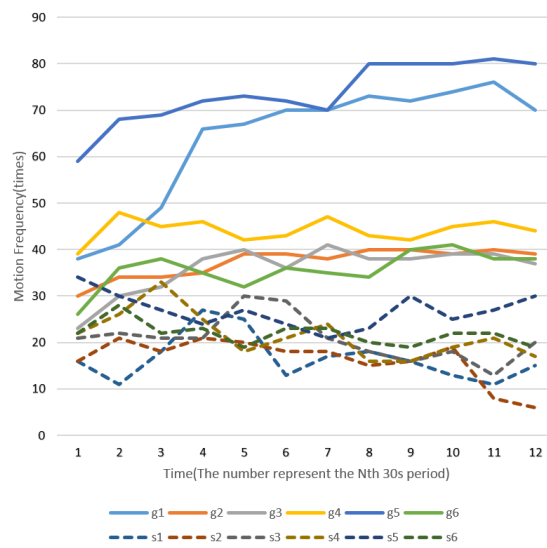


Figure 6. The motion frequency comparison of single player mode and the group player mode.

The following Figure 6 is a motion frequency comparison diagram in these two modes. The horizontal axis represents time, and the unit is in thirty seconds. The process takes six minutes. The vertical axis represents the motion frequency. It mainly records the users' waving hand times in the thirty second periods. The full line from g1 to g6 represents the six users' motion frequency in the group player mode, while the dotted line from s1 to s6 represents the six users' motion frequency in the single player mode.

Table 1. Summary of response variables means (min, max, std. dev) across the experiment, HR is short for heart rate and MF is short for motion frequency.

	Single player mode		Group player mode	
	Resting	Playing	Resting	Playing
HR(beats/min)	78.00(72.54,83.96,4.58)	85.65(78.27,88.90,4.25)	80.93(72.5585.55,4.79)	104.67(94.91,119.02,10.07)
MF(times/min)	0(0,0,0)	41.33(32.67,53.67,7.74)	0(0,0,0)	96.86(71.50,147.33,32.68)
NASA TLX				
TLX MD	0(0,0,0)	51.67(40,65,8.76)	0(0,0,0)	54.17(15,80,22.45)
TLX PD	0(0,0,0)	70.00(60,85,8.37)	0(0,0,0)	74.17(60,85,10.68)
TLX TD	0(0,0,0)	54.17(40,65,8.61)	0(0,0,0)	54.17(35,75,12.81)
TLX OP	0(0,0,0)	62.50(45,80,13.32)	0(0,0,0)	75.00(70,85,10.11)
TLX EF	0(0,0,0)	59.00(50,70,7.42)	0(0,0,0)	81.67(75,90,6.06)
TLX FR	0(0,0,0)	37.50(25,50,9.35)	0(0,0,0)	61.67(35,75,13.43)
TLX WWL	0(0,0,0)	58.17(55,62,67,2.91)	0(0,0,0)	69.78(61,79,5.60)

As we can see from the diagram, the maximum motion frequency in single player mode is not more than 40 times(s5-1), but the maximum motion frequency in group player mode is larger than 80(g5-11). The minimum motion frequency in single player mode is less than 10(s2-12), but the minimum motion frequency in group player mode is more than 20(g3-1). We can extract that the users' participating enthusiasm in group player mode is higher than that of single player mode.

Observing from the whole tendency, the motion frequency in group player mode presents a tendency to gradually increase at the beginning, and then become stable without any obvious decline tendency. But in the single player mode, the users' motion frequency has the tendency to decline. It is perhaps because when players play together, user competition improves their enthusiasm. On the contrary, users motion frequency may decline with less competition.

The following Table 1 is an overall information table including the NASA-TLX result. The table can be divided into two parts, which separately represents the single player mode and the group player mode. Every mode can be divided into resting and playing part. Above of the table and below of the table represent heart rate(HR), motion frequency(MF) and NASA-TLX respectively. In the table, every data represents the value in the present mode. The values in brackets represent the minimum, the maximum and the standard deviation of the data. For example, the left-top data 78.00(72.54, 83.96, 4.58) represents the six participants' resting heart rate and the average is 78beats/min, the minimum is 72.54 beats/min, the maximum is 83.96, and the standard deviation is 4.58.

From the data of motion frequency, in group player mode the average of users' motion frequency is 96.86 times/min, far higher than the average 41.33 times/min in single player mode. As we can see from the NASA-TLX, the average of the six values in group player mode is higher than that of single player mode. The difference of MD, PD, TD in the two modes is not obvious and the values of OP, EF, FR is higher in group than in single. Therefore, in the group player mode, competition and cooperation are fiercer. Users in this mode have a higher winning desire. Though the frustration in this mode is higher, the users are still more satisfied with their performances.

6. Conclusions

In this paper, we study the effects of group interaction in large display systems and propose a novel application framework called Groupnect. We design a game to compare the standard and group interaction mode. In the process of playing, the users have more body action and language event, and the users' heart rate is higher when in group interaction mode than the single player mode.

As for future work, we can use more complex models to improve the segmentation and recognition accuracy, and we shall use accelerated technologies such as CUDA to ensure the system can response in real time. In the future, we may drop the markers and directly recognize users' hands to make the interaction method more natural. And we may do experiments that more people participate in to test and optimize our system. We will also build more VR applications based on Groupnect and apply them in different fields.

Acknowledgements

This work was in part supported by the National Key Technology R&D Program of China (Grant No. 2015AA016405 and 2013AA013902), the Natural Science Foundation of China (Grant No.61100086 and 61272322), the Open Project Program of the Engineering Laboratory of Digital Preservation Technology of Traditional Settlements of Hunan Province (CT16K01) and the Key Laboratory of Digital Protective Technologies & Application of Cultural Heritages of Settlements of Hunan Province (JL16K02).

References

1. R. Aspin, "An Initial Study into Augmented Inward Looking Exploration and Navigation in Cave-Like Ipt Systems," *In Virtual Reality Conference, 2008. VR '08. IEEE*, pages 245–246, March 2008.
2. X. Cao and R. Balakrishnan, "Visionwand: Interaction Techniques for Large Displays Using a Passive Wand Tracked in 3d," *in Proceedings of the 16th annual ACM symposium on User interface software and technology*, pages 173–182. ACM, 2003.
3. O. Chapuis, A. Bezerianos, and S. Frantzeskakis, "Smarties: An Input System for Wall Display Development," *In Proceedings of International Conference on Human Factors in Computing Systems*, pages 2763–2772, 2014.
4. K. Cheng and K. Pulo, "Direct Interaction with Large-Scale Display Systems Using Infrared Laser Tracking Devices," *in Proceedings of the Asia-Pacific symposium on Information visualisation*, Volume 24, pages 67–74. Australian Computer Society, Inc., 2003.

5. C. Cruz-Neira, D. J. Sandin, T. A. DeFanti, R. V. Kenyon, J. C. Hart, "The Cave: Audio Visual Experience Automatic Virtual Environment," *Commun. ACM*, 35(6):64–72, June 1992.
6. A. Febretti, A. Nishimoto, V. Mateevitsi, L. Renambot, A. Johnson, and J. Leigh, "Omegalib: A Multi-View Application Framework for Hybrid Reality Display Environments," *In Virtual Reality (VR) IEEE*, pages 9–14, March 2014.
7. T. K. Fredericks, D. C. Sang, J. Hart, S. E. Butt, and A. Mital, "An Investigation of Myocardial Aerobic Capacity as a Measure of Both Physical and Cognitive Workloads," *International Journal of Industrial Ergonomics*, 35(12):10971107, 2005.
8. S. G. Hart and L. E. Staveland, "Development of Nasa-tlx (task load index): Results of Empirical and Theoretical Research," *Advances in Psychology*, page 139183, 1988.
9. B. Izatt, K. Scholberg, and R.P. McMahan. "Super-Kave, an Immersive Visualization Tool for Neutrino Physics," *In Virtual Reality (VR), 2013 IEEE*, pages 75–76, March 2013.
10. R. Jota, M. A. Nacenta, J. A. Jorge, S. Carpendale, and S. Greenberg, "A Comparison of Ray Pointing Techniques for Very Large Displays," *In Proceedings of Graphics Interface 2010*, pages 269–276. Canadian Information Processing Society, 2010.
11. J. Kim, D. Gracanin, K. Matkovic, and F. Quek, "Iphone/Ipod Touch as Input Devices for Navigation in Immersive Virtual Environments," *In Virtual Reality Conference, 2009. VR 2009. IEEE*, pages 261–262, March 2009.
12. R. Kopper, D. A. Bowman, M. G. Silva, and R.P. McMahan, "A Human Motor Behavior Model for Distal Pointing Tasks," *International Journal of human-computer studies*, 68(10):603–615, 2010.
13. J. J. Laviola, Jr. Daniel, F. Keefe, R.C. Zeleznik, D.A. Feliz, "Case Studies in Building Custom Input Devices for Virtual Environment Interaction," *VR Workshop Beyond Glove & Wand Based Interaction*, pages 67–71, 2004.
14. H. K. Lee and J. H. Kim, "An Hmm-Based Threshold Model Approach for Gesture Recognition," *IEEE Transactions on Pattern Analysis Machine Intelligence*, 21(10):961–973, 1999.
15. Y. Li, "Protractor: A Fast and Accurate Gesture Recognizer," *In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI '10*, pages 2169–2172, New York, NY, USA, 2010. ACM.
16. M. Nancel, O. Chapuis, E. Pietriga, X. D. Yang, P. P. Irani, M. Beaudouin-Lafon, "High-precision Pointing on Large Wall Displays Using Small Handheld Devices," *In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI 13*, pages 831–840, New York, NY, USA, 2013. ACM.
17. M. Nancel, J. Wagner, E. Pietriga, O. Chapuis, and W. Mackay, "Mid-air Pan-and-zoom on Wall-Sized Displays," *In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 177–186. ACM, 2011.
18. T. Ni, G.S. Schmidt, O.G. Staadt, M.A. Livingston, R. Ball, and R. May, "A Survey of Large High-Resolution Display Technologies, Techniques, and Applications," *In Virtual Reality Conference, 2006*, pages 223–236, March 2006.
19. M. Nunes, L. Nedel, and V. Roesler, "Motivating People to Perform Better in Exergames: Collaboration vs. Competition in Virtual Environments," *In Virtual Reality (VR), 2013 IEEE*, pages 115–116, March 2013.
20. R. Pausch, J. Snoddy, R. Taylor, S. Watson, E. Haseltine, "Disney's Aladdin: First Steps Toward Storytelling in Virtual Reality," *In Proceedings of the 23rd Annual Conference on Computer Graphics and Interactive Techniques, SIGGRAPH '96*, pages 193–203, New York, NY, USA, 1996. ACM.
21. B. Peng and G. Qian, "Online Gesture Spotting from Visual Hull Data," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 33(6):1175–1188, June 2011.
22. D. Roberts, A. S. Garcia, J. Dodiya, R. Wolff, A. J. Fairchild, and T. Fernando, "Collaborative Telepresence Workspaces for Space Operation and Science," *Institute of Electrical & Electronics Engineers*, 2015.
23. A. Simon and S. Scholz, "Multi-viewpoint Images for Multi-user Interaction," *In Virtual Reality, 2005. Proceedings. VR 2005. IEEE*, pages 107–113, March 2005.
24. D. Vogel and R. Balakrishnan, "Distant Freehand Pointing and Clicking on Very Large, High Resolution Displays," *In Proceedings of the 18th Annual ACM Symposium on User Interface Software and Technology, UIST '05*, pages 33–42, New York, NY, USA, 2005. ACM.
25. H. D. Yang, S. Sclaroff, and S. W. Lee, "Sign Language Spotting with a Threshold Model based on Conditional Random Fields," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 31(7):1264–1277, July 2009.