Proceedings

3rd International Conference on Digital Interactive Media in Entertainment and Arts

DIMEA 2008

10-12 September 2008, Athens Information Technology (AIT), Athens, Greece

Jointly organized by:

In cooperation with:

Editors

Sofia Tsekeridou, Athens Information Technology
Aristodemos Pnevmatikakis, Athens Information Technology
Kevin Wong, Murdoch University
Thanassis Tiropanis, University of Southampton
Ryohei Nakatsu, National University of Singapore

The Association for Computing Machinery, Inc.
2 Penn Plaza, Suite 701
New York, NY 10121-0701

ACM ISBN: 978-1-60558-248-1

Printed in Athens, Greece.

Augmented Reality for Games

Filipe Costa Luz
movlab – University Lusófona
Av. do Campo Grande, 376
1749-024 Lisbon
+351 217 515 500 ext:2389
filipe.luz@ulusofona.pt

Vasco Bila
movlab – University Lusófona
Av. do Campo Grande, 376
1749-024 Lisbon
+351 217 515 500 ext:2389
vasco.bila@ulusofona.pt

José Maria Dinis
movlab – University Lusófona
Av. do Campo Grande, 376
1749-024 Lisbon
+351 217 515 500 ext:2389
jose.dinis@ulusofona.pt

ABSTRACT

In this text, we intend to explore augmented reality as a means to visualise interactive communication projects. With ARToolkit, Virtools and 3ds Max applications, we aim to show how to create a portable interactive platform that resorts to the environment and markers for constructing the game's scenario.

We plan to show that the realism of simulation, together with the merger of artificial objects with the real world, can generate interactive empathy between players and their avatars.

Categories and Subject Descriptors

H.5.2. [**Information Interfaces and Presentation**]: User Interfaces; I.3.6 [Methodology and Techniques]: Interaction Techniques

General Terms

Experimentation

Keywords

Augmented Reality, motion capture, 3d animation, Building Blocks

1. INTRODUCTION

The possibility of simultaneously inhabiting two spaces, the artificial and the real, to communicate both here and there at the same time, aggressively expands our power of communication.

The "interface culture" that Steven Johnson [1] proposes is precisely an attempt to understand the very individual that today we feel is diluted along very diffused real and virtual borders. Many of the conventional functions of the human eye are being replaced by techniques where images do not position themselves in the traditional manner that we observe them [2], or in the way we perceive the real world. The digitalization of the real world to a new informational layer over objects, people or environments, needs to be processed and mediated by tools that amplify the natural human senses.

At the outset of modern civilization, writing (perhaps one of the first technologies invented by humans) extended the human race's logical thinking to clay plates or papyrus. Likewise, drawing was represented in several supports in a clear expansion of the territories where we express ourselves.

On the other hand, the drawing experience through a dark chamber is an individualisation's process, in the sense that the observer is isolated from the outside world, enclosed in a dark space, with the projection of a single piece of the outside world [2]. The relationship with technique is thus much more intimate, because our natural senses in this example are enhanced by the "machine".

If the realism of Jan Van Eyck's paintings was achieved with the help of a dark chamber, at present the realism of movements represented in video animation projects is obtained through rotoscopy or motion capture (mocap) techniques. We are able to expand our senses through corporate work with the machine, as in the example of a mocap system resource, "we can see" up to 470 frames per second.

Thus, the new tools enhance human senses and can project the interaction space onto a new plane. Like an architect who imagines a dwelling on a sheet of paper, virtualising the spaces to be built using drawing techniques, the "new architect" can enter the drawing through a virtual reality system, and using sensorially more sophisticated media, he/she can interact with digital objects that virtualise real objects. However, the human-machine interaction in a VR system is still not very transparent, which makes it difficult to feel one's presence in the digital space. Curiously, in Ivan Sutherland's *Ultimate Display* project, this author called his helmet "Democles sword", since it seemed like the user's decapitation was imminent. If on one hand, the possibility of interacting and moving in space caused feelings of immersion, on the other, this helmet's excess hardware and complex connections caused great opacity in this human-machine relationship that intended to be as transparent as possible.

In the film, *Strange Days* (Kathryn Bigelow, 1995), the users of *Wire* used a much more ergonomic helmet than current ones; however, it is Cronenberg, (e*XistenZ*, 1999) who gives us an indication of an organic connection to a new world that is indistinguishable from the original.

In *Strange Days,* Lanny Nero, the actor Ralph Finnes, is a dream salesman. Through virtual reality (VR) technology, his clientes could feel sensations of other people through a virtual helmet that connected them to the network (*wire).* Subjective planes introduce the spectator to the VR simulation proposed by the film. On the other hand, Cronenberg distances himself from this thriller (ou science fiction film) and proposes a much more interesting problem: «Are we still in the game?»

eXistenZ, besides showing what VR and computer games could become in the future, is a paradigmatic film in that it presents an issue with transparency. The game "*eXistenZ"* presented by Jennifer Jason Leigh, as Allegra Geller, proposes experiences where it is not possible to distinguish between the real and artificial. This illusion is so complete that Cronenberg projects the same sense of disorientation to the spectator. After the film is over, we don't understand when the game started, nor if it ever finished.

With current technology, it is still difficult to have the total transparency sought by VR enthusiasts, because virtual worlds are meant to be immersive spaces that assume total interaction transparency, which means that these universes should be "natural" where the only interface is one's own body.

VR is an interface that tries to connect all human senses to a single totally transparent media where elimination of the mediation is understood, that is, the contradiction Grusin and Bolter defined as the *Double Logic of Remediation* [3].

On our work it's not a goal to follow in conducting a historical study since Ultimate Display, or Morton Heilig's Sensorama (1962) to the modern caves, but rather to reinforce the idea of the computer's dominating power.

Experiences in the act of playing [4] help to understand this strong attraction that computers possess. While traditional television "offers content", computer games provide a space to interact and modify. There is a type of "transformation" in the fact we emerge in the space of the game, because anxiety and pleasure are fed "from behind" the screen, in small interactive worlds through a reflex: - the avatar.

On avatars, we see ourselves mirrored in unexpected or desired actions. Perhaps that is where the name ID came from for the company that develops "bloody games", like Doom. We visualize our actions through cameras in the first or third person, which generates a greater or lesser approximation to the space. In order for us to better visualize what is about to happen in a certain game, we switch to the aerial camera (God View), thus expanding our view of space. As we defended in a previous study, the chosen point of view can determine the feeling of presence in space, action (event) or story being told [5].

In cinema as well as in computer games, the point of view is the user's main connection to the space being represented. The manner in which we perceive this space is a limited way to better understand it. Several perspectives permit the multi-angle perception of space, which means that general planes reveal more information, whereas tighter planes can provide greater detail of part of this information. Thus, the chosen point of view is both limiting and broadening at the same time, to the point where programmers, or designers, define the angle of view to cause greater levels of immersion for users. In order to generate greater spectacle while enjoying a narrative or game, the point of view is radically altered so the objects represented can be better transmitted. In Grand Theft Auto, as we drive a car at great speed over a ramp, the point of view changes to a cinematographic representation of the leap moment and the image is shown in slow motion. This cinematic effect has the player pull away from the space, by the subjective plane or in the third person of the car being driven, and substituted by another of contemplation, a general plane.

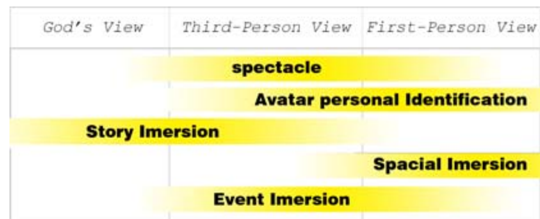

Figure 1. Camera angles and different senses of presence

In the augmented reality project we present, the point of view is placed on the world being represented. Since the webcam being used is supported by a tripod, our angle of vision is a God-View like point of view from above, because we intend to use the application for an audience and not for a single individual's contemplation.

The use of HMD (Head-mounted display) helmets results in using technology of little interest due to the excessive weight and cost of the available hardware, which is why it was not considered from the outset. Through video projections, which we have been accustomed to for so long now, innovations in AR permit new interfaces for communication that are much more ergonomic than HMD. Likewise, we know that the forefathers of Augmented Reality (AR) foresaw communication possibilities, many of which are represented in the film Minority Report (Steven Spielberg, 2002), which permit merging real with artificial images in an entirely new way. However, in this study, we intend to show how augmented reality projects can function as playful applications with a pedagogical nature or not, but that promote the removal of barriers between real and artificial objects, enchanting audiences and encouraging interaction with the system.

2. REALISM OF THE APPLICATION

The objective of this application is to demonstrate the potential use of AR for visualising human movement. The project consists of the digital representation of a track and field track and a 3D character (graphic simulation of an athlete). We intend to project the movements of an athlete set to begin a 400m race over real space using AR techniques. Thus, we determined the need to capture the movements of a professional athlete over motion capture, apply the movements to an avatar using 3D animation applications (3ds Max version 9), integrate the objects in real time (we used Virtools) and integrate the graphic processing to real video (ARToolKit).

We sought the realism of the simulation by capturing athlete movements at the Digital Animation and Biomechanics Laboratory for human movement (movlab) at the Lusófona University, and for avatar displacement we sought a formula to make the representation of the animations random.

At movlab, we used 8 infrared cameras that capture up to 470 frames per second with a resolution of 1.3 mpixels (black and white images up to 1280x960). We used 14mm reflection markers (tied to a training suit and directly on the athlete's skin) and we captured the following movements: - 10 walking movements, 10 running, 10 of the walking-running transition, 10 of the running-stopping transition, 10 jumps over a 55cm high obstacle and 10 poses at rest (standing).

In the Vicon system, we used the Vicon IQ software to filter the information, and with 3ds Max software, we created 4 footsteps cycles because due to the limitations of graphic processing in real time, in general, animations are shortened, so that, in cycles, they simulate displacement with greater visual realism and less computer processing.

Figure 2. Motion Capture Studio (movlab - Universidade Lusófona)

As we know, a person does not always move the same way. Depending on the surface friction, inclination or each person's physical state, we adapt our daily displacements to the environment. The perception of the surrounding environment results in individual interpretations of each person's experience with the stimulations provided by the environment. Thus, a place is defined by possible interaction and also by the capacity it has to transform us into an inhabitant of this space: - the world as a series of interconnected objects and individuals [6,7]. We do not intend to explore the differences between space and place here, but we accept the study by Yi-Fu Tuan [8] or the definition that space is the place of action, that is, the experience of the place offers the consistency of the world [9]. Thus, if in the interaction of an avatar with the environment, we recognize daily tasks, the realism of movement is absorbed by the spectators.

From this idea, we understand that the cyclical movement of an avatar is not natural and in order to hide the nature of the programming we try to mix the greatest number of possible movements to camouflage this technical requirement.

This is a concern in computer games, so that through artificial intelligence, they can prepare the agents for the interactions that will occur in the digital space. The interactivity that occurs between agents results from the analysis we make of the real world, thus, the relation avatars have with the environment is of utmost importance to perceptively understand human movement [10]. Therefore, we add objects that can present themselves as obstacles for the characters' movement.

The captures were edited using Vicon IQ, trying to eliminate all capture noise, and using Character Studio from 3ds Max software, we created the different cycles For example, for walking, the capture presented a 10 footsteps movement. We selected a 4 footsteps cycle where it was necessary to correct the initial and final position of the hands and head so the cycle would not be perceived during repetition.

Figure 3. Interacting with the application

We validated the animations generated by showing them to a group of twenty-five students from Cinema and Multimedia course (Lusófona University) who were between 21 and 25 years of age. They were told the walking movement had been captured in mocap studio, at movlab, by an athlete that runs over 16 footsteps. What really happened was a capture of only 4 footsteps, and later, it was multiplied on 3ds max (motion mixer) to produce a single animation of 16 footsteps.

When the students were interacting with the system, all of them accept the animation as a cycle of 16 footsteps, which legitimizes the realism of the real cycle being used (only 4 footsteps).

The motion capture systems allow us to capture real movements, so we can se in this application a avatar running like us. The avatar is not absorbed as real, cause is a simple graphic project on a laptop monitor, but graphics are not important for a application looks real, what matter is if the action of user is real. If the user moves the markers and the avatar follow the user instruction without delays, the agency of the application will be very strong [11]. The notion of realism resides in application's tactility and the player's experience. The realism of the application is not understood in the sense of the graphical representation on screen, but rather on pleasure of the user participation.

This is an important characterization on video games – the realism of action. First was the realism in literature (narratives), after on produced images (Mirrors, paintings, photography and movies), now we live the third order of realism – the realism of action [12]

So, to achieve a nice game play in this application, it was developed a basic avatar with a simple texture, but we oriented our efforts on animation post production (looking for real cycle movements), interface user friendly (markers with instrutions) and flow interaction (at least 70 frames per second). We were very concerned if the interaction with markers could be similar to a interaction with real objects. The main goal was to guarantee realism in interaction with a avatar like we play with a ant, trying to block her to follow is path. Every time the user put an obstacle in front of the avatar it should stop and try to contour the obstruction.

3. DEVELOPMENT OF THE APPLICATION

The application aims to contextualise a space for interaction between people and artificial beings that cohabitate the same place. As we mentioned earlier, we sought a level of realism in the movements of a 3D character so that, through the similarity of the movements, some level of empathy/charm could be generated by interacting with a being that is somehow similar.

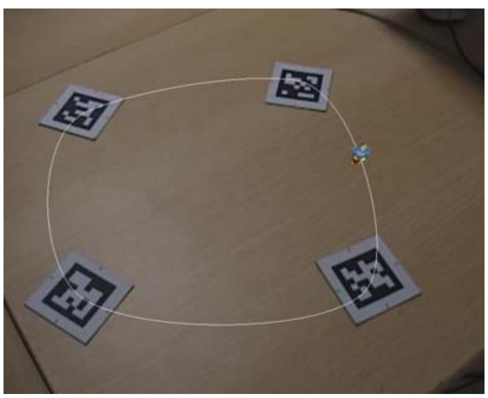

Figure 4. Path defining markers

We used Virtools software to integrate the 3D character (modelled and animated with 3ds Max) and to program its movements. Through ARToolKit Plus plug-in libraries, developed by Graz University of Technology (Austria) and the use of a Microsoft XBox 360 Live webcam, we created the application based on augmented reality.

This library (ARToolkitPlus) is an extension of the ARToolkit library, originally developed by Hirokazu Kato (Hiroshima City University) and continued by Human Interface Technology Laboratory (HIT LAB), at the University of Washington. ARToolkit uses C and C++ languages to permit the detection of markers (patterns) and the identification of their position, angle or direction, through video. The captured data are later transferred in 3D spatial references, where tracking between the real and the digital worlds is possible. For such to happen, this library resorts to graphic computation algorithms to calculate the position and direction of the real camera in relation to the marker.

In the next phase, it is possible to generate graphs on the marker's position and direction.

In the application developed, the computation process follows these steps:

1. Video captured by a webcam is analyzed in real time.

2. The software seeks patterns, that is, the identification of markers, frame by frame.

3. When the previous point has been validated, a trajectory is drawn (track and field track) over markers 1-4.

For this to happen, it was necessary to develop the individual detection of each marker. When validated, a curve is created comprised of four points, which correspond to each of the four markers.

Tracking of the real perspective with the digital is the biggest problem in combining the two contents in an AR system [13]. Alignment of the artificial camera must be fully synchronized to the movements of the real camera, and in order to solve this we used the curve markers for the respective tracking.

Due to image processing and the respective precise values, the decimal values (float) for the position and direction are alternatively unstable and regularly conflict with the curve. In order to solve this problem, a Building Block (AR Filter) was created in VSL language to be used as a filter with the position and direction values to make them stable.

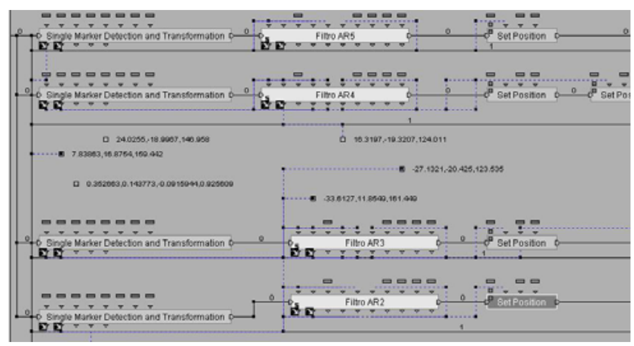

Figure 5. Virtools Building Blocks created with Virtools Script Language (VSL)

The filter detects the variation of input values, and if it detects a transformation (translation, rotation and scale) greater than what we consider noise, it is altered: - new values (output) are used, making it possible to alter the design of the trajectory when we move the markers.

An occlusion function was also created in the filter in order to maintain the current positions whenever a marker is no longer visible. This is an essential function for transparency in system mediation because whenever we move markers or block the view of the camera by moving our hands, the system remains calibrated thus avoiding delays in graphic processing or trajectory deconstruction.

4. When the curve was created, consonant with the position of the markers, we programmed such that the character followed the direction of the curve. Thus, the 3D character is placed on marker 1 (processed in render) and the randomly scheduled pose animations immediately begin.

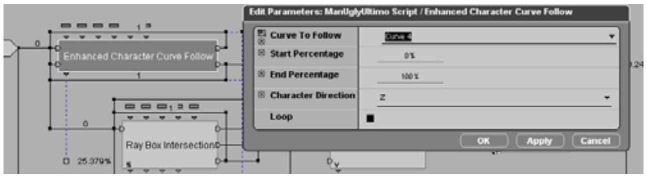

Figure 6. Assign Character movement to curve 2D

A function was also created that generates the average of the direction angles for the diverse markers. This average is associated with the character's direction in order to obtain a more precise direction over the space.

5. When we add markers 5 and 6, images of obstacles are processed graphically.

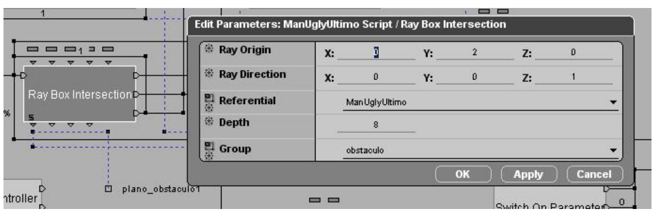

Figure 7. Defining collisions between 3d character and digital objects

6. If markers 5 and 6 are placed over the alignments for the trajectory markers, collisions will be processed between the character and the obstacles. The obstacles are divided into two types: those that cannot be hurdled (*muppi*) and those that can (placard).

With the obstacle that can be hurdled, the avatar will jump and hurdle without any difficulty, however, since the *muppi* cannot be hurdled, the avatar will need to look for the closest path to the trajectory's next marker.

This small example of artificial intelligence is initially processed by video analysis as a sensorial component, such as proximity sensors used in robots.

In this function, the origin of the "sensorial axis" is defined as well as its direction in relation to the pivot of the character. The sensorial ray (depth) dimension is also defined as well as the group that contains all of the detectable objects. When there is an object (detectable) that is closer than the ray, the function returns the identification of this object. The conditions that activate the actions the 3D character should take are then processed from this output.

4. CONCLUSION

The presentation of this project always generates a special empathy in the audience due to the strangeness of interacting with 3D entities in real time and through simple movements of our hands over the project markers. By interacting with the 3D character that runs from marker to marker, positioning the obstacles in front of it, we play and observe its reactions. When the same students who validated the animations generated by the motion-capture system in the initial phase of this project saw the final application, most of them asked (78%) to interact with the character.

The enchantment that these beings evoke in people place these types of applications in more organic spaces and modes of human-machine interaction, elevating the level of presence in mixed reality spaces. In a perceptual point a view, borders between real and artificial could be dissolved with the advent of artificial intelligence, graphic processing and the reduction of the processor sizes.

The future becomes progressively more hybrid, where digital artefacts will be mixed into real world, allowing us to communicate in broad band. In order to explain to an animation film student how human movement happens in everyday life, we think this application provides a glimpse at the possibility of realistically showing movements of different characters captured through mocap that interpret and react to the surrounding environment. By developing the databases for this application, we came up with a fun application that is pedagogical at the same time.

5. ACKNOWLEDGMENTS

We would like to thank Professor João MCS Abrantes and Ivo Roupa for the support with motion capture technology and biomechanic's studies.

To *Fundação para a Ciência e Tecnologia* by supporting Project INFOMEDIA – Information Acquisition in New Media (PTDC/CCI/74114/2006).

6. REFERENCES

[1] JOHNSON, Steven. 1997. Interface Culture: How new technology transforms the way we create and communicate. 2ed. New York. Basic Books.

[2] CRARY, J. 1992. Techniques of the Observer: On vision and modernitythe nineteenth century. Cambridge. MIT Press.

[3] BOLTER, D. e GRUSIN, R. 2000. Remediation: Understanding New media. Cambridge. MIT Press.

[4] TURKLE, S. 1989. O Segundo Eu, os Computadores e o Espírito Humano. Lisboa. Presença.

[5] LUZ, F. 2006. Mediação Digital como Jogo: Transparência e Imersão. Masters thesis. Universidade Nova de Lisboa,

[6] HALL, Edward. 1966. The Hidden Dimension. New York. Doubleday.

[7] RYAN, M. 2001. Narrative as Virtual Reality, Baltimore, John Hopkins.

[8] TUAN, Y. 1976. Space and Place: The perspective of Experience. Minneapolis. University of Minnesota Press.

[9] MCCULLOUGH, M. 2004. Digital Ground: architecture, pervasive computing, and environmental knowing. Cambridge. MIT Press.

[10] LEE, W., CHAI, J., REITSMA, P., HODGINS, J., POLLARD, N. 2002. Interactive Controls of Avatars Animated with Human Motion Data. Special issue: Proceedings of ACM SIGGRAPH 2002. PP. 491-500.

[11] MURRAY, J. 1997. Hamlet on the Holodeck - The Future of narrative Cyberspace. New York. The Free Press.

[12] GALLOWAY, A. 2006. gamming Essays on Algoritmic Culture, Electronic Mediations Séries, University of Minnesota Press, Minneapolis.

[13] BEHRINGER, R., KLINKER, G., MIZELL, D. 1999. Augmented Reality: Placing artificial objects in real scenes, proceedings fo IWAR’98, Natick, A K Peters