Mixed Reality on Mobile Devices *

Adel Rizaev
University of Fribourg
Departement of Computer Science
DIVA Group
CH-1700 Fribourg, Switzerland
adel.rizaev@unifr.ch

ABSTRACT

The paper represents a desk research, which is done for the seminar *Multimodal Interaction on Mobile Devices* in the University of Fribourg, Switzerland. Different aspects of Mixed Reality on Mobile Devices, such as displays of mixed reality, tracking technologies and techniques, mixed reality applications for mobile devices (which making use of humans different modalities), etc., are reviewed in that paper. Of course, each of the topic in the paper can be considered as a separate theme for discussion. But that paper also can be useful, because it does an overview of the work, which is already done and the main tendencies in mobile mixed reality domain.

Categories and Subject Descriptors

H5.1 [**Information Systems**]: Multimedia Information Systems—*Artificial, augmented, and virtual realities*

General Terms

Human Factors, Exprerimentation

1. INTRODUCTION

Since Ivan Sutherland has described his augmented reality system [2] in 1968, which is also the first such a kind of system developed, a great effort has been done [4]. Many *Mixed Reality* applications have been implemented within non-portable computer setups. But increasing computational power of mobile devices, equipped with high-resolution cameras and displays, and ubiquitous availability of the mobile devices, like mobile phones, make them a perspective platform for: 1) representation of real-world objects, augmented with useful geotagged virtual information, 2) offering new interaction capabilities with virtual world objects for entertainment, education, collaborative work on the same virtual space with remote users and so on.

*The Paper is a delivrabled of the Msc Research Seminar *Multimodal Interaction on Mobile Devices* in 2012, DIVA Group University of Fribourg, supervised by Dr. D. Lalanne

The rest of the paper will be structured as the following. At first, we will briefly discuss the concept of mixed reality in general and give the definition. The next section of the paper is an overview of the mobile mixed reality from the perspective of the system, where main design issues and some technical solutions will be described. In the further sections we will discuss some interesting examples of the mobile mixed reality applications and make a conclusion.

2. MIXED REALITY

In 1994 P. Milgram and F. Kishino published their "Taxonomy of Mixed Reality Visual Displays" [3]. In that paper they define a *reality-virtuality* continuum. At the right side of the continuum *virtual environments* are placed, and at the opposite, left end of the continuum there are *real environments*. A virtual environment is an environment, consisting only from virtual objects (*virtual objects are objects that exists in essence or effect, but not formally or actually*). For instance, some computer graphic simulation, according to the definition given in [3], is an example of the virtual environment. In contrast, real environments consist only from real objects (*real objects are any objects that have an actual objective existence*). The set of real environments include, for example, a real-world scene, which is observed through some conventional video display, or the same real-world scene, which is observed without any help of any particular display system. If virtual objects and real objects are appearing together within a single display, i.e. every environment, which is in between these two opposite poles of the reality-virtuality continuum, then we reference it as a *mixed reality* [3]. Hence, one can speak about *augmented reality* (AR) and *augmented virtuality* (AV) systems. In summary, mixed reality (MR) is a common term for AR and AV:

- AR is in the case, where a system augments the user's view of the real world with supplementary virtual information.
- AV systems immerse the reality (i.e. user) in virtual world.

Azuma et al. named some requirements for implementation of MR system in order to be successful in [4]:

- system combines virtual and real information,
- system runs interactively and in real time, system registers (aligns) virtual and real objects with each other.

For a long time full compliance with these requirements has been infeasable, especially with limited resources of mobile devices. But the recent achievements in technology of mobile devices with high-resolution displays, cameras, broadband connectivity, graphical capabilities and growing computational power give good opportunity to speak about mobile mixed reality.

3. MOBILE MIXED REALITY: ENABLING TECHNOLOGIES

In this section of the paper we will try to present some of the most important enabling technologies for the mobile mixed reality systems. Of course, the basic enabling technology for mobile mixed reality (MMR) is mobile device itself. But also additional technologies for MMR are needed. As listed [4], such advances in the basic enabling technologies are: displays, tracking sensors, tracking techniques, calibration and registration techniques.

During the evolution of mobile devices different platforms were used by researchers for MMR applications[1]. The increasing computational power of personal mobile devices has made it possible to migrate the MR applications from backpack platforms [5], weighting few kilograms, to PDA's [6], Tablet PC's, and now mobile phones.

3.1 Displays

MR visual information could be presented using various displays [4]. One of the basis requirements for MR displays has been already proposed by I. Sutherland in [2]: "a perspective image, presented to the user must be changing, as the user moves". Supporting that claim, it is possible to suppose the following three types of displays, which can be available for MMR applications:

- *head-attached displays*: head-mounted displays (HMD), retinal displays (utilizes low power laser projection onto retina of human eye), head-mounted projectors (not widely accepted outside of research labs, because of their cumbersomeness). HMD's can be of the two types: optical see-through and video-through. Issues associated with video-through displays are latency, field of view and perception factors.
- *hand-held displays* (for example, LCD displays of mobile phones). That kind of displays may use an attached camera to provide video-through augmented information.
- *pocket projection displays*[2].

It is important to notice, that despite of the fact, that visual displays are more prevailing in MMR [7], designing displays of other modalities may be a challenge for researchers. In addition to the visual displays, auditory displays and haptic displays may be supposed to use in MMR systems [3]. For example, Albrecht et al. describe microphone-hearthrough augmented reality hardware system, consisting of headphones with built-in microphone for each ear (i.e. binaural microphone), in [12].

3.2 Tracking Sensors and Techniques

To provide stable visual mixed reality illusion, estimating of the camera position and information about actual coordinates in global space are required. Any 3D body may have three coordinates in the global space (x, y ,z) and can be rotated about three orthogonal axes (x, y, z), which is corresponding to the *Euler angles* of the 3D object: *roll*, *pitch* and *yaw*. Therefore, Mixed Reality requires 6 DOF pose tracking for mobile devices [9]. The requirements for pose tracking are inexpensiveness, robustness and possibility to operate in real time with changing environment conditions [9]. For that purpose different tracking sensors and tracking techniques are employed. In the context of MMR the following tracking sensors are wide spread:

- GPS is a *Time-of- Flight* (TOF) system, which measures the distance between mobile device with GPS receiver and the features associated with the reference object (e.g. coordinates of the satellite). The measurements are done by calculating the time delay in the propagation of the signal between satellite with known coordinates and antenna of the GPS receiver. GPS system uses 24 satellites out of 30 (6 are reserved) and 12 ground stations spread around the world [8].
- Optical sensors (e.g. cameras)..
- Mechanical gyroscopes (uses physical phenomena of attempt to conserve a given rotation axis) and accelerometers (uses physical phenomena of attempt to conserve a given position) are employed in inertial sensing.
- Magnetic field sensors (e.g. RFID) have high update rate and have small weight, but any nearby metallic object with its magnetic field can distort indication of the sensor [7, p.195].
- Acoustic sensors.

Each of the above listed sensors has its own advantages and disadvantages [8].

Due to the limitations of the paper size, we will not cover all of the implemented tracking techniques, but good reviews can be founded in [7] and [8]. In brief, actual tracking techniques could be combined in three groups: *sensor-based* tracking techniques, *vision-based* tracking techniques and *hybrid* tracking techniques. The most accurate sensor-based tracking techniques combine different type of sensors and perform dynamic data fusion [7].

Vision-based tracking techniques use image processing algorithms to estimate camera position relative to real world objects [7]. Respective to computer vision algorithms, vision-based techniques can be of two types: *feature-based* and *model-based* [7]. In a nutshell, feature-based tracking techniques are trying to find correspondence between the plane features of the image and their coordinates in the frame of the real-world. Then the camera pose can be founded

[1] `https://www.icg.tugraz.at/~daniel/HistoryOfMobileAR/` (checked at May 2012)
[2] `http://web.media.mit.edu/~raskar/pocket.html` (checked at May 2012)

by computing projection of the 3D feature coordinates on the observed 2D image coordinates, while minimizing the distance to their corresponding 2D features [7, p.195]. For computing that homography, the correspondence between minimum four features must be found. Due to the limited capacity of computational power in older mobile devices, the most efficient implementation of vision-based techniques used artificial markers (*marker tracking*). For the tracking purposes real environments are supplied with special *fiducial* markers, which are less computationally demanding (see [6], application "Signpost"). Later, tracking techniques running on mobile devices and *based on the natural features* (such as, points, lines, edges, textures) were introduced [7, p.195], [10]. For example, Wagner et al. modified SIFT and Fern features to implement natural feature tracking on mobile phones [10]. The resulting tracking techniques (*PhonySIFT* and *PhonyFerns*) were capable to provide the tracking with real-time frame rates (30 Hz) and six degrees of freedom using only built-in camera of the phone. Another vision-based tracking technique is a model-based tracking method. In that kind of tracking techniques CAD models or 2D templates of the tracked object's features are used. For computing the camera pose edges of the object in the frame of the real-world detected and matched with their corresponding edges of the 3D model or 2D template. Wagner et al. presented *PatchTracker*, the model-based tracking technique, which employs 2D image as a template, in [10]. They also combined each of the both: *PhonySIFT* and *PhonyFerns* with *PatchTracker* and obtained faster and more robust tracking system. 3D model-based tracking for mobile AR applications has been proposed by Seo et al. in [14]. For estimating camera poses the method starts with edge detection using Canny algorithm. Then lines on the 3D model are filtered using visibility tests and projected on the scene image. After, the correspondence between the edges, observed on the frame image, and the ones on 3D model, can be found as the closest edge scene. Then, the camera pose estimation can be computed.

Hybrid tracking techniques are very often used for outdoor AR applications, where vision-based technique alone can't provide robust tracking. For example, in [4] Azuma et al. propose, that MMR application used in outdoor, where it is not practical to cover the environment with markers, can employ compass and gyroscope tracking in addition to vision-based tracking to provide more robust results. Actually, tracking on unprepared environments is very depending on the visibility of the natural features. And if the database of the environment exists, then tracking could be based on the visible horizon silhouette, but in order to access that database some initial tracking information, such as GPS and compass data can be needed [4], [7].

4. MOBILE MIXED REALITY: APPLICATIONS OVERVIEW

Computational cost of augmented virtuality makes it very challenging task to implement on mobile devices, like mobile phones. Therefore the major part of mixed reality applications for mobile devices, introduced by researchers by now, are augmented reality applications. Depending on the purpose of the use, MMR applications can be classified in one or several of the following categories: learning, tourism, navigation and path finding, maintenance and inspection, military, collaborative applications, games and entertainment, etc. Another, more systematical way to classify MMR applications is to order them: 1) from the technical parameters point of view (for example, taxonomy proposed by P. Milgram et al. in [3] is a classification according to visual displays, although that taxonomy can not be widely applied to classify MMR applications), 2) from the information presenting point of view (is the information 2D or 3D, temporal aspects, etc.), 3) from the user point of view (which modalities are involved in MMR application, sociality, functions of MMR application, etc.). It is also possible to classify MMR application, by mixing different criteria from the above listed categories. Let's briefly discuss some MMR applications.

Signpost was one of the earlier AR applications, implemented for PDA handheld platform [6]. The presented system worked fully autonomously on PDA (HP iPAQ 5450) and used optical tracking with optional back-end server support for the gain in the application's performance (by outsourcing to the server some computationally-demanding tasks). The application was able to navigate a user through unknown building, using fiducial markers attached to the building's wall.

An interesting paper [12], presenting mobile audio AR system, is written by Albrecht et al. The *mic-through* AR system uses binaural microphone embedded in the user's headphones. The real-world sounds, sensed by binaural microphone, can be augmented by virtual audio information. One of the possible use cases of such a system, which is not mentioned in paper, could be employing it in application, which navigates blind-people in the real-world environments. The auditory space perception of the human is very high. Hence, the virtual audio information can be used to indicate the real-world objects in the acoustic space. But tracking of such kind of devices is still a challenging task [12].

A way of interaction, allowing mobile phone user to manipulate virtual objects using hand gestures, introduced in [13] by Caballero et al. This way of interaction with virtual objects, called *behand*, is based on mobile augmented virtuality technology. The camera, placed on the back of the user's mobile phone, captures the gestures of the user, while he moves his hand behind the phone. The hand of the user is transported into the virtual world, which is "inside" of the mobile phone. Then the authors present some prototypes, which demonstrate the use of their *behand* interaction method in different contexts, like games, museums, interaction with smart objects, and collaborative interaction in some shared virtual space.

MobGeoSen is an example of application, which is not directly related to MMR, at first sight, but it may help to deliver content for MMR applications [11]. The application collects data from different internal and external sensors of mobile device, such as: microphone, camera, GPS receiver, temperature sensor, accelerometer, electric field strength sensor, carbon monoxide level measurer and so on. Further, collected from different mobile devices data can be stored on some host storage, some methods of statistics can be applied to it, additional information could be extracted. Finally, processed geotagged data can be employed in different visualizations of the real environments, also using with MMR technology.

Technology		Some of the recent achievements	Future work
Tracking	Sensor-based	GPS, optic sensors,gyroscopes, accelerometers, acoustic sensors, ultrasonic sensors, RFID location sensing[3]	ubiquitous tracking (using automatically and dynamically fused data from heterogeneous and widespread sensors); implementing effective sensor data fusion algorithms at sensor data, feature and decision levels
	Vision-based	Feature-based: marker (ARToolKit[6]) / markerless(from natural feautes);Model-based: 2D template [10]/3D model with edge detection[14]	temporal continuity relations between frames, tracking in previously unknown environments (SLAM); achieving of real time frame rates with 3D model-based tracking
	Hybrid tracking	Combination of inertial and vision-based tracking techologies; GPS sensor with vision-based tracking.	implementing algorithms independent from some specific sensors; effective using of the sensor data as initial estimate for vision-based tracking
Displays	HMD	Relative low cost	demanding of higher resolution, wide FOV, higher speed and flexible head mobility, smaller size and weight, social acceptance
	Projectors	Geometrically awareness and self-calibration[4]	Undistorted projection, low power, compact size, multiple viewing angles; embedding in mobile phones
	Hand-held	Small sizes and weights	new interaction metaphors and techniques

Table 1: MMR Technologies: some of the recent achievements and future work.

5. CONCLUSION

Concluding this paper, let's briefly discuss some challenges and future work for MMR. At first, let's take a look at the Table 1, including some of MMR technologies and techniques with the results, which are already achieved and the work, which can be considered as a future work. The Table 1 is based on [7], [8] and [9]. We see, that the ideas, proposed by early MR researchers in 60-70's, are highly developed and the set of the problems, associated with mixed reality, now can be considered as a separate research area, which involves many other disciplines. D. Schmalstieg et al. have been described the idea of *Augmented Reality 2.0* (AR 2.0) in [1]. AR 2.0 is a platform, which joins ubiquitous availability of mobile hardware and social networking power of Web 2.0 software. The main difference between AR 2.0 applications and current AR applications is that the former enable the users to create location-aware MMR experiences and modify them. The authors show, that the AR technology has developed to the point, that it can be deployed on mobile consumer-level hardware. In addition with user-generated content the idea of AR2.0 can be realizable [12]. May be some times in the future, with enough powerful mobile hardware and more sophisticated software, the community will be fascinated with idea of MR2.0, which contains also VR2.0 and makes use of multimodality.

6. REFERENCES

1. D. Schmalstieg, T. Langlotz, M. Billinghurst, "Augmented Reality 2.0", in "Virtual Realities", Dagstuhl Seminar, Springer, 2008.

2. I. Sutherland, "A Head-Mounted Three Dimensional Display", in "Proceedings of Fall Joint Computer Conference", 1968, pp. 757-764.

3. P. Milgram and F. Kishino, "Taxonomy of Mixed Reality Visual Displays", in "IEICE Transactions on Information and Systems", Vol E77-D, No.12 December 1994

4. R. Azuma, Y. Baillot, R. Behringer, S. Feiner, S. Julier, B. MacIntyre, "Recent Advances in Augmented Reality" in "IEIE Computer Graphics and Applications", November-December 2001.

[3] http://web.media.mit.edu/~raskar/Sig04/ (checked at May 2012)
[4] http://web.media.mit.edu/~raskar/pocket.html (checked at May 2012)

5. S. Feiner, B. MacIntyre, T. Hollerer, A. Webster, "A touring Machine: Prototyping 3D Mobile Augmented Reality Systems for Exploring the Urban Environment", in "Proceedings of the 1st IEEE International Symposium on Wearable Computers", 1997.

6. D. Wagner, D. Schmalstieg, "First Steps Towards Handheld Augmented Reality", in "Proceedings of the 7th IEEE International Symposium on Wearable Computers", 2003.

7. F. Zhou, H. B-L. Duh, M. Billinghurst, "Trends in augmented reality tracking, interaction and display: A review of ten years of ISMAR", in "Proceedings of the 7th IEEE/ACM International Symposium on Mixed and Augmented Reality", 2008.

8. J. P. Rolland, L. D. Davis, Y. Baillot, "A Survey of Tracking Technologies for Virtual Environments", in "Fundamentals of Wearable Computers and Augmented Reality", W. Barfield, T. Caudell, 2001, pp. 67-112.

9. D. Wagner, D. Schmalstieg, "History and Future of Tracking for Mobile Phone Augmented Reality", ISUVR, 2009.

10. D. Wagner, G. Reytmair, A. Mulloni, T. Drummond, D. Schmalstieg, "Real-Time Detection and Tracking for Augmented Reality on Mobile Phones", in "IEICE Transactions on Visualization and Computer Graphics", 2010.

11. E. Kanjo, S. Benford, M. Paxton, A. Chamberlain, D. Stanton Fraser, D. Woodgate, D. Crellin, A. Woolard, "MobGeoSen: facilitating personal geosensor data collection and visualization using mobile phones", in "Personal Ubiquitous Computing", November 2008.

12. R. Albrecht, T. Lokki, L. Savioja, "A mobile augmented reality audio system with binaural microphones", in "Proceedings of Interacting with Sound Workshop: Exploring Context-Aware, Local and Social Audio Applications", 2011

13. M. Luz Caballero, T. Chang, M. Menéndez, V. Occhialini, "Behand: augmented virtuality gestural interaction for mobile phones", in "Proceedings of the 12th international conference on Human computer interaction with mobile devices and services", 2010.

14. Byung-Kuk Seo, Jungsik Park, Jong-Il Park, "3-D visual tracking for mobile augmented reality applications", 2011