Exploring User Experience in "Blended Reality": Moving Interactions Beyond the Screen

David F. Huynh
MIT CSAIL, the Stata Center
32 Vassar St., 32-350
Cambridge, MA 02139, USA
dfhuynh@mit.edu

Yan Xu
Renmin Univ. of China
Zhixing 1, 1203#
Beijing 100872, P.R.China
vivi.xy@gmail.com

Shuo Wang
Microsoft Research Asia
5F, Beijing Sigma Center
No.49, Zhichun Road
Beijing 100080, P. R. China
shuowang@microsoft.com

CHI 2006, April 22–27, 2006, Montreal, Canada.
ACM 1-xxxxxxxxxxxxxxxxxx.

Abstract

In many video games, players map their physical actions to control their on-screen avatars. In our work, we removed this mapping by applying screen display as a "window" through which virtual objects enter the player's physical space, and the player interact with them directly without the mediation of an avatar. We define this interaction as "Blended Reality" (BR).
We designed, developed, and evaluated a BR game prototype called "Apple Yard." A camera was used to track the positions of the player's eyes and wand. The 3D game scene was rendered view-dependently to create the illusion of looking through a window. A user experiment conducted on the prototype showed BR's potential in camera-based entertainment.

Keywords

Blended reality, augmented reality, physical interaction, camera-based game.

ACM Classification Keywords

H.5.2 [User Interfaces]: Interaction styles, Prototyping, Evaluation/methodology.

Introduction

While physical interaction is being promoted by computer vision technology, tangible user interface and ubiquitous computing, most of the interactions still rely on mapping physical movements to on-screen avatars or mirrored images. For instance, in QuiQui's Giant Bounce game, children's gestures were amplified, extrapolated, or re-interpreted when applied on the virtual dragon [5]. Such translations from players' movements to the avatar's actions depend on the limited number of behavior patterns that the system previously learned. Another example is the arcade shooting game that support light guns. A player only has to aim at the pixels on the screen rendered as a target. The depth and direction information of both the gun and target are not considered in determining whether the virtual bullet hits the target.

While these mappings do have merits, our work explores the interaction that is more direct and natural. In this paper, we demonstrate a new concept of Blended Reality (BR) through a camera-based game prototype called Apple Yard (Figure 1). This prototype is designed for direct interactions between users and virtual objects. And the 3D game scene is rendered and projected onto the screen from users' perspectives. User experiment results on Apple Yard show that BR systems can provide enriched experiences and more favorable interactions, even though accuracy is lowered.

Blended Reality

Blended Reality deals with physical interaction between virtual objects and users, and it occurs in the realm where the physical and virtual worlds blend as one.

In Milgram's "virtuality continuum" spectrum [6], Blended Reality is closer to Augmented Reality (AR) than Virtual Environment (VE). While users are almost disconnected from the physical world they actually inhabit in VE [2], AR and BR are dominated by the real world. At present, most AR research is concerned with the use of live video imagery, which is digitally processed and "augmented" by the addition of computer-generated graphics. However, BR keeps the two worlds mostly apart, but augments a subspace of the physical world in which the two worlds are made to seemingly blend into one, allowing the user's physical actions directly influence the virtual world. Besides, AR usually requires heads-up displays, but BR creates the possibility for non-intrusive interaction.

Prototype Design: Apple yard -- a Camera-based Blended Reality Game system

In "Apple Yard" (Figure 1), virtual apples in a virtual yard "fly out" of the screen and the player is asked to hit them with a hand-held wand. Because we do not render the apples after they enter the physical space, the player must imagine the virtual apples' trajectories. If the apples are hit, they will fly back and are rendered as apple cores after they arrive in the virtual world behind the screen.

Figure 1. "Apple Yard" and its elements

Features of the system

Our prototype “Apple yard” differs from other AR systems, such as TouchSpace (which recaptures human touch and physical interaction as essential elements of game play [2]), AR bowling (which enhance game realism by an integrated real-time kinematics multi-body system simulation [4]). ”Apple Yard” has following unique features:

- The player directly interacts with virtual objects in the physical world;
- The display screen is rendered as the “window” that connects the physical and virtual world;
- Less unencumbered interaction.

System Overview

Our system was implemented in C++. Microsoft DirectShow was used to capture images from a camera

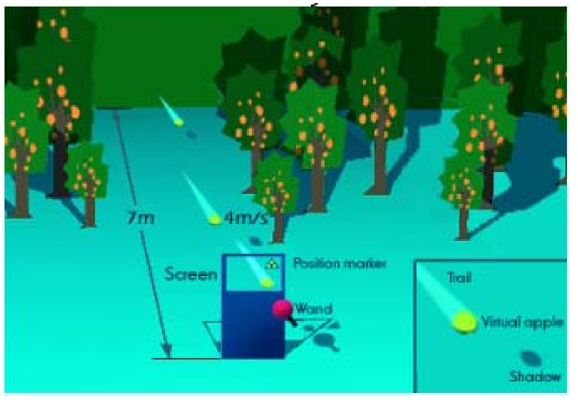

Figure 2. 3D model of Apple Yard

and Microsoft Direct3D [1] to render three-dimensional (3D) graphics. We used a wide-angled CCD security camera with 320 x 240 resolution, 24-bit color depth for high-fidelity image capture. Major aspects of the systems are as below:

- 3D game world

Figure 2 shows the 3D virtual world model, including a virtual apple yard, flying virtual apples, a window, and representations of users’ eyes and wand. The physical objects (screen, users’ wand and position marker) are modeled in the virtual world on a 1:1 scale. We choose to use a camera to track real-time 3D position information of the wand and users’ eyes because of its unencumbered nature [7].

- Rendering

To create the illusion of being inside the game scene for users, we locate the rendering camera of the virtual world exactly at the coordinates of users’ eyes. A user can view different aspects of a 3D object on a screen by moving his head, instead of controlling a mouse, keyboard or joystick. This matches the player’s everyday experiences of looking around their physical environment.

- Eye tracking

We track a position marker worn on the player’s forehead to calculate the approximate position of the player’s eyes. The marker is a triangle pattern painted blue and with three small green circles attached in three corners. The center of the triangle was located by hue analysis. Then, the centers of the three circles were located by k-means clustering algorithm. Finally, we used the largest pixel distance between the circles to infer z-displacement of the user’s eyes.

- Projection

In traditional first-person view games, players use either a mouse or joystick to control perspective change while their eye-positions remain stationary. However, in Apple Yard, players' eye-positions change when they move around. To adapt to the movement, we first formulate a standard perspective projection matrix such that the viewing frustum's front face covers the same area as the virtual window's opening. Then, we shift the viewpoint by multiplying the standard perspective projection matrix with a translation matrix that shifts projected images. So the 3D game world can be correctly projected considering the displacement of a player's eye-position to the center of the screen.

Our prototype adopted simple image analysis strategies to reduce computational cost. Our hue-based detection method trades off accuracy for speed. In effect, we applied a low-pass filter to dampen jitter at the cost of some lag in time

User Experiment

We targeted to find: (1) How do players perform in games of "interacting with virtual objects in the physical world;" (2) How does the rendering camera perspective influence user experience; and, (3) Can users readily accept interactions with invisible virtual objects in the physical world.

Participants

18 volunteers (1 female, 17 males) from our lab participated this experiment. They ranged in ages from 22 to 28, averaged 25, and were all right-handed. None had any prior experience with the system.

Experiment Design

We tested two factors:

- ***View-dependent/independent*** decided whether the virtual world images were rendered and projected on the screen according to a user's real-time eye position.
- ***First/third person view*** decided whether the users viewed the virtual world as being "inside" that realm.

Of the four combinations of the two factors, view-dependent and third-person view combination was invalid because when the players viewed the virtual world outside of it, their perspectives were fixed. The other three modes we tested included:

(a) **IT:** view-**I**ndependent & **T**hird-person;
(b) **IF:** view-**I**ndependent and **F**irst-person;
(c) **DF:** view-**D**ependent and **F**irst-person

Figure 3. Experiment setup of Apple Yard.
(a) **IT**, (b) **IF**, (c) **DF**,
Camera (1), Plasma display (112 x 63 cm^2) (2), Wand (3), Position marker (4), and participant (1.5 to 2 meters to the screen) (5).

The test followed a within-subjects design. Each of the participants finished three tasks, one task for each of the three modes (Figure 3). The order effect was counterbalanced by a 3 x 3 Latin square.

Quantitative Measures are: Wand scope and Accuracy. The coordinate of the wand was logged when it hit a virtual object (x, y, z). The scope of a user's wand is the Max(x)-Min(x), Max(y)-Min(y) and Max(z)-Min(z). Accuracy is the number of objects hit by the player's wand divided by the total number of objects. We also gathered qualitative information from a questionnaire, through interviews and observations, to analyze and understand users' behaviors.

Results

Three major results are reported below and qualitative reasons are presented:

(1) The majority of the participants understood and accepted the concept of "interacting with virtual object in physical world."

Through the interview, four of the users (22.22%) said "I understood the interaction happened in the real world space immediately I started playing the game"; 10 of them (55.56%) said that "I accepted this concept when learning to play"; four (22.22%) did NOT accept this concept at all.

In interviews, the large display (112cm x 63cm) and the high speed of flying objects (4 m/s) were two frequently mentioned parameters that enhanced immersion and helped to convince people of the BR concept.

(2) There is no correlation between the accuracy and the preference rank.

By correlation analysis, there is no correlation between the accuracy and the preference rank (p=0.384, two-tailed).

- Eight (44.44%), Five (22.78%), Five (22.78%) participants chose DF, IF, IT as their most preferable mode respectively.

Our interview also found that there were three major reasons why people preferred DF to IF: First, it was more similar to real-life experience; second, players enjoyed enriched interactivity brought by augmented ways of changing perspectives. Finally, player could more easily estimate trajectory of the virtual objects by naturally turning to better observation angles. Participants who preferred the IT mode more than the other two modes claimed that they preferred easier game play and would like the whole process of interaction visible by seeing the wand's avatar in the virtual world.

- According to ANOVA, the factor of first/third person view significantly influence the accuracy (F=14.168, p<0.001) (Accuracy: IT=52.63%, IF=37.23%, DF=29.84%)

The significant difference in accuracy was expectable because in IF and DF, users need to imagine the trajectory after virtual objects enter physical world and became invisible. Two major reasons summarized from participants' complaints about IF and DF were: (1) Estimations of object trajectories were different from actuality. (2) No hint was given when the user missed the objects.

(3) Players' physical movements were more intensive and diverse in DF than IT.

In DF mode, 3D movement scopes are significantly larger than IT (paired-samples T-test, two tailed x: t=4.832, p<0.001; y: t=6.225, p<0.001; z: t=3.576, t=0.002). Also, it was observed that players sweat much more in DF than in IT.

And it is observed that participants held more clam and unified gestures in IT mode. But in DF, they showed the diversity of playing, including jumping aside to hit the object, switching the wand to the other hand etc. And they differently use perspective change: some would like to front-face the objects, while some prefer bigger deflective angle.

Conclusions
As found in the user experiment, the rendering camera in 3D game world does significantly influence a player's experience. While the third-person view assures higher accuracy, first-person view could encourage more intense and diverse movements.

But the factor of view dependent/independent rendering and projection did not show significant advantages in either accuracy or movement scope. From the interview, we found it was the system's performance (jittery and delay) that harmed user experience with the view-dependent feature.

There are three major problems shown in the current prototype: (1) System Performance: jittery and delay; (2) Lack of feedback when players fail to hit the target; (3) Lack of stereopsis with 2D displays when players try to estimate the trajectory.

Summary and Future work

Besides solving the problems found in the user experiment, we plan to research: (1) how to adopt the concept of Blended Reality in multi-player video games and (2) how the other parameters, such as screen size, feedback, virtual objects' attributes influence user experience in BR systems.

We hope that "Blended Reality" can inspire research and practice community to incorporate more physical and natural interaction to video games. BR itself is one of these explorations. As proved by the user experiment, BR has great potential in camera-based entertainment with enriched interactivity. We also found that BR games can bring health benefits, which can be applied to physical exercise or sports training, especially in homely settings where physical space is limited.

Reference

1. Microsoft DirectX, http://www.microsoft.com/windows/directx/.
2. Adrian David Cheok, Xubo Yang, Zhou Zhi Ying, Mark Billinghurst, Hirokazu Kato, Touch-Space: Mixed Reality Game Space Based on Ubiquitous, Tangible, and Social Computing, *Personal and Ubiquitous Computing*(2002), Volume 6 Issue 5-6.
3. Bimber, O., Wetzstein, G., Emmerling, A., and Nitschke, C., Enabling View-Dependent Stereoscopic Projection in Real Environments, *Proc. ISMAR'05.*
4. Carsten Matysczok, Rafael Radkowski and Jan Berssenbruegge, "AR-Bowling: Immersive and Realistic Game Play in Real Environments Using Augmented Reality", *ACM SIGCHI, Con. A CET 2004.*
5. Hämäläinen, P., and Höysniemi, J. A, Computer Vision and Hearing Based User Interface for a Computer Game for Children. *Pro. ERCIM 2002*, 23-25.
6. Milgram, P. and Kishino, F., A Taxonomy of Mixed Reality Visual Displays, IEICE Transactions on Information Systems, vol. E77-D (12), Dec. 1994.
7. Pattie Maes, Trevor Darrell, Bruce Blumberg, Alex Pentland, The ALIVE system: wireless, full-body interaction with autonomous agents, MIT Media Laboratory, 20 Ames Street, Cambridge, MA 02139, USA, Multimedia Systems (1997) 5: 105–112.