Augmented Mirror: Interactive Augmented Reality System based on Kinect.

Lucía Vera,Jesús Gimeno,Inmaculada Coma,Marcos Fernández
Instituto de Robótica. Universitat de València.
{Lucia.Vera,Jesus.Gimeno,Inmaculada.Coma,Marcos.Fernandez}@uv.es

Abstract. In this paper we present a virtual character controlled by an actor in real time, who talks with an audience through an augmented mirror. The application, which integrates video images, the avatar and other virtual objects within an Augmented Reality system, has been implemented using a mixture of technologies: two kinect systems for motion capture, depth map and real images, a gyroscope to detect head movements, and control algorithms to manage avatar emotions.

Keywords: Augmented Reality, Motion Capture, Virtual Characters.

1   Introduction

In the field of HCI, combining the real world with virtual information and introducing virtual characters interacting with a real audience are elements that add richness and complexity to presentations and public events.

In this context we have developed an Augmented Reality (AR) application that allows the audience to interact and talk with a virtual character through a large screen called augmented mirror. The avatar movements are controlled by an actor in real-time using a Motion Capture (MoCap) system. That avatar, together with video images from the real world and other virtual objects, are integrated in a system that also incorporates the voice of the character and the manipulation of virtual objects.

In a review of the literature we find different MoCap systems used for the animation of virtual characters. The MoCap devices use different technologies to acquire positions. Thus, in Waldo C., a virtual puppet, a *mechanical system* was used to track body joint angles using an exo-skeleton structure [1]. Other systems use *electromagnetic sensors* [2], or *optical systems* with markers and cameras. Finally, some use optical systems without markers using an image recognition algorithm to detect positions and movements [3].

Within AR applications, MoCap devices are also used to control the interaction between the user and the virtual objects. Sometimes optical systems with markers [4] are used to detect positions but, in other cases, more complex systems are utilized that are based on ultrasounds or inertial sensors that detects movements in a wide range of use [5, 6].

Aiming to achieve a system that captures the actor's movement in terms of limited space and lighting conditions, including facial expressions and lip movements, we

used a combination of technologies: a Kinect device for MoCap, a gyroscope for head movements, and a WiiMote with control algorithms for facial expressions. Moreover, it has been necessary to create an augmented video image blending the avatar and the real environment, using another Kinect depth map for that purpose. However, Kinect presents some restrictions that we have solved with different solutions described in the next section.

2   System description

The proposed system is composed of two main parts, called scenarios (see Figure 1). In the first one, the control scenario, the actor tracking is performed using a MoCap system. In the second one, the augmented scenario, the audience can interact with the avatar and other virtual objects. These two scenarios are connected using a Client-Server model, via an Ethernet connection, so that the avatar can be remotely controlled.

With regards to the control scenario, a number of input devices are combined into an enhanced MoCap system (see Figure 2) which serves the actor to control the avatar movements, facial expressions, lips movement and also some predefined virtual objects. Moreover, input data from these devices is received and processed by a server application to control the augmented scenario (the avatar and the virtual objects).

The main MoCap device is a Kinect camera that controls the avatar movement. By using OpenNI (Open Natural Interface) SDK, Kinect performs actor tracking in real time, with an 11 bones avatar structure (see Figure 1). In our system, this control was not enough to obtain a realistic avatar and, therefore the tracking was enhanced in 4 different ways: head orientation, lips movement while talking, facial expressions and automatic gestures (blink, hands, feet, etc.).
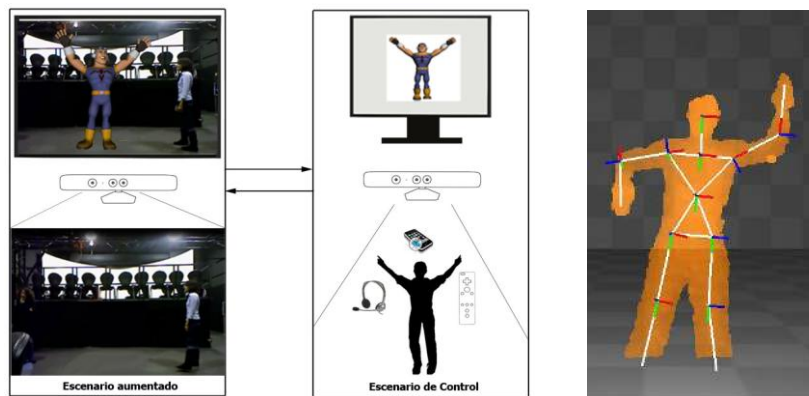
Fig. 1. General system structure (left); Kinect avatar skeleton (right).

Head orientation is estimated using a mobile phone attached to a cap. This can be achieved with an android application that computes heading, pitch and roll from the phone sensors data (accelerometer and compass) and send it to the server application

through a Wi-Fi connection. This global information, referenced to the earth gravity and North Pole, is computed to calculate head orientation according to the actor orientation. For lips movement while talking with the audience, the actor is provided with a wireless microphone. An amplitude based algorithm has been developed to calculate, lips aperture values (horizontal and vertical), which simulates the real mouth movement while talking. A WiiMote is used to control avatar facial expressions and 3D virtual objects. The actor can easily choose the avatar facial expression, using the WiiMote buttons, from five different predefined emotions. In the same way, 5 virtual animated objects can be shown or hidden. Finally, other avatar movements are computed automatically. Blinking velocity is controlled depending on the selected emotion; hand and finger movements are calculated from the arm movement. All this information is sent to the augmented scenario in real time, making possible a live interaction between the real actor and an audience.
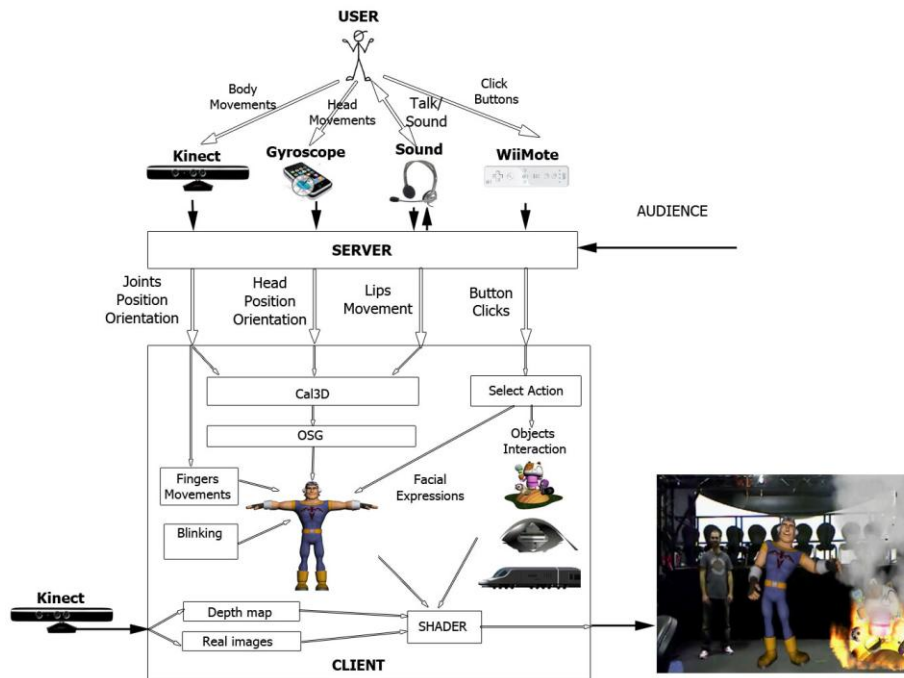

Fig. 2. Scheme of the overall application.

The augmented scenario is where the audience interacts with the avatar and the virtual objects via the augmented mirror. The audience can talk with the avatar or walk around him and the virtual objects while they can see themselves in a large LED screen (Figure 4). An important element of our augmented visualization is the correct occlusion between real and virtual information. This augmented images displayed on a 4x3 meter LED screen provide a more immersive multiuser augmented reality

experience. A Kinect camera is used to capture the real image of the audience which is used in both scenarios. In the control scenario the real actor can see the audience to interact with them, while this image is merged with the avatar and the virtual elements to generate the augmented mirror image. The occlusion handling is implemented on a GPU shader, which takes the real image and the depth information captured by Kinect, mixing it with the virtual scene. The implementation of the augmented visualization is based on OpenSceneGraph, Cal3D and our own developed libraries. The avatar skeleton (see Figure 3), has been designed to perform all the gestures and facial expressions desired.

Fig. 3. Augmented mirror (left); avatar skeleton (right).

3   Conclusions

This paper has presented an augmented mirror where a virtual actor interacts with an audience. Kinect motion capture capabilities have been used and enhanced with other 3 input devices to perform a MoCap system which allows the actor to control the avatar movements, facial expressions, lips movement and also some predefined virtual objects. The system has been successfully used in a real marketing presentation involving touristic contents (FITUR'11), where nearly 100 person enjoyed the augmented mirror experience.

4   Bibliografía

1. Walters,G. The story of Waldo C. Graphic. Course Notes: 3D Character Animation by Computer, ACM SIGGRAPH '89, Boston, (1989)
2. Huang, Y. Kallman, M. Interactive Demonstration of Pointing Gestures for Virtual Trainers. LNCS 5611. pp178-187. (2009)
3. Wren,R., Azarbayejani, A., Darrell,T. and Pentland, A. P. . Pfinder: real-time tracking of the human body, Trans. Pattern Anal. Mach. Intelligence 19(7), 780–785 (1997)
4. Dorfmüller, K. Robust tracking for augmented reality using retroreflective markers. Computer & Graphics. Volume, 23, 6, 795-800 (1999).
5. Foxlin, E., Harrington, M., Pfeifer, G. Constellation: A Wide-Range Wireless Tracking System for Augmented Reality and virtual set applications. Proc. of SIGGRAPH'98. (1998).
6. Vlasic D. et al. Practical motion capture in everyday surroundings. ACM Transactions on Graphics, Vol. 26,3.(2007)

Augmented Mirror: Interactive Augmented Reality System based on Kinect

Autors: Lucia Vera, Jesús Gimeno, Inmaculada Coma, Marcos Fernández
IRTIC Institute, University of Valencia

IRTIC [♀-☐]

Institut de Robòtica i de Tecnologies
de la Informació i les Comunicacions

VNIVERSITAT
ID VALÈNCIA

Introduction and goals

- We have developed an Augmented Reality application that allows the audience to interact and talk with a virtual character through a large screen called **augmented mirror**.

- The avatar movements are controlled by an actor in real-time using a Motion Capture (MoCap) system.

- The avatar, video images and other virtual objects, are integrated in a system that also incorporates the voice of the character and the manipulation of virtual objects.

System description

- The system is composed of two main parts: the **control scenario**, where the actor tracking is performed using a MoCap system; the **augmented scenario**, where the audience can interact with the avatar and other virtual objects.

- The two **scenarios are connected using a Client-Server model**, via an Ethernet connection, so that the avatar can be remotely controlled.
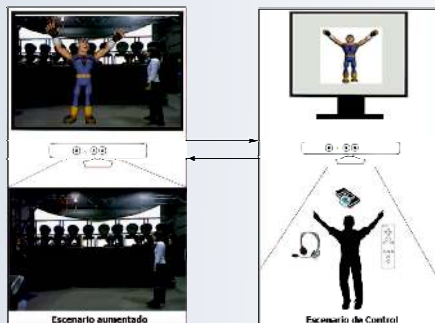

Augmented scenario

- The augmented scenario is where the audience interacts with the avatar and the virtual objects via the augmented mirror.

- A Kinect camera is used to capture the real image of the audience .

- The **occlusion handling** is implemented on a **GPU shader**, which takes the real image and the **depth information** captured by Kinect, mixing it with the virtual scene.

- The **avatar skeleton** has been designed to perform all the gestures and facial expressions desired using **OpenSceneGraph and Cal3D**.
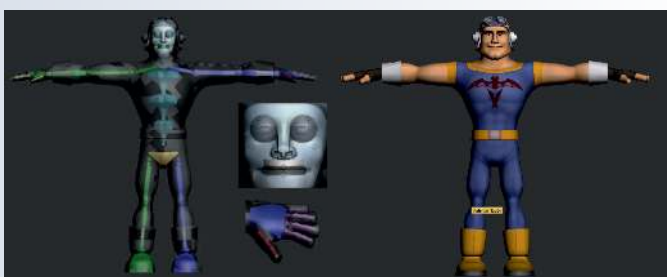
Control scenario

- The main MoCap device is a **Kinect camera** that controls the avatar movement by using OpenNI (Open Natural Interface) SDK.

- **Head orientation** is estimated using **an accelerometer and a compass** of an Android mobile phone attached to a cap.

- For **lips movement** the actor is provided with a wireless microphone. An **amplitude based algorithm** has been developed to calculate lips aperture values (horizontal and vertical).

- A **WiiMote** is used to control avatar facial expressions and 3D virtual objects. The actor can choose from **five different predefined emotions** and five virtual animated objects.

- Other **avatar movements are computed automatically**. Blinking velocity is controlled depending on the selected emotion; hand and finger movements are calculated from the arm movement.

- All this information is sent to the augmented scenario in real time, making a live interaction between the real actor and an audience possible.
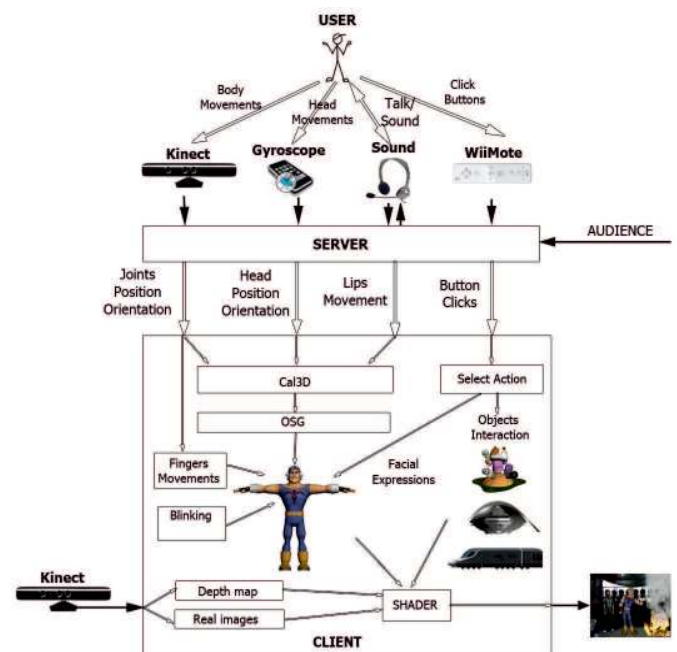
Conclusions

- The system has been successfully used in a real marketing presentation involving touristic contents (FITUR'11), where nearly 100 person enjoyed the augmented mirror experience.

- Kinect motion capture capabilities have been used and enhanced with other 3 input devices to perform a MoCap system which allows the actor to control the avatar movements, facial expressions, lips movement and also some predefined virtual objects.