2008

A survey of mobile and wireless technologies for augmented reality systems

Papagiannakis, George; Singh, Gurminder; Magnenat-Thalmann, Nadia

A survey of mobile and wireless technologies for augmented reality systems

George Papagiannakis*, Gurminder Singh**, Nadia Magnenat-Thalmann*,
*MIRALab, University of Geneva, 7 Route de Drize, CH-1227, Carouge, Switzerland,
george.papagiannakis@miralab.unige.ch, nadia.thalmann@miralab.unige.ch
** Center for the Study of Mobile Devices and Communication, Naval Postgraduate School, Monterey, California 93943, USA, gsingh@nps.edu

ABSTRACT

Recent advances in hardware and software for mobile computing have enabled a new breed of mobile AR systems and applications. A new breed of computing called "augmented ubiquitous computing" has resulted from the convergence of wearable computing, wireless networking and mobile AR interfaces. In this paper we provide a survey of different mobile and wireless technologies and how they have impact AR. Our goal is to place them into different categories so that it becomes easier to understand the state of art and to help identify new directions of research.

Keywords

Augmented-Mixed Reality, Mobile Systems, Wireless networking

1. INTRODUCTION

For the last forty years, interactive 3D graphics have focused on the "kinetic depth effect" so that the image presented by the three-dimensional display changes in exactly the same way that the image of a real object would change for similar motions of the user's head (Sutherland et al [1]). This basic metaphor has been the driving force behind "Virtual Reality" and the immersion in virtual environments. This base idea was further enhanced to "augment" the visual field of the user with information necessary in the performance of the current task, enabling an "Augmented Reality" (AR) Caudell et al [2]. Although AR was meant to include mobility, it was not until "The Columbia Touring Machine" by Feiner et al [3] that the first outdoor Mobile Augmented Reality System (MARS) was created. Around the same time as the development of MARS, research on wearable computers and personal imaging got started Mann [4]. Wellner et al [5] introduced the term "Computer AR" to include systems such as the augmented digital desk interface which enhance the physical world by superimposing computer generated scenes. The Mixed Reality-Virtuality Continuum has been consequently defined by Milgram et al [6] based on the Extent of World Knowledge Continuum as depicted in the following **Error! Reference source not found.**, i.e. the level that the depicted World is modeled in 3D. Since this early work, researchers have been working to improve the methods and algorithms to allow convincing traversal of this MR continuum (Figure 1). Azuma et al [7] have surveyed the MR continuum that included the notions of VR, AR and AV (Augmented Virtuality).

At about the same time during the 1990s that AR research experienced the above renaissance, Weiser [8] conceptualized the idea of "ubiquitous computing": an environment in which computing technology is embedded into all kinds of everyday physical objects, (such as appliances, doors, windows, or desks) which results in the "computer disappearing into the background". The opposition between the notion of virtual reality and ubiquitous, invisible computing is so strong that Weiser coined the term "embodied virtuality" to refer to the process of drawing computers out of their electronic shells. Recently miniaturized mobile devices have extended their capabilities from simple communication devices to wearable, networked computational platforms. Mobile AR (Figure 2) can be viewed as the meeting point between AR, ubiquitous computing and wearables.

Thus within the scope of this work, we define a mobile AR system as the one which:

- Combines real and virtual objects in a real environment
- Runs in real-time and mobile mode
- Registers(aligns) real and virtual objects with each other
- The virtual augmentation is based on dynamic, 3D objects (e.g. interactive, deformable virtual characters)

The basic components of such a mobile AR system include: a) h/w computational platform, b) display, c) tracking, d) tracking, e) wireless network, f) wearable input and interaction, and g) software.

A successful mobile AR system would enable the user to focus on the use of the system rather than its computing equipment. Ultimately the user should not be aware of presence computing equipment (Weiser [8]). With lightweight, wearable or mobile devices, and tracking technology embedded in the environment, it is becoming possible to achieve Weiser's vision. Recent advanced in wireless technology is further supporting the creation of such environments. Since the focus of our paper is mobile AR systems, we present the following enabling technologies:

- Mobile computational platform devices
- AR system architecture and Content
- Wireless networking

In our study, AR is treated as a user interface for both ubiquitous mobile computing as well as wearable computing since the real world is leveraged as in interface itself. An ideal mobile AR system would include a pair of stylish sunglasses equipped with high-resolution 3D graphics capabilities, built-in computer with wireless network support, and accurate 6 Degree of Freedom (DOF) tracking (Azuma et al [9]). A mobile user would not need to wear or carry any further equipment in order to experience mobile AR (Hollerer [10]). All computation and sensing could be embedded in the environment itself as infrastructure. For example, cameras could cover all possible spaces, and distributed computation and displays stationed in the environment could provide AR augmentations. Already today wireless networking as well as GPS & Galileo coverage has started following this trend.

We present taxonomy of the research work that has been carried out in Mobile AR systems since the last significant survey of Azuma et al [1] published in 2001. In the last few years there has been an explosive proliferation of wireless technologies, mobile devices, networking standards and distributed computing power, allowing for new forms of such Augmented Ubiquitous Computing. The rest of this paper presents work in this area. While we try to be as complete as possible in our coverage, we have not attempted to cover all systems developed to date.

Section 2 presents an overview of the various mobile AR enabling technologies. Since mobility is our focus, we begin by reviewing the state-of-the-art in mobile devices. There is a wide variety of hardware computing platforms used in mobile AR. We review these in Section 2.1. The cost and effort in developing mobile AR systems is quite high. To reduce these, a number of different software architectures and toolkits have been proposed. We provide a review of these in section 2.2. Since mobility is a critical part of mobile AR systems, a significant attention is devoted to wireless technologies in section 2.3. At the heart of AR are the registration of virtual augmentation correctly in the real world (Section 2.4) and appropriately displaying the composition of the two worlds (Section 2.5). The recent advances in these mobile AR enabling technologies have allowed for a complete new breed of applications covered in Section 3. Finally in Section 4, we discuss and compare all recent advances in the previous domains and provide our recommendations for future research directions and synergies in the area of mobile AR systems.

2. Enabling Technologies for Mobile AR Systems

2.1 Computing Hardware for Mobile AR

2.1.1 Mobile workstation & wearable PC

A number of systems (Tamura et al [28], Cheok et al [27], Piekarski et al [23], Papagiannakis et al [12], Hughes et al [32], Hollerer [8], Scmeil et al [40], Wagner et al [25] and Egges et al [11]) have employed mobile workstations (Figure 3), often aggregated together with other mobile equipment, in the form of a backpack (weighting 1-6 Kgs), so that the user can freely move in the real environment and have their hands free for input and interaction. These backpacks include amongst others, mobile workstations such as Dell™ Inspiron & Precision, Alienware™ and JVC™ sub-notebooks. Although severe ergonomic issues are apparent due to the size and weight of the backpack, it allows researchers to focus on their research without the constraints that smaller devices often present, namely in the computational power, operating system and hardware connectivity. Almost all of the desktop computing system can be made mobile by using high-end notebook computers. However, due to the backpack setup, the use of head mounted displays (HMDs) is enforced as opposed to handheld display that other devices can offer. The next step towards this direction is the employment of ultra mobile PCs (UMPCs), discussed in section 2.1.3, that could provide both handheld as well as HMD capabilities.

2.1.2 Tablet-PC

In attempt to use mobile powerful handheld displays for AR, tablet-PCs have been employed in a number of mobile AR systems Klein et al [20], Stork et al [26], Zauner et al [31] and Renevier et al [42]. A tablet-PC is a notebook or slate-shaped mobile computer which allows to be operated via a fingertip or stylus thus offering a more convenient way of interaction. Special tablet-PC editions of the Microsoft Windows™ and Linux™ OS have been mostly involved in the AR systems. The use of tablet-PCs eliminates the operating system and hardware shortcomings of small-size devices and the ergonomic issues of laptops inside backpacks.

2.1.3 UMPC

A very recent trend in mobile AR systems is the usage of ultra mobile PCs (Figure 4). UMPCs are based on the Microsoft Origami™ specification released in 2006 and have been developed jointly by Microsoft™, Intel™, and Samsung™, among others. UMPCs are basically small mobile PCs running Microsoft Windows XP. A number of researchers have started employing them in AR simulations such as Wagner et al [25], Newman et al [46] and specifically the Sony Vaio™ U70 and UX180, as well as Samsung™ Q1. Elmqvist et al [29] have employed the Xybernaut™ Mobile Assistant, which, although shares some common characteristics with UMPCs, does not belong in the UMPC category.

2.1.4 PDA

Before the recent introduction of UMPCs or SmartPhones with CPUs of significant compute power, personal digital assistants (PDAs) were the only truly mobile alternative for AR researchers. PDAs now have enhanced color displays, wireless connectivity, web-browser and GPS system. However, a number of computational issues make use difficult for AR due to lack of dedicated 3D capability and floating point computational unit. Goose et al [15], Reitnayr et al [18], Wagner et al [25], Barakonyi et al [33], Wagner et al [39], Gausemeier et al [47] have all employed them as handheld display devices for AR applications, whereas Makri et al [49] and Peternier et al [24] allowed for a custom-made connections with a special micro-optical display as an HMD. The majority of these applications were developed on top of the Microsoft Windows Mobile™ OS.

2.1.5 Smartphone

Currently mobile phones are the most widely used device. Hence, their usage in mobile AR systems would allow extending even more their range of applications and capabilities. Smartphones are fully featured high-end mobile phones featuring PDA capabilities, so that applications for data-processing and connectivity can be installed on them. Rashid et al [30], Wagner et al [25], Henrysson et al [22], Olwal et al [17] utilize them as final mobile AR displays. They are also often used in conjunction with a stationary server, as in the case of Rauhala et al [48], where a notebook PC was also used as an intermediate central data processing unit. As the processing capability of smartphones is improving, their application use is increasing. Jonietza [60] use their smartphone to calculate the location of just about any object its camera is aimed at. As the smartphone changes location, it retrieves the names and geographical coordinates of nearby landmarks from an external database. The user can then download additional information about a chosen location from the Web--say, the names of businesses in the Empire State Building, the cost of visiting the building's observatories, or hours and menus for its five eateries. The most often used operating systems upon which these AR applications where build include Symbian™, Windows Mobile™ and the Linux™ OS.

2.2 Software architectures for mobile AR

In this section we review the recent AR software system architectures beyond those initially introduced by Feiner et al [3], Tamura [28] and Billinghurst [35] covered in the survey of Azuma et al [1]. The reviewed frameworks in this section feature basic kernel, networking, display, tracking and registration components based on hybrid methods (sensor and vision based), allowing for the prototyping of different AR applications as a result of their toolkit, plug-in architecture. This is how most new architectures different from old mobile AR architectures that were built as single monolithic pieces of software.

Schmalstieg et al [45] introduced the "Studierstube" collaborative AR platform, based on a heterogeneous distributed architecture. Studierstube's software development environment has been realized as a collection of C+ + classes built initially on top of the Open Inventor (OIV) scenegraph toolkit and later on top of Coin3D. Applications are written and compiled as separate shared objects and dynamically loaded into the runtime framework. A safeguard mechanism makes sure that only one instance of each application's code is loaded into the system at any time. Besides decoupling application development from system development, dynamic loading of objects also simplifies distribution, as application components can be loaded by each host whenever needed. Studierstube is intended as an application framework that allows the use of a variety of displays with a variety of tracking devices with the concept of a distributed shared scene graph, similar to distributed shared memory. From the application programmer's perspective, multiple workstations share a common scene graph. Numerous mobile AR applications have been built based on this architecture such as Wagner et al [39], Reitmayr et al [18]and Newman et al [46].

Ponder et al [43] aimed to address some of the most common problems related to the development, extensions and continuous need of integration of heterogeneous simulation technologies under single system roof. VHD++ combined both framework (complexity curbing) and component (complexity hiding) based development methodologies. In effect large-scale architecture and code reuse is achieved adding to development efficiency and robustness of the final both AR and VR virtual character applications. The (now open-source under LGPL license) VHD++ framework provides an efficient research environment a) offering full power to researchers and at the same time b) allowing them to publish their research results in form of ready to use, plug-able Services (plug-ins encapsulating heterogeneous technologies) that are plugged to the generic Runtime Engine. Currently the list of Services that encapsulate the heterogeneous technologies emphasize on virtual character simulation such as facial and body animation, deformation, speech, cloth simulation, AR marker and markerless tracking, scripting etc. A number of AR applications where based on this framework such as those by Papagiannakis et al [12] and Egges et al [11].

Building complex VR/AR applications usually is a time-consuming task, even if only a small part of the system functionality is to be evaluated. Using the MORGAN framework developed by Ohlenburg et al

[41], distributed multi-user VR/AR applications can be implemented much faster. MORGAN is a component-based framework that relies on the CORBA middleware for network communication. It currently supports many devices, including mouse and keyboard as well as haptic input devices, object tracking systems and speech recognition libraries. Thus, multi-modal user interfaces can be rapidly developed and evaluated. Additionally, MORGAN provides a distributed render engine with automatic scene graph synchronization capabilities where all components are accessible from remote computers.

Hollerer [8] built a series of Mobile AR systems (MARS) prototypes, starting with extensions to the 1997 "Touring Machine" from Feiner et al [3] and leading up to their most recent system, MARS 2002. This featured a shared central Java and Java3D infrastructure which enables AR, VR and desktop-based indoor/outdoor communication. Important features of this architecture are the central database and networking server for inter-process communication and the relational database backend for persistent data storage. The platform allowed for indoor/outdoor tracking, navigation and collaboration based on hybrid User Interfaces (2D and 3D). As AR interfaces have to consider both virtual and physical objects and potentially a multitude of devices (input and output) visuals become easily cluttered and overwhelming. A main innovation from this system was a rule-based architecture for adaptive MARS interfaces and UI management.

Hughes et al [32] utilize their own MR Software Suite (MRSS) for the development and delivery of their MR experiences. The main system consists of four subsystems called Engines: three rendering engines for visual, audio and special effects multimodal simulation and a fourth integration engine combining the above in interactive, non-linear scenarios. The central networking protocols receive and integrate sensor data such as tracking, registration, and orientation with input from other sources, for example, artificial intelligence and specialized physics engines, and execute a nonlinear, interactive script. This then produces a multimodal simulation situated within real world conditions based on the rendering and display technology available. The key technologies used in this MR system are Open Scene Graph and Cal3D for graphics, Port Audio for sound, and a DMX chain for talking to special effects devices. The network protocol is built on top of TCP/IP. Authoring of stories is done in XML, which can include C or Java-style advanced scripting. The MR system can run stand-alone (one user) or in combination with multiple MR systems (each managing one or more users).

Wagner et al [38] recently introduced Muddleware, a communication platform for mixed-reality multiuser games that is light-weight and highly portable, as shown in several MR game projects that was employed. A hierarchical database built on XML technology allows convenient prototyping and simple, yet powerful queries. Server side-extensions address persistence and autonomous behaviors through hierarchical state machines. The core of Muddleware is a memory mapped database that provides persistence and can be addressed associatively using XPath. Clients connect to the server by any of four APIs: Immediate C++, Shared Memory C++, Java and Muddleware Script. All data elements are stored as nodes of a modified XML DOM. Clients store arbitrary messages as XML fragments, and use XPath to specify query or update operations. A simple benchmark by the authors showed that its XML server can easily handle thousands of complex requests per second.

An integrated and uniform approach for building mobile applications for mixed reality environments has also been presented by Piekarksi et al [50] as an evolution of previous AR applications. The architecture uses a data flow methodology with an object-oriented design aiming to provide a simple model for programmers to re-use as well as compose new AR applications. Using this toolkit approach, a number of kernel features such as distributed programming, persistent storage, and run time configuration are possible. The design is based on the C++ language and the capabilities of this software architecture are demonstrated by the Tinmith-Metro mobile outdoor modeling application, as well as other AR examples (AR-Quake game etc.).

2.3 Wireless Networking for Mobile AR

This section will discuss the impact of wireless networking in AR. Wireless network characteristics differ quite markedly from wired in latency, bandwidth, bandwidth fluctuations and availability. These have direct impact on the performance and quality of user experience in AR. In addition, there are many different types of wireless networks available. These impact the types of applications that can be developed. To support a usable AR system, a wireless network should provide sufficient data rate, low latency and support for mobility. The ability to be mobile by itself introduces several new application possibilities but when combined with the knowledge of location, the same applications can become more useful and exciting. In the following sections, we present how different types of networks support these key requirements.

2.3.1 Wireless WANs and 3G Networking

The wireless wide area networks (WWANs) are ideal for AR systems that need to support large scale mobility, for example nation-wide or in a large city. Systems that provide location-based services are a good a very good example of such an application. The prototype system developed by Nokia in which the phone can calculate the location of just about any object its camera is aimed at (Jonietza [60]) falls in this category. As the phone changes location, it retrieves the names and geographical coordinates of nearby landmarks from an external database. The user can then download additional information about a chosen location from the Web--say, the names of businesses in the Empire State Building, the cost of visiting the building's observatories, or hours and menus for its five eateries. To be useful the system needs to support large scale mobility.

In the WWAN category, there are several choices available from the slow speed of 9.6 kbps to high speeds of the third generation (3G) of 2 mbps. While the availability of slow speed 2G WWANs such as GSM and CDMA is widespread, due to their slow speed and high latency, they are limited in their use for AR. It is possible to use such networks to implement applications where much of the data is local and little needs to be sent over the network. This scenario is very similar to early days of networked virtual reality systems which used modem-based slow speed connectivity for sharing data such as in NetEffect (Das et al [57]).

Viktorsson et al [64] make an interesting use of the SIM (Subscriber Identity Module) card used in GSM phones. SIM card is typically used by the user to store his personal information such as contact information of people he knows, his preferences and other personal information. Information about avatar characteristics for a user can also be stored in a SIM card. It can then be moved from one access terminal to another. A virtual world, which the avatar is designed to enter, can then be accessed from many different access terminals by means of inserting the SIM card and entering a personal identity number (PIN) code. Thus, besides making it possible to access a virtual world from different access terminals, this technique also makes it possible to use avatars in new applications.

Beyond 2G, we have the 2.5G and 3G WWANs which are designed to support multimedia applications by providing better network infrastructure. As a result, these networks should also provide better support for networked AR systems. An example of a system which runs on 3G phones and requires the high data rates of 3G networks but not necessarily their support for mobility is the virtual disco system by Artificial Life (http://www.3g.co.uk/PR/Jan2005/9022.htm). The 3G virtual disco is presented to the users as a 3-D animated virtual building with several floors representing different music clubs and styles of music according to the selection and tastes of the listeners. In the intro sequence the user can select an animated 3-D character (avatar) as his or her virtual persona and visit the different music rooms in the virtual disco. Users can download or stream music in combination with high quality 3-D animation clips showing synchronized dancing avatars.

GPRS (General Packet Radio Service) is considered as a 2.5G network and is built on top of a GSM network. The theoretical maximum data rate of a GPRS network can be as high as 171.2 kbps, but in practice, per connection, typically much lower data rate is available. Often, this is because of the limitations of the network infrastructure. In addition, the downlink and uplinks data rates are also significantly different, with uplink data rate being lower. Beyond data rates, the telecommunication operators have significant control over other Quality of Service (QoS) parameters, including priorities for services, transmission delay between different stations, and reliability of packet transmission. By controlling these parameters the performance of the network can be tuned for different purposes such as the number of connections the network can support, and QoS for services running on the network. As expected, the operators would optimize their network for supporting maximum connection as it results in higher revenue for the operator. This leads to not only low data rates but also high latency. Highly-interactive, multi-user VR environments require that end-to-end latency remain less than 100ms (Smed et al [61]), (Pantel et al [62]). This low level of latency is not supported by most 2.5G networks.

A few evaluation studies have been reported that measure the performance of network VR services on wireless WANs. Pervoic et al [59] have tested the performance of three VR systems over a GPRS network. In their GRPS network, the uplink data rate supported by the network was 44.8kpbs while the uplink was 11.2kpbs. Their VR applications included a multi-user community where users could see each others avatars, a conversational virtual character and a 3D multi-user game. As expected, with low data rates and high latency, they found the GPRS network to be insufficient for supporting the networked VR applications.

Gorseta et al [58] have done an experimental performance evaluation of networked VR systems in UMTS (Universal Mobile Telecommunication System) network. UMTS is considered a 3G network. The UMTS core network is based on GPRS network toology but differs in its air interface transmission. The air interface access method for UMTS is based on wide-band CDMA (code division multiple access) where as GPRS is based on TDMA (time division multiple access) and FDMA (frequency division multiple access). The theoretical maximum data rate in a UMTS network is 2mbps. However, due to the current mobile phone limitations, data rates higher than 384kbps cannot be supported. For a large number of interactive, multi-user VR environments, this data rate works well, but high latency (round trip time 300-580ms) in UMTS networks cause problems in implementing real-time services in VR environments. A number of location-based mobile games have been developed over the last few years which either use GPS or triangulation techniques based on cell-towers for location and a WWAN for networking.

2.3.2 WLANs

Wireless local area networks (WLANs), as the name suggests, are wireless networks implemented in a local area such as a home or an office building. WLANs typically will support much higher data rates (between 11-54 mbps) and lower latency than WWANs but their support for mobility is limited than in WWANs. Currently, WLANs can be built using any of the IEEE 802.11a/b/g/n standards compliant equipment.

Human Pacman (Cheok et al [27], [66]) is an interactive role-playing, physical fantasy game integrated with human-social and mobile-gaming that emphasizes on collaboration and competition between players. By setting the game in a wide outdoor area, natural human-physical movements become an integral part of the game. In Human Pacman, Pacmen and Ghosts are human players in the real world who experience mixed reality visualization from wearable computers. Virtual cookies and actual physical objects are incorporated to provide novel experiences of seamless transitions between real and virtual worlds and tangible human computer interface respectively. Human Pacman uses WLAN technology to enable mobility in small scale environments.

While the transmission range of one WLAN base station or access point is typically 100 meters, a number of access points can be used to provide coverage in much larger areas. Using the approach, currently there are efforts underway to provide WLAN coverage in entire cities. In this scenario, WLANs can be seen to be competing with 3G networks which are designed to provide WWAN capability (Ferguson, 2007). As the two wireless networking technologies begin to co-exist, it is useful to allow the VR environment to operate on whichever connection is connection is available at the time of operation and to be able to operate on cheaper connections. In the current state of the art, WLANs are significantly cheaper in cost than WWANs.

McCaffery et al [68] have developed a wireless overlay network concept which allows the application to select the most appropriate of a number of different network technologies as the mobile device moves around in the environment. The migration between different networks happens transparently to the application. To test their overlay network, they have developed an Compaq iPAQ-based first person shooter game called Real Tournament (Mitchell [69]). For this game, their overlay network works on top of IEEE 802.11b and GPRS.

WLAN capability in game machines has become very popular. As a result, many of the new game machines now come with built-in support for WLANs or provide attachments for WLANs. WiFi Max for Sony PSP is a good example of such an attachment. By plugging-in WiFi Max into an internet enabled PC, one can create a wireless access point. Five PSP machines can wirelessly connect to this access point and play games with one another as well as with other internet PSP players. Similarly Microsoft Xbox 360 can be made WLAN ready by installing a WiFi adapter card in it.

2.3.3 WPANs

The wireless personal area networks (WPANs) are short-range (typically a few meters), high-bandwidth wireless networks used for applications such as printing, file transfer and remote control. Often WPANs are implemented using Bluetooth or infra-red communication technologies. In VR, WPANs have been extensively used in combination with PDAs to interact with 3D VR environments. To control 3D environments, users often need to provide inputs through buttons, sliders and menus. Such input can be provided through a handheld device which can communicate with the VR environment through WPANs.

Watsen et al [70] have investigated the contention between 2D and 3D interaction metaphors and involved the use of a 3Com PalmPilot handheld computer as an interaction device to the VE, allowing the use of 2D widgets in a 3D context. Tweek (Hartling et al [71]) presents users with an extensible 2D Java graphical userinterface (GUI) that communicates with VR applications. Using Tweek, developers can create a GUI that provides capabilities for interacting with a VE. Tweek has been used in VR Juggler (Bierbaum [72]), an open source virtual reality development tool. More recently, use of handheld devices and WPANs has been extended to interaction with real-world scenarios. For example, Ubibus (2004) is designed to help blind or visually impaired people to take public transport. The application allows the user to request in advance the bus of his choice to stop, and to be notified when the right bus has arrived. The user may use either a PDA (equipped with a WLAN interface) or a Bluetooth mobile phone. The system is designed to be integrated discretely in the bus service via ubiquitous computing principles. It tries to minimize both the amount of required changes in the service operation, and explicit interactions with the mobile device. This is done by augmenting real-life interactions with data processing, through a programming paradigm called spatial programming.

2.4 Tracking and Registration for Mobile AR

AR requires very accurate position and orientation tracking in order to align, or register, virtual information with the physical objects that are to be annotated. Without this, it is rather difficult to trick the human senses into believing that computer-generated virtual objects co-exist in the same physical space as the real world objects. There are several possibilities for classifying tracking methods. First,

technological characteristics can be used to differentiate between the approaches. Another criterion is the applicability in different environments like indoor or outdoor, or the granularity of the determination of the position or the inclusion of the position together with the orientation within the physical space can be administered.
This section outlines the tracking strategies used in recent mobile AR systems. A wide range of both visual and non-visual tracking technologies, such as magnetic and ultrasound, have been applied to AR as already described in recent surveys from Azuma et al [1], Azuma et al [9] and Hollerer [8]. More specifically, Piekarksi [50] provides an extensive comparison of recent tracking methods and sensors, categorized based on their a) accuracy, b) resolution, c) delay, d) update rate, e) Infrastructure and operating range f) cost g) degrees of freedom and h) portability and electrical power consumption. However, the low cost of video cameras and the increasing availability of video capture capabilities in off-the-shelf PCs has inspired substantial research into the use of video cameras as means for tracking the position and orientation of a user (Klein [21]). A recent comparison from DiVerdi et al [53] has been adapted for our survey and presented in Table 1.

2.4.1 Tracking with GPS, GSM, UMTS

Probably, the most predominant system for outdoors tracking is the Global Positioning System (GPS) (Intermedia [16]). GPS is a time measurement based system and can be applied in almost all open space environments except narrow streets or covered sight to the sky due to trees or other obstacles to receive the signals from at least 4 satellites. The accuracy of the localization can vary between 3 and 10 meters depending on the satellite connection and the continuity of the navigation of the receiver. The accuracy can be increased by so called Differential GPS (D-GPS) by terrestrial stations to an accuracy of 2 to 5 meters. GPS receivers are becoming less and less expensive as they are introduced in mass-market devices such as PDAs and mobile phones. Schmeil et al [40] employed standard GPS for outdoors location tracking and a 3-DOF orientation tracker mounted on the HMD for orientation tracking and registration of a virtual guide on the real outdoors environment. Azuma et al [19] presented a method to enhance the position tracking accuracy of GPS, for more accurate and believable registration for mobile AR systems. They propose a new hybrid tracking system of improved accuracy for military operations, where an AR helmet has three rate gyroscopes, two tilt sensors, a GPS sensor and an infrared camera that occasionally observes small numbers of mobile infrared beacons added to the environment which help to significantly correct the sensor errors.
Another upcoming solution is locating users by triangulating signals of their GSM mobile phones. However, the accuracy of this localization method is quite crude and subject to huge variations. In particular in rural areas with wide phone ID cells the accuracy is not acceptable. With the advent of the third generation mobile standard UMTS, the accuracy of localization will improve significantly.

2.4.2 Outside-in and Inside-out Tracking

Tracking a user with an external camera is an example of outside-in tracking, where the imaging sensor is mounted outside the space tracked. Outside-in tracking can be used to produce very accurate position results - especially when multiple cameras observe the tracked object. In inside-out systems, the imaging sensor is itself head-mounted and any rotation of the user's head causes substantial changes in the observed image. Klein [21], Hollerer [8] and Azuma et al [1] provide a comprehensive overview of latest inside-out as well as outside-in tracking methods, not limited only to mobile AR systems.

2.4.2.1 Visual Marker-based tracking

A still common approach for more demanding augmented reality applications is to make use of fiducials: easily recognizable landmarks such as concentric circles placed in known positions around the environment. Such fiducials may be passive (e.g., a printed marker) or active (e.g., a light-emitting diode); both types of fiducial have been used in AR applications. While many passive fiducial-based tracking implementations for AR exist, none can match the ubiquity of the freely available ARToolkit system. Tamura et al [28], Billinghurst et al [35], Wagner et al [25], Newman et al [46], Henrysson et al

[22], Cheok et al [27], Barakonyi et al [33] and Goose et al [15] have employed various versions of ARToolkit in the range of mobile devices that are covered in Section 2.1. Makri et al [49] have employed their own visual marker-based tracking methods on mobile AR systems.

2.4.2.2 Visual Markerless tracking

A number of recent visual tracking algorithms as described in Klein [21] can provide realistic real-time camera tracking based on different approaches (natural feature detection, edge detection, planar methods etc.) but require large amounts of processing power posing difficulties on the additional AR rendering tasks.

Stork et al [26] employed a planar surface tracking algorithm, where 3D planes of building facades are used to recover the camera pose, by tracking natural features extracted from them. Vacchetti et al [13] employ another markerless tracking algorithm suitable for mobile AR, where it starts from 2D matching of interest points, and then it exploits them to infer the 3D position of the points on the object surface. Once the 3D points on the object are tracked it is possible to retrieve the camera displacement in the object coordinate system using robust estimation.

In Papagiannakis et al [12], a robust markerless real-time visual tracking method was introduced based on the Boujou system from 2D3™ which can recover from complete occlusion or extreme motion blur within one frame. At a first pre-processing stage of the scene-to-be-tracked, a model of the scene which consists of 3D coordinates together with invariant descriptors for the feature appearances is automatically created based on Structure-from-Motion techniques. During real-time operation, this database is traversed and compared against the real-time detected scene features, providing the estimated camera matrix.

Klein et al [20] employed an edge-based tracking system mounted on a tablet-PC for visual inside-out tracking. The tracking system employed relies on the availability of a 3D model of the scene to be tracked. This 3D model should describe all salient edges and any occluding faces. Using a predicted estimate of the camera pose, an estimate of the tablet camera's view of the model can be recovered at each frame.

2.4.3 Sensor based tracking

Infra-red LEDs can output light across a very narrowly tuned wave-band, and if this band is matched at the sensor with a suitable filter, ambient lighting can be virtually eliminated. This means the only thing visible to the imaging sensor is the fiducials, and this vastly reduces the difficulty and computational requirements for tracking. For this reason, LEDs have long been used in commercially available tracking systems and real tracking applications; Olwal et al [17] employed IR-LEDs for robust tracking of mobile phones, based on the vision of spatially aware handheld interaction devices. Based on outside-in tracking methods they allowed for the augmentation of a real map with digital content in a focus + context fashion.

RFID (radio frequency identification) tags have also been recently used in mobile AR systems. RFIDs consist of a simple microchip and antenna which interact with radio waves from a receiver to transfer the information held on the microchip. RFID tags are classified as either active or passive, with active tags having their own transmitter and associated power supply, while passive tags reflect energy sent from the receiver. Active RFID tags can be read from a range of 20 to 100m where passive RFID tags range from a few centimeters to around 5m (depending on the operating frequency range). Rashid et al [30] employed mobile phones that incorporate RFID readers for creating games in which players interact with real physical objects, in real locations.

A recent promising technology for wide-area indoor tracking is the commercially available Ultra-Wide-Band (UWB) local positioning system by Ubisense™ (Steggles et al [52]). Based on network of small-size sensors and tags this system allows for estimating the 3D position of a tag within 15cm accuracy of tens of meters distance of a tag from a sensor.

2.4.4 Wireless-LAN tracking

Due to the fact that networked mobile AR users are enabled with wireless radio communication network interfaces (such as Wi-Fi), protocols that provide location estimation based on the received signal strength indication (RSSI) of wireless access points have been recently becoming increasingly accurate, sophisticated, and hence, popular. The main benefit of RSSI measurement-based systems is that they do not require any additional sensor/actuator infrastructure but use already available communication parameters and downloadable wireless maps for the position determination. Their shortcoming for mobile AR is precision and often multiple access points as well as tedious training offline phases for the construction of the wireless map. Peternier et al [24] employed a WiFi-localization based method for a PDA-based Mixed-Reality system for visualizing virtual character 3D content. Liu et al [51] describe a WiFi-localization algorithm based on a single access-point infrastructure as navigational aid.

2.4.5 Hybrid tracking systems

The use of inertial sensors such as rate gyroscopes and accelerometers is wide-spread in virtual and augmented reality applications. Visual tracking systems perform best with low frequency motion and are prone to failure, especially given rapid camera movements, such as may occur with a head-mounted camera. Inertial sensors measure pose derivatives and are better suited for measuring high-frequency, rapid motion than slow movement where noise and bias drift outweigh. The complementary nature of visual and inertial sensors has led to the development of a number of hybrid tracking systems. Their ultimate goal is "anywhere augmentation" as specified in DiVerdi et al [53], where their hybrid tracking system for mobile AR consists of a camera facing the ground and orientation tracker inspired by the workings of an optical mouse, providing very accurate tracking resolution both indoors/outdoors in the expense of specific h/w setup. In order to tackle efficiently the mobility and robustness of an indoor mobile AR-system, an aggregation of tracking sensors/methods such as UWB position and accelerometer orientation sensors together with fiducial markers was exhibited by Newman et al [46].

2.5 Displays

There are many approaches to displaying information to a mobile person and a variety of different types of displays can be employed for this purpose, such as, personal hand-held, wrist-worn, or head-worn displays; screens and directed loudspeakers embedded in the environment; and, image projection on arbitrary surfaces; to name but a few. Several of these display possibilities may also be used in combination. Augmented reality displays utilized in recent mobile AR systems can be fundamentally split into two categories: optical see-through displays with which the user views the real world directly (such as Micro-Vision Nomad, TekGear Icuiti or EyeTop), and video see-through displays with which the user observes the real world in a video image as acquired from a mounted camera (such as Trivisio AR-Vision and i-glasses PC). There are various issues associated with both types of displays as recently reviewed by Piekarksi [50] such as a) technological: latency, resolution-distortion, field of view and cost, as well as b) perceptual: depth of field, qualitative, and finally, c) human factors: social acceptance and safety. One of the current trends for mobile AR is the fusion of different display technologies with wearable computing (Hollerer [8],). Head-worn displays provide one of the most immediate means of accessing graphical information since the viewer does not need to divert his or her eyes away from their object of focus in the real world and if they are worn as part of a wearable system (i.e. not as part of a helmet) can be even assume social acceptance beyond the AR prototype stage. The immediacy and privacy of a personal head-worn display is complemented well by the high text readability of hand-held displays in collaboration with wall-sized displays. The attractiveness of mobile AR relies on further progress in this area, as for example, the new dedicated for AR OLED based HMDs that appear in Stork et al [26]) as well as Makri et al [49] by Trivisio™.

2.6 Wearable input and interaction technologies

Piekarksi [50] defines a wearable computer to be a self powered computing device that can be worn on the body without requiring the hands to carry it, and can be used while performing other tasks. It should thus be worn like a piece of clothing, as unobtrusive as possible. Key factors amongst others are comfort, weight, size, mobility, and aesthetics. How to interact with wearable computers effectively and efficiently is an area of active research. Mobile interfaces should try to minimize the burden of encumbering interface devices. The ultimate goal is to have a free-to-walk, eyes-free, and hands-free interface with miniature computing devices worn as part of the clothing. This ideal wearable mobile AR system cannot be reached yet with current mobile computing and interface technology. Hollerer [8], Piekarksi [50], Kölsch et al [54] and Revenier et al [42] provide overviews of recent approaches on multimodal input and interaction technologies for wearables which extends the scope of the current survey. Readers are encouraged to refer to these works for further details. Some devices already nicely meet the size and ergonomic constraints of mobility: auditory interfaces for example, can already be realized in a inconspicuous manner, with small wireless hands-free earphones and microphones are barely noticeable. AR characters are envisaged to assume new roles as central interfaces on such wearable systems, based on natural language-based communication and interaction patterns similar to the ones observed between real humans.

Furthermore, recent clothes manufacturers have gradually started creating clothes with special focus on allowing the efficient incorporation, carriage and operation of mobile devices such as iPods, PDAs or SmartPhones such as the jackets from Zegna™ (i-jacket) or Scottvest™ (v3 jacket). As a wearable computer is a very personal device, it should be worn like a piece of clothing, as unobtrusive as possible so that a user could interact with this computer based upon context. Mann [4] in his early vision of wearable computing presents the miniaturization of hardware components and their fusion with clothing, as the main factor allowing individuals to freely move about and interact supported by their personal domain. In the forthcoming mobile AR networked media environments, overcoming information overload and access complexity via natural conversation with wearable, incorporated in everyday clothes companions can play decisive roles as new personalized information interfaces.

3. Mobile AR: Applications and Challenges

This section presents the recent advances as well as new additions to the applications areas where mobile AR systems are used. This is not an extensive chronological list as we mainly aim to complement the most recent surveys from Azuma et al [1], [9] studying the convergence of the AR, ubiquitous and wearable computing. The main mobile AR applications studies that this survey covers are:

- Virtual Character-based applications for AR
- Cultural Heritage
- Edutainment and Games
- Navigation and Path-Finding
- Collaborative assembly and design
- Industrial maintenance and inspection

The above constitute research directions that are already identified and aimed to drive further the area of augmented ubiquitous computing. In this manner, augmented ubiquitous computers will help overcome the problem of information overload. There is more information available at our fingertips during a walk in the woods than in any computer system, yet people find a walk among trees relaxing and computers frustrating (Weiser [8]). Machines that fit the human environment, instead of forcing humans to enter theirs, will make using a computer as refreshing as taking a walk in the woods.

3.1 Virtual Characters in AR

Virtual Characters have already been synthesized with real actors in common non-real-time mixed reality worlds, as illustrated successfully by a number of cinematographic storytelling examples. One of the earliest research-based examples was and the virtual 'Marilyn Monroe' as appearing in the research film "Marilyn by the lake" by MIRALab, University of Geneva as well as Balcisoy et al [34]. However these 'compositing' examples involve non-real-time (offline) pre-rendered simulations of virtual characters and mostly are rendered and post-processed frame by frame in an ad-hoc manner by specialized digital artists or compositors as they are termed in the industrial domain of special effects (SFX). The active SFX sector with applications in film, television and entertainment industry has exemplified such compositing effects in a constantly growing list of projects.

In this survey we study the recent surge of employing virtual characters in mobile AR systems (Figure 5). A first such real-time example in a mobile setup employed virtual creatures in collaborative AR games (Hughes et al [32]) as well as a conversational and rigid-body animated characters, during a construction session (Tamura et al [28], Cheok et al [27]). In cultural heritage sites, a recent breed of mobile AR systems allows for witnessing ancient virtual humans with body animation, deformation and speech, re-enacting specific context-related scenarios (Papagiannakis et al [12]) as well as allowing visitors to interact and further enquire on their storytelling experience (Egges et al [11]). An important aspect of such AR examples is that these virtual characters are staged in scenario-based 'life-size' scaling, position orientation as a result of markerless AR tracking and registration. Further recent examples of marker-based tracking such as ARToolkit, various researchers employed dynamic content on top of such markers, such as 3D storytelling book content (Billinghurst et al [35]) and other interactive characters reacting to user's actions (Barakonyi et al [33], Wagner et al [39]). Very recent examples include also the use of virtual characters as outdoor navigation guides (Schmeil et al [40]).

3.2 Cultural heritage

Mobile AR systems are increasingly being tested in rich content environments, as they can enable visualization of 'unseen' valuable and complex 3D content as well as provide added edutainment-value in today cultural heritage sites. The shift that the cultural heritage sector is currently facing in its traditional economic paradigm combined with the increasing digitization efforts allow for AR interfaces to be used as ideal and novel showcases of both tangible (objects, static edifices) and intangible (ceremonies, customs, myths) cultural artifacts. In particular, mobile AR guides were employed in the site of ancient Olympia, Greece in order to visualize the non-existing ancient temple edifices (Vlahakis et al [36]), and in Pompeii, Italy to visualize ancient Roman characters reenacting stories based on the site frescoes (Papagiannakis et al [12]). **Error! Reference source not found.** and **Error! Reference source not found.** illustrate examples of a mobile AR system in ancient Pompeii.

3.3 Navigation & Path Finding

Mobile AR systems have also been widely employed for mobile navigation assistance. Such systems typically involve a hardware device as described in Section 2.1 and based on an AR platform similar to those as described in Section 2.2, they allow for multimodal navigation AR aid while traversing physical buildings or outdoor locations. As shown in **Error! Reference source not found.**, different approaches are followed based primarily on whether indoors or outdoors AR navigation is needed. Hollerer [8], Elmqvist et al [29], Olwal et al [17] and Newman et al [46] work indoors while Bell et al [14], Reitmayr et al [18] and Azuma et al [19] are employed outdoors. Absolute tracking and registration remains still an open issue and recently it has mostly been tackled by no single method, but mostly with an aggregation of tracking and localization methods, mostly based on handheld AR. A truly wearable, HMD based mobile AR navigation aid for both indoors and outdoors with rich 3D content remains still an open issue and a very active field of multi-discipline research.

3.4 Edutainment & Games

Magerkurth et al [37] presents an overview of pervasive gaming containing a section on AR Games (Figure 6). Recently AR multi-user games appeared based on generic AR frameworks (Wagner et al [38]). Traditional 2D games also find their application in mobile AR, based on the well-known 'Pac-Man' gaming genre (Cheok et al [27], Rashid et al [30] and Klein et al [20]). Mobile phones have also been used as kineasthetic AR interfaces in an AR tennis game Henrysson et al [22]. Based on the "Shoot'em up" computer gaming genre, several AR games have been realized using mobile AR systems, such as those described in Hughes et al [32] and Piekarksi et al [50]. The main unsolved research issues include multiple, interactive virtual characters in AR, common-vision collaborative games as well as convincing illumination registration and real-time rendering.

3.5 Collaborative assembly and construction

One of the main features of mobile collaborative AR is that the augmentation of the real-world is adapted according to multiple-user location and knowledge. Renevier et al [42] exhibited such a mobile AR system for both archaeological field work as well as asynchronous collaborative game. Furthermore, the industrial CAD design field has also recently benefited from mobile AR systems (Stork et al [26]) allowing multiple users to reviews complex 3D CAD architectural or automotive industry models. Finally in the filed of on-site collaborative AR and construction, Piekarksi et al [50] employed their generic AR software framework for novel 3D construction on real sites.

3.6 Maintenance and inspection

On of the early mobile systems employed on industrial maintenance or inspection for service personnel as well as consumers were introduced by Gausemeier et al [47]. It involved a PDA-based handheld AR solution where the inside-out visual tracking was distributed on a stationary network server and the final augmentation transmitted to the mobile device. In the same domain the approach from Vacchetti et al [13] allowed for interactive virtual humans augmenting the real-scene in order to provide new training capabilities for professionals, beyond the traditional video material approach, however based on a mobile workstation so that tracking, registration and simulation are performed in real-time on site. On-site mobile AR augmentation for industry professionals was also implemented on PDA based on mobile inside-out visual tracking as handheld as well as HMD AR by Goose et al [15], Makri et al [49]. Recently a handheld AR interface to a sensor network has been proposed by Rauhala et al [48] based on SmartPhones for on-site humidity inspection.

4. Discussion and Conclusions

The following Table 3 exhibits complete AR Systems as employed in different application areas.
In this table the abbreviations shown in Table 2 were used.

From the above comparison it is clear that there is no single ideal mobile AR system approach but rather different AR systems according to location (indoors or outdoors), type of display (handheld or hmd), content augmentation (static 3D, virtual characters) as dictated by each application domain.

Of course still important challenges lie in all three areas such as highlighted by the reviewed literature:

- Limited computational resources: While mobile devices are transforming from simple communication or dedicated multimedia devices to more powerful computational platforms, there will still be need for more computing power. More powerful processor chips are available but their high power consumption and high heat generation present challenges for their use in mobile platforms.
- Size, weight: wearable AR systems should not be a burden but as unobtrusive as possible
- Battery life: an important factor of the sustainability of the above AR applications. Except the smart phone category, most other devices suffer significantly from this aspect, limiting the AR application to few hours. Components that have large electrical energy requirements need more batteries, which adds to both weight and size. The power consumption of devices is directly proportional to clock speed and heat dissipation.
- Ruggedness: all above mentioned mobile AR systems are early prototypes and depending on the display setup (handheld or HMD), device materials, cables, connectors and cases normally used

indoors may be unsuitable outdoors. Sensitive electronic equipment needs to be protected from the environment or it will be damaged easily.

- Tracking and Registration: these are the basic components of a mobile AR system as specified before. The current trend shows that a combination of tracking modalities is employed for best results, such as, vision or sensor based. This helps avoid the problems of a single tracking approach. Since different methods are employed for indoors or outdoors tracking, the target of supporting unprepared environments is still elusive.
- 3D graphics and real-time performance: one of he limiting factors for rich mobile content was the absence of dedicated 3D processing units in mobile devices. Nowadays more such as functionality is incorporated into such devices as well as the emergence of new powerful small-factor PCs (such as UMPC) allow for new possibilities in more compelling AR content and applications.
- Social acceptance and mobility: the miniaturization of devices as well as their aggregation into wearable systems will contribute to gathering social acceptance momentum. The current AR prototypes are quite bulky or intrusive. Another important factor is the evolution of mobile phones where very soon GPS, graphics acceleration and other sensors would allow them to be fully transformed into first class mobile AR computational systems. Thus mobility is tightly intertwined with social acceptance of future mobile AR systems
- Networked Media: with the increasing expansion in bandwidth new breed of audiovisual networked media applications are envisaged and mobile AR systems can profit significantly. However, issues such as content adaptation (Singh [73]) and sharing, user modeling and personalized interfaces will also need to be addressed for compelling AR applications

The recent advances in mobile augmented ubiquitous interfaces are envisaged to be built to emulate and extend basic biological principles and communication patterns. Multimodal interfaces which enable multisensory perception through fusion of different information sources using embedded computing capacity are the ultimate goal. Such radical research in new fusion patterns between AR, wearables and ubiquitous computing can lead to new ways of unsupervised audiovisual common-cause techniques for perception of coverage, context awareness and improved and augmented visual performance in a variety of tasks and take advantage of such new 'cyborgian' interfaces (Clynes [56]). The cyborgian mode of interaction manifestates itself when a human and other external process interact in a manner that does not require conscious thought or effort. A person riding a bicycle is one cyborgian example, since after time the rider operates the machine (bicycle) without conscious thought or effort, and, in some sense, the machine becomes an extension of the wearer's own body. Sometime soon we should such compelling examples involving mobile augmented ubiquitous interfaces.

5. Acknowledgments

The provided work has been partially supported by the EU FP6 IST Programme, in the framework of the INTERMEDIA Network of Excellence Project IST-38419.

6. References

[1] Sutherland, I., "A head-mounted three dimensional display", In *Proc. FJCC 1968*, pages 757–764, Washington, DC. Thompson Books, 1968

[2] Caudell, T. P. and Mizell, D. W., "Augmented reality: An application of heads-up display technology to manual manufacturing processes", In *Proceedings of 1992 IEEE Hawaii International Conference on Systems Sciences*. IEEE Press, 1992

[3] Feiner, S., MacIntyre, B., Hollerer, T., Webster, A., “A touring machine: Prototyping 3D mobile augmented reality systems for exploring the urban environment”, In *Proc. ISWC ’97 (First Int. Symp. on Wearable Computers)*, pages 74–81, Cambridge, MA, 1997

[4] Mann, S., "Wearable Computing: A First Step Toward Personal Imaging", IEEE Computer, February 1997

[5] Wellner, P., Mackay, W., Gold, R., “Back to the real world”, Communications of the ACM, Volume 36, Issue 7, pp. 24-27, 1993

[6] Milgram, P., Colquhoun, H., “A taxonomy of real and virtual world display integration”, In Y. O. H. Tamura (Ed.), Mixed Reality: Merging Real and Virtual Worlds. (pp. 5-30). Springer-Verlag, 1999

[7] Azuma, R., Baillot, Y., Behringer, R., Feiner, S., Julier, S., MacIntyre, B., “Recent Advances in Augmented Reality”, IEEE Computer Graphics and Applications, Volume 21, Issue 6, pp. 34-47, November/December 2001

[8] Weiser, M., “The computer for the 21st century”, *Scientific American*, 265(3):94–104, 1991

[9] Azuma, R., Billinghurst, M., Schmalstieg, D., Hirokazu, K., "Developing Augmented Reality Applications", ACM SIGGRAPH 2004 Course Notes, 2004

[10] T. Hollerer, “User Interfaces for Mobile Augmented Reality Systems”, Columbia University, PhD Thesis, 2004

[11] Egges, A., Papagiannakis, G., Magnenat-Thalmann, N., “Presence and Interaction in Mixed Realities”, *The Visual Computer, Springer-Verlag (accepted)*, 2007

[12] G. Papagiannakis, S. Schertenleib, B. O’Kennedy , M. Arevalo-Poizat, N. Magnenat-Thalmann, A. Stoddart, D. Thalmann, “Mixing Virtual and Real scenes in the site of ancient Pompeii”, *Computer Animation and Virtual Worlds, p 11-24, Volume 16, Issue 1*, John Wiley and Sons Ltd, February 2005

[13] L. Vacchetti, V. Lepetit, M. Ponder, G. Papagiannakis, P. Fua, D. Thalmann,N. Magnenat-Thalmann, “Stable Real-time AR Framework for Training and Planning in Industrial Environments”,*Virtual Reality and Augmented Reality Applications in Manufacturing, Ong, Soh K., Nee, ISBN: 1-85233-796-6, Springer-Verlag, A.Y.C.*. May 2004

[14] B. Bell, S. Feiner, T. Hollerer, “Visualization viewpoints”, *IEEE Computer Graphics & Applications*, July/August 2002

[15] S. Goose, S. Guven, X. Zhang, S. Sudarsky, N. Navab, “Paris: fusing vision-based location tracking with standards-based 3d visualization and speech interaction on a PDA”, *Proc. IEEE DMS 2004 (International Conference on Distributed Multimedia Systems)*, p 75-80, San Francisco, CA, September 8-10, 2004.

[16] Intermedia D3.1.a: Draft on State-of-the-Art Report on Approaches Toward User-Centric Multimedia, INTERMEDIA NoE, http://intermedia.miralab.unige.ch, January 2007

[17] Olwal, A., “LightSense: Enabling Spatially Aware Handheld Interaction Devices”, *Proceedings of ISMAR 2006 (IEEE and ACM International Symposium on Mixed and Augmented Reality)*, Santa Barbara, CA, Oct 22-25, 2006, pp 119-122

[18] G. Reitmayr, T. Drummond, “Going out: Robust Model-based Tracking for Outdoor Augmented Reality” , *Proceedings of ISMAR 2006 (IEEE and ACM International Symposium on Mixed and Augmented Reality)*, Santa Barbara, CA, Oct 22-25, 2006, pp 119-122

[19] R. Azuma, J. Leonard, “Performance Analysis of an Outdoor Augmented Reality Tracking System that Relies Upon a Few Mobile Beacons”, *Proceedings of ISMAR 2006 (IEEE and ACM*

International Symposium on Mixed and Augmented Reality), Santa Barbara, CA, Oct 22-25, 2006, pp 119-122

[20] G. Klein and T. Drummond, "Sensor Fusion and Occlusion Refinement for Tablet-based AR", *in Proc. Third IEEE and ACM International Symposium on Mixed and Augmented Reality,* ISMAR'04, Arlington, 2004

[21] G. Klein, "Visual Tracking for Augmented Reality", University of Cambridge, PhD Thesis, 2006

[22] A. Henrysson, M. Billinghurst, M. Ollila,"Face to Face Collaborative AR on Mobile Phones", *Proceedings of ISMAR 2005 (IEEE and ACM International Symposium on Mixed and Augmented Reality)*, 2005

[23] W. Piekarski, R. Smith, B. H. Thomas, "Designing Backpacks for High Fidelity Mobile Outdoor Augmented Reality", *in Proc. Third IEEE and ACM International Symposium on Mixed and Augmented Reality,* ISMAR'04, Arlington, 2004

[24] A. Peternier, F. Vexo, D. Thalmann,"Wearable Mixed Reality System In Less Than 1 Pound", *Eurographics Symposium on Virtual Environments' 06*, 2006

[25] D. Wagner, D.Schmalstieg, "ARToolKitPlus for Pose Tracking on Mobile Devices", *Computer Vision Winter Workshop 2007,* Austria, February 6–8, 2007

[26] A. Stork, P. Santos, T. Gierlinger, A. Pagani, C. Paloc, I. Barandarian, G. Conti, R. Amicis, M. Witzel, O. Machui, J. Jiménez, B. Araújo, J. Jorge, G. Bodammer, "IMPROVE: An innovative application for collaborative mobile mixed reality design review*", Proceedings of Virtual Concept '06*, 2006

[27] A. Cheok, S. Fong, K. Goh1, X. Yang, W. Liu, F. Farzbiz, Y. Li, "Human Pacman: A Mobile Entertainment System with Ubiquitous Computing and Tangible Interaction over a Wide Outdoor Area", *Human-Computer Interaction with Mobile Devices and Services, Lecture Notes in Computer Science,* Springer Berlin / Heidelberg, 2003

[28] Tamura, H., Yamamoto, H., Katayama, A., "Mixed Reality: Future Dreams Seen at the Border between Real and Virtual Worlds," *IEEE Computer Graphics and Applications*, pp. 64-70, November/December 2001

[29] Elmqvist, N., Axblom, D., Claesson, J., Hagberg, J., Segerdahl, D., So, Y., Svensson, A., Thoren, M., Wiklander, M., "3DVN: A Mixed Reality Platform for Mobile Navigation Assistance", Technical Report no. 2006-12, Chalmers University of Technology and Goteborg University, 2006

[30] Rashid, O., Bamford, W., Coulton, P., Edwards, R., Scheible, J., "PAC-LAN: Mixed-Reality Gaming with RFIDEnabled Mobile Phones", ACM Computers in Entertainment, Vol. 4, No. 4, October 2006

[31] Zauner, J., Haller, M., "Authoring of Mixed Reality Applications including Multi-Marker calibration for Mobile Devices", in 10th Eurographics Symposium on Virtual Environments, EGVE 2004, pp. 87-90

[32] Hughes, C. E., Stapleton, C. B., Hughes, D. E. Smith, E., "Mixed Reality in Education, Entertainment and Training: An Interdisciplinary Approach," IEEE Computer Graphics and Applications, 26(6), November/December 2005, 24-30

[33] Barakonyi, I., Schmalstieg, D., "Ubiquitous Animated Agents for Augmented Reality", *Proceedings of ISMAR 2006 (IEEE and ACM International Symposium on Mixed and Augmented Reality)*, Santa Barbara, CA, Oct 22-25, 2006

[34] Balcisoy, S., Torre, R., Ponder, M., Fua, P., Thalmann, D., "Augmented Reality for Real and Virtual Humans", Proc. CGI00, IEEE Computer Society Press, 2000

[35] Billinghurst, M., Kato, H., Poupyrev, I., “The MagicBook – Moving Seamlessly between Reality and Virtuality”, IEEE Computer Graphics and Applications, Volume 21, Issue 3, pp. 6-8, May/June 2001

[36] Vlahakis, V., Ioannidis, N., Karigiannis, J., Tsotros, M., Gounaris, M., Stricker, D., Gleue, T., Daehne, P., Almeida, L., “Archeoguide: An Augmented Reality Guide for Archaeological Sites”, IEEE Computer Graphics and Applications, vol. 22(5), pages 52-60, 2002

[37] Magerkurth, C., Cheok, A. D., Mandryk, R.L., Nilsen, T., ”Pervasive games: bringing computer entertainment back to the real world”, Computers in Entertainment 3(3): 4, 2005

[38] Wagner, D., Schmalstieg, D., “Muddleware for Prototyping Mixed Reality Multiuser Games”, IEEE VR2007

[39] Wagner, D., Billinghurst, M., Schmalstieg, D., “How real should virtual characters be?”, Conference on Advances in Computer Entertainment Technology(ACE) 2006

[40] Schmeil, A., Broll, W., "MARA - An Augmented Personal Assistant and Companion", ACM SIGGRAPH 2006 Sketch, 2006

[41] Ohlenburg, J., Herbst, I., Lindt, I., Fröhlich, T., Broll, W., “The MORGAN Framework: Enabling Dynamic Multi–User AR and VR Projects”, in Proc. of the ACM Symposium on Virtual Reality Software and Technology (VRST 2004), pp. 166-169, ACM, 2004

[42] Renevier P., Nigay L, Bouchet J., Pasqualetti L., “Generic Interaction Techniques for Mobile Collaborative Mixed Systems”, Fifth International Conference on Computer-Aided Design of User Interfaces CADUI'2004, ACM, Funchal, 2004, pp 307-320

[43] Ponder, M., Papagiannakis, G., Molet, T., Magnenat-Thalmann, N., Thalmann, D., ”VHD++ Framework: Extendible Game Engine with Reusable Components, for VR/AR R&D featuring Advanced Virtual Character Simulation Technologies”, Proceedings of Computer Graphics International03, pp. 96-104, IEEE Computer Society Press, 2003

[44] Papagiannakis, G., Schertenleib, S., Ponder, M., Arevalo, M., Magnenat-Thalmann, N., Thalmann, D., “Real-Time Virtual Humans in AR Sites”, Proc. of IEE Visual Media Production04, pp. 273-276, IEE Society, 2004

[45] Schmalstieg, D., Fuhrmann, A., Hesina, G., Szalavári, Zs., Encarnação, M., Gervautz, M., Purgathofer, W., “The Studierstube Augmented Reality Project”, PRESENCE -Teleoperators and Virtual Environments, MIT Press, 2002

[46] Newman, J., Schall, G., Barakonyi, I., Schürzinger, A., Schmalstieg, D., "Wide-Area Tracking Tools for Augmented Reality",Advances in Pervasive Computing 2006, Adjunct Proceedings of Pervasive 2006, Dublin, 2006

[47] J. Gausemeier, J. Fründ, C. Matysczok, B. Brüderlin, D. Beier, "Development of a Real Time Image Based Object Recognition Method for Mobile AR-Devices," in proceedings of the ACM SIGGRAPH AFRIGRAPH 2003: 2nd International Conference on Computer Graphics, Virtual Reality, Visualisation and Interaction in Africa, Cape Town, Feb. 3-5, 2003

[48] Rauhala, M., Gunnarsson, A., Henrysson, A., Ynnerman, A., "A Novel Interface to Sensor Networks using Handheld Augmented Reality", ACM MobileHCI06, 2006

[49] Makri, A., Arsenijevic, D., Weidenhausen, J., Eschler, P., Stricker, D., Machui, O., Fernandes, C., Maria, S., Voss, G., Ioannidis, N., "ULTRA: An Augmented Reality system for handheld platforms, targeting industrial maintenance applications", 11th International Conference on Virtual Systems and Multimedia, Ghent, Belgium, 2005

[50] Piekarski, W. and Thomas, B. H. An Object-Oriented Software Architecture for 3D Mixed Reality Applications. In 2nd Int'l Symposium on Mixed and Augmented Reality, Tokyo, Japan, Oct 2003

[51] Liu, A. L., Hile, H., Kautz, H., Borriello, G., Brown, P. A., Harniss, M., Johson, K., "Indoor Wayfinding: Developing a Functional Interface for Individuals with Cognitive Impairments", ACM ASSETS06, 2006

[52] Steggles, P., Gschwind, S., "The Ubisense Smart Space Platform: A Ubisense White Paper", http://www.ubisense.net/SITE/UPLOAD/Document/WhitePapers/Ubisense_a_smart_space_platform.pdf, retrieved at 01/04/07, 2007

[53] DiVerdi, S., Höllerer, T., "GroundCam: A Tracking Modality for Mobile Mixed Reality," IEEE Virtual Reality, Mar. 10-14, 2007

[54] Kölsch, M., Bane, R., Höllerer, T., Turk, M., "Multimodal interaction with a wearable augmented reality system," IEEE Computer Graphics and Applications, Vol. 26, No. 3, pp. 62 -71, May/June 2006.

[55] Wellner, P., Mackay, W., Gold, R., "Back to the real world", Communications of the ACM, Volume 36, Issue 7, pp. 24-27, 1993

[56] Clynes, M., "Cyborg II: Sentic Space Travel", *The Cyborg Handbook*, Chris Hables Gray, Steven Mentor, Heidi J. Figueroa-Sarriera (eds), Routledge, October 1995

[57] Das, T., Singh, G., Mitchell, A., Kumar, P.S., McGee, K., "Developing Social Virtual Worlds using NetEffect", Proc. WET-ICE (Workshop on Enabling Technologies on Infrastructure for Collaborative Enterprises), Cambridge, MA, June 18-20, 1997, pp:148-154.

[58] Goseta, M., Cipar, I., Mikic, D., "Exeprimental Performance Evaluation of Networked Virtual Reality Services in UMTS Network", Proc. IEEE MELECON, Dubrovnik, Croatia, May 12-15, 2004, pp: 661- 664

[59] Perovic, I., Skorin-Kapov, L., VilendeEiC, D., "Performance Evaluation of Networked Virtual Reality Services in a GPRS Network", Proceedings of the 7th International Conference on Telecommunications - ConTEL 2003, June 11-13 2003, Zagreb, Croatia, pp: 359-366

[60] Jonietza, E., "10 Emerging Technologies – Augumented Reality", Technology Review, March/April 2007, available at http://www.technologyreview.com/read_article.aspx?ch=specialsections&sc=emerging&id=18291 (accessed on June 18, 2007)

[61] Smed, J., Kaukoranta, T., Hakonen, H., "Aspects of Networking in Multiplayer Computer Games", Proc., International Conference on Application and Development of Games in the 21st Century, 2001

[62] Pantel, L., Wolf, L.C., "On the Impact of Delay on Real-time Multiplayer Games". Proc. NOSSDAV'02, May 12-14, 2002, Miami, Florida, pp: 23-29.

[63] Torabi, M., "Mobile Virtual Reality Services". Bell Labs Technical Journal. 7(2), Summer 2002, pp: 185-193

[64] Viktorsson, P., Borg, K., "Method and devices for use in a virtual environment". US Patent number 6,397,080, Awarded: May 28, 2002

[65] Ferguson, T., "3G and Wi-Fi go head to head", CNET UK, June 19, 2007, http://news.cnet.co.uk/networking/0,39029686,49291134,00.htm (accessed on June 19, 2007).

[66] Cheok, A.D., et al., "Human Pacman: A Mobile Entertainment System with Ubiquitous Computing and Tangible Interaction over a Wide Outdoor Area", Proc. Mobile HCI 2004 – 6th International Symposium, Glasgow, UK, September 13 - 16, 2004, pp: 209-223.

[67] Banâtre, M., Couderc, P., Pauty, J. and Becus, M., "Ubibus: Ubiquitous Computing to Help Blind People in Public Transport", Proc. Mobile HCI 2004 – 6th International Symposium, Glasgow, UK, September 13 - 16, 2004, pp:310-314

[68] McCaffery, D.J. and Finney, J. "The need for real time consistency management in P2P mobile gaming environments". Proc. 2004 ACM SIGCHI International Conf of Advances in Computer Entertainment Technology – ACE'04. June 3-5, 2004, pp: 203-211.

[69] Mitchell, K., et al., "Six in the City: Introducing Real Tournament – A Mobile IPv6 based Context-Aware Multiplayer Game", Proc. 2nd Workshop on Network and System Support for Games, ACM Press, 2003, pp: 91-100.

[70] Watsen, K., Darken, D. P., Capps, W.V., "A Handheld Computer as an Interaction Device to a Virtual Environment," Third International Immersive Projection Technology Workshop (IPT 1999), Stuttgart, Germany, 1999.

[71] Hartling, P.L., Bierbaum, A.D., Cruz-Neira, C., "Tweek: Merging 2D and 3D interaction in immersive environments". 6th World Multiconference on Systemics, Cybernetics and Informatics (SCI 2002), 2002.

[72] Bierbaum, A. "VR Juggler: A Virtual Platform for Virtual Reality Application Development", Master's thesis, Iowa State University, 2000.

[73] Singh, G. (ed.) Content Repurposing. *IEEE Multimedia Systems*. 11(1), Jan-Mar 2004.

[74] Papagiannakis, G., Magnenat-Thalmann, N., "Mobile Augmented Heritage: Enabling Human Life in ancient Pompeii", International Journal of Architectural Computing, Multi-Science Publishing, April 2007 (accepted), issue 02, volume 05, pp.395-415, 2007

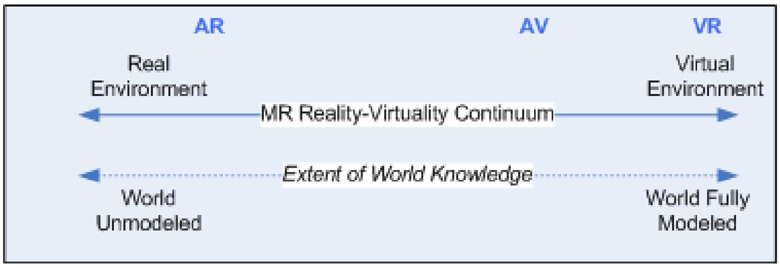

Figure 1 View of the MR Continuum as originally defined by Milgram et al [6]

Figure 2 Examples of Real, AR and VR environments from a mobile AR system (images courtesy of Papagiannakis et al [12])

Figure 3 Example of a mobile AR system based on a laptop, backpack, and video see-through HMD (images courtesy of Papagiannakis et al [74])

Figure 4 Examples of handheld mobile AR devices : UMPC, PDA and SmartPhone
(images courtesy of Wagner et al [25])

Figure 5 Real-time Virtual Characters in AR (Images Courtesy of Papagiannakis et al [74] (left) and Wagner et al [39] (right))

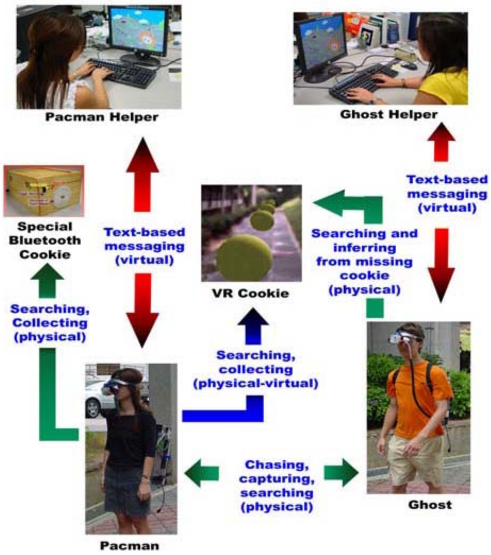

Figure 6 mobile AR Games(Images Courtesy of Cheok et al[27])

Table 1 Comparison of tracking technologies (adapted from DiVerdi et al [53])

Technology	Range (m)	Setup (hr)	Resolution (mm)	Time (s) (in which useful tracking occurs i.e. before drift)	Environment
Magnetic	1	1	1	∞	In / out
Ultrasound	10	1	10	∞	in
Inertial	1	0	1	10	In/out
Accelerometer	1000	0	100	1000	In/out
UWB	100	10	500	∞	in
Optical: outside-in	10	10	10	∞	in
Optical: marker-based	10	0	10	∞	In/out
Optical: markerless	50	0-1	10	∞	In/out
Hybrid	10	10	1	∞	in
GPS	∞	0	1000	∞	out
WiFi	100	10	1000	∞	In/out

Abbreviation	Meaning	Abbreviation	Meaning	Abbreviation	Meaning
I/O	*Indoors/ Outdoors*	IO	*Inside-Out*	OI	*Outside-In*
mb	*Marker-based*	Ml	*Marker-less*	UWB	*Ultra-wide band*
3Ddyn	*Dynamic (animated) 3D*	3Ddyn*	*Advanced Dynanic, Deformable 3D*	3Dstat	*Static 3D*
IR	*Infrared sensor*	Inert	*Inertial sensor*	PAN	*Personal Area Network*

Table 2 Abrevviations

Table 3 Mobile AR systems Comparison

Application Domain of mobile AR systems	Method	Computing Devices					Indoor/Outdoor, Wireless Networking			Tracking and Registration						Displays	AR Content
		MPC	TPC	UMPC	PDA	Phone	I /O	WLAN /gprs	PAN	gps	io	oi	sensors	uwb	wlan	Hmd / handheld	3Ddyn/ 3Dstat / 2D
Virtual Character based	Billinghurst et al [35]	x					I		x		mb					handheld	3Ddyn
	Tamura et al [28]	x					I		x		mb					hmd	3Ddyn
	Cheok et al [27]	x					O	x			mb					hmd	3Ddyn
	Papagiannakis et al [44]	x					I,O		x		ml					hmd	3Ddyn*
	Hughes et al [32]	x					I	x			mb					hmd	3Ddyn
	Barakonyi et al [33]	x			x		I	x			mb	x		x		handheld	3Ddyn
	Peternier et al [24]				x		I,O	x							x	hmd	3Ddyn
	Wagner et al [39]	x			x		I,O	x			mb	x				handheld	3Ddyn
	Scmeil et al [40]	x					O		x	x						hmd	3Ddyn*
	Egges et al [11]	x					I,O		x		ml					hmd	3Ddyn*
Navigation and Path finding	Bell et al [14]	x					O		x	x						hmd	2D
	Hollerer [8]	x					O		x	x			x			hmd	2D, 3Dstat
	Olwal et al [17]					x	O	x				x	IR			handheld	2D
	Reitnayr et al [18]				x		O	x		x			x			handheld	2D
	Azuma et al [19]	x					O	x		x			x			handheld	2D, 3Dstat
	Elmqvist et al [29]			x			O	x							x	hmd	2D
	Newman et al [46]			x			O	x			x		inert	x		handheld	3Dstat
Edutainment and Games	Klein et al [20]		x				I		x		x		x			handheld	3Ddyn
	Henrysson et al [22]					x	I		x		mb					handheld	3Ddyn
	Rashid et al [30]					x	I	gprs					RFid			handheld	2D
	Piekarksi et al [50]	x					I,O	x			mb					hmd	3Ddyn*
Cultural Heritage	Vlahakis et al [36]	x					O	x		x	x		x			hmd	2D, 3Dstat
	Papagiannakis et al [12]	x					I,O		x		ml					hmd	3Ddyn*
Collaborative assembly	Revenier et al [42]		x				I		x	x			x			hmd	3Dstat
	Stork et al [26]		x				I				ml		x			hmd	3Dstat
Maintenance & inspection	Gausemeier et al [47]				x		I	x			ml					handheld	2D/3Dstat
	Vacchetti et al [13]	x					I		x		ml					hmd	2D, 3Ddyn*
	Goose et al [15]				x		I	x			mb					handheld	2D/3Dstat
	Rauhala et al [48]	x				x	I		zigBee, bluetooth		mb					handheld	2D/3Dstat
	Makri et al [49]				x		I	x			both					both	2D/3Dstat