

# Augmented Reality Using a 3D Motion Capturing Suit

Ionut Damian  
Human Centered Multimedia  
Augsburg University, Germany  
damian@hcm-lab.de

Mohammad Obaid  
Human Interface Technology  
Lab New Zealand  
Christchurch, New Zealand  
mohammad.obaid@hitlabnz.org

Felix Kistler  
Human Centered Multimedia  
Augsburg University, Germany  
kistler@hcm-lab.de

Elisabeth André  
Human Centered Multimedia  
Augsburg University, Germany  
andre@hcm-lab.de

## ABSTRACT

In the paper, we propose an approach that immerses the human user in an Augmented Reality (AR) environment with the use of an inertial motion capturing suit and a Head Mounted Displays system. The proposed approach allows for full body interaction with the AR environment in real-time and it does not require the use of any markers or cameras.

## 1. INTRODUCTION

Augmented Reality (AR) is a technology which augments the surroundings of the user with virtual elements, such as 3D objects or virtual agents. This technology has been used for over a decade in various of fields, including medicine, engineering, social science and gaming, to enhance the user's perception of the real world [1].

Typically, AR systems use a computer vision based tracking approach. This approach "is accurate, but only works over short distances"[3] and usually requires distinct markers placed throughout the real world [1, 3]. There have been attempts to remedy these problems, such as using GPS or compass sensors [9], tracking 3D structures instead of 2D images [6, 7] or using inertial sensors [2]. However, in all of these approaches the user is just a single point in the virtual space.

This paper proposes a novel approach to AR in which the user's whole body is immersed into the AR environment with the use of a real time inertial motion capturing (Mocap) system<sup>1</sup>. This approach greatly increases the interaction possibilities between the user and the virtual world. In addition it does not require the use of markers or cameras as the tracking and synchronization between real and virtual world is handled by the Mocap system. This gives the user more mobility as the tracking happens continuously regardless of the user's position or orientation.

In our previous work [8] we used a Microsoft Kinect<sup>2</sup> to

<sup>1</sup><http://www.xsens.com>

<sup>2</sup><http://http://www.microsoft.com/en-us/kinectforwindows/>



Figure 1: The AR setup featuring the Xsens MoCap system, the Vuzix HMD and a virtual agent

achieve AR tracking. A similar tracking approach has been presented by Izadi et al. [5] where they use a Kinect camera to generate a 3D representation of an environment with which the user can then interact. However, the Kinect's small field of view and limited tracking distance (at max 4m) hinders the mobility of the user. In our previous setup, we additionally noticed that the rather high amount of tracking noise resulted in a "jittery" virtual scene. Using the MoCap, the operation range is increased to 150m outdoors and 50m indoors and the tracking is less noisy.

A possible application for our approach is AR-based training, coaching or rehabilitation, where the system can accurately evaluate and analyze the user's movements due to the MoCap data. The approach can also be used to increase the user's sensation of presence by having the AR environment respond to the user's body.

## 2. THE SYSTEM

### 2.1 The Setup

The proposed AR setup uses the Xsens MVN inertial motion capturing suit, the "see-through" Vuzix STAR 1200

Head Mounted Displays (HMD)<sup>3</sup> and the AAA framework [4] to generate the virtual scene.

The Xsens MVN system is a wireless full body suit consisting of 17 inertial sensors which track the user's skeleton in real time. This information is then sent to our AR system which manipulates the virtual objects and renders them on the HMD. Figure 1 illustrates the setup.

## 2.2 User Tracking

The tracking within the AR setup is handled by the Mocap system. This is composed of two parts: (1) real time tracking of the position of the user's body in the real environment and (2) real time tracking of the user's head position and orientation. The position of the user in the environment is determined by the Xsens MVN Studio software using complex bio-mechanical simulations, whereas the head orientation is obtained from an inertial sensor attached to the user's head. The head position is computed relative to the user's position. This data enables the system to accurately determine where the user is situated in the real world and where she or he is looking at. This is done in absence of any cameras or markers in the real world. This means that the user's movement is not restricted and the synchronization between the real world and the virtual world can happen regardless of the user's position or orientation as long as she or he is within the system's range of 150m (outdoors) or 50m (indoors).

In addition, the system receives data about the rest of the user's body which can be used to further enhance the interaction with the virtual world. For example, the position of the hands can be used for gesture recognition or for precise object manipulation. The data can also be used for an in-depth analysis of the user's movement, such as computing various social cues (i.e. energy, fluidity, spatial extent) to give insight into the user's mental state.

## 2.3 Virtual Scene and Display Device

The virtual scene is rendered using the Horde3D graphics engine<sup>4</sup>. For the synchronization of the virtual scene with the real environment we use a representation of the user in the virtual world, a virtual agent. The position of this agent is updated every frame to match the position of the user in the room. The virtual camera's position and orientation is continuously synchronized with the user's head. We then render the scene stereoscopically using two different camera positions, one for each eye, to simulate binocular vision.

Using this approach we are able to place any 3D object in the virtual environment and its position and orientation will be continuously updated so that it matches the user's perspective to generate the AR effect without having an anchor point in the real environment (i.e. marker).

The AAA framework is designed to realize simulations of social situations using virtual agents. In the proposed setup, we use AAA to control virtual agents in the AR environment so that they behave and interact realistically with the user and their surroundings.

## 3. IMPRESSIONS AND FUTURE WORK

In this paper we presented a new approach to Augmented Reality based on a motion capturing system. The main ad-

<sup>3</sup><http://www.vuzix.com>

<sup>4</sup><http://www.horde3d.org>

vantages over common AR setups are (1) greatly improved interaction possibilities between the user and the AR environment due to the full immersion of the user's body in the virtual scene and (2) increased mobility for the user through the elimination of markers and cameras.

Initial tests were performed with a few users who were asked to navigate through an AR scene which included a virtual agent. The users stated that the tracking was very accurate and they felt as if they are part of the same environment as the virtual agent. A main limitation noted when using the AR technology is the limited field of view of the HMD for the virtual scene. This effect can be greatly reduced by using HMD's with larger displays, such as the upcoming Oculus Rift<sup>5</sup>. Other tests are planned to further evaluate the capabilities of this approach. We also plan to use this setup to study various elements of the human behavior in social interactions, such as proxemics, postures, involuntary reflexes and presence.

**Acknowledgment.** This work was partially funded by the European Commission within the 7th Framework Program FP7-ICT-2011-7 (grant agreement no. 288578).

## 4. REFERENCES

- [1] R. Azuma. A survey of augmented reality, 1997.
- [2] R. Azuma and G. Bishop. Improving static and dynamic registration in an optical see-through hmd. In *Proceedings of the 21st annual conference on Computer graphics and interactive techniques*, SIGGRAPH '94, pages 197–204, New York, NY, USA, 1994. ACM.
- [3] M. Billinghurst. The Future of Augmented Reality in Our Everyday Life. In *Proc. of the 19th Int. Display Workshops*, Nagoya, Japan, December 2011.
- [4] I. Damian, B. Endrass, P. Huber, N. Bee, and E. André. Individualizing Agent Interactions. In *Proc. of 4th Int. Conf. on Motion in Games (MIG 2011)*, 2011.
- [5] S. Izadi, D. Kim, O. Hilliges, D. Molyneaux, R. Newcombe, P. Kohli, J. Shotton, S. Hodges, D. Freeman, A. Davison, and A. Fitzgibbon. Kinectfusion: real-time 3d reconstruction and interaction using a moving depth camera. In *Proc. 24th ACM Symp. on User interface software and technology*, UIST '11, pages 559–568, NY, USA, 2011.
- [6] G. Klein and D. Murray. Parallel tracking and mapping for small AR workspaces. In *Proc. 6th IEEE and ACM Intl. Symp. on Mixed and Augmented Reality (ISMAR'07)*, Nara, Japan, November 2007.
- [7] V. Lepetit and P. Fua. Monocular model-based 3d tracking of rigid objects. *Found. Trends. Comput. Graph. Vis.*, 1(1):1–89, Jan. 2005.
- [8] M. Obaid, I. Damian, F. Kistler, B. Endrass, J. Wagner, and E. André. Cultural behaviors of virtual agents in an augmented reality environment. In *Proc. of the 12th Int. Conf. on Intelligent Virtual Agents, IVA'12*, pages 412–418, Berlin, Heidelberg, 2012. Springer-Verlag.
- [9] G. Reitmayr and T. Drummond. Going out: robust model-based tracking for outdoor augmented reality. In *Mixed and Augmented Reality, 2006. ISMAR 2006. IEEE/ACM International Symposium on*, pages 109–118, oct. 2006.

<sup>5</sup><http://www.oculusvr.com/>