Accelerating the Mass-Adoption of Augmented Reality through High-Fidelity Prototyping

Christian Sandor
University of South Australia
Wearable Computer Lab
christian.sandor@unisa.edu.au

1. Introduction

So far, Augmented Reality (AR) systems have been confined to research prototypes and niche applications. However, recent advances in AR research make it seem feasible that AR can go beyond this stage. We envision that it can find broad adoption on consumer devices; therefore, ubiquitously enriching peoples' life.

It is very likely that mobile phones are going to be the first delivery platform of AR for the masses. They already include cameras (one of the preconditions for doing AR) and they are already being carried on a daily basis by many people. As a result, several research groups and companies have recently been focussing their efforts on this target platform.

Most of the time, these efforts consist of porting existing AR applications and tracking algorithms to mobile phones. While this approach helps to make AR popular, we doubt that these efforts alone are sufficient for a mass-adoption of AR. In this paper, we present an alternative, complementary research strategy for achieving this goal.

Fundamentally, current mobile phones have not been designed to run AR applications. Therefore, the user interfaces and applications that can run on them can achieve a low-fidelity only. As a result, it is inherently very difficult to create compelling user experiences on current mobile phones. We believe in a different approach: by simulating the capabilities of future, powerful mobile phones, we can create rich, high-fidelity user interfaces and applications. These explorations can help to convince mobile phone companies to create mobile phones with specialized AR hardware. As a result, the mass-adoption of AR would be accelerated.

The rest of this paper is structured as follow: Section 2 presents a brief overview of tracking technology for AR, since tracking is the main technical constraint for running AR applications. Section 3 details the motivation and our plan for creating high-fidelity prototypes. Section 4 draws conclusions and details future work.

2. Tracking for AR

In this section, we give a brief overview of the state of the art for tracking for AR. First, we present tracking technologies that have been demonstrated on mobile phones. So far, these approaches have been limited to running marker-based tracking on mobile phones. Second, we present more powerful tracking approaches, such as natural-feature tracking and sensor-based approaches. Due to the current computational limitations of mobile phones, these have been demonstrated only on laptops or wearable computers that are based on laptops.

Several researchers have ported marker-based tracking algorithms to mobile devices. For example, Daniel Wagner an colleagues[17] have ported ARToolkit[6] to a PDA in 2003. In 2008, Wagner and colleagues have presented a further improvement: they have demonstrated[16] the use of texture-based markers, based on Ferns[10] and SIFT[8] on a mobile phone.

However, for compelling AR experiences, marker-based tracking approaches are too limited. Ideally, a tracking system could work in a totally unprepared environment. The most simple approach to achieving this is to use a combination of GPS for localization and an inertial sensor coupled with a magnetic compass to provide orientation. This approach has been used in several systems, such as Feiner et. al's MARS[4] and Piekarski's TINMITH[11].

While this approach is simple and pragmatic, its accuracy and applicability are limited. Reitmayr et. al[14] have presented a robust and accurate tracking solution by combining these sensors with an edge tracker based on RAPID[5]. While this approach improves tracking accuracy dramatically, there are still two shortcomings: First, it does not work indoors due to the reliance on GPS. Second, the edge tracker requires a textured 3D model of the environment, which requires a significant preparation effort.

Probably the most promising approach for AR tracking is SLAM[15], because it can work in totally unprepared environments. Georg Klein has presented PTAM[7], which is an adaption of SLAM to the needs of AR tracking. However, PTAM is limited to a workspace of a table-top size. Recently, the same research group has presented PTAMM[3], which overcomes some of the limitations by creating several workspaces that can be distributed on a wider area, for example a building.

3. High-Fidelity Prototyping

As we have already briefly explained in Section 1, we assume that high-fidelity prototypes are a key element for achieving a mass-adoption of AR. Section 3.1 gives more details about the rationale behind this assumption. Section 3.2 details how we intend to create such prototypes.

3.1. Motivation

Bill Buxton describes in his book "Sketching User Experience"[2] a methodology for exploring tomorrow's user interfaces and applications by creating working prototypes with today's technology. He suggests that by putting user experience first, technological developments can be driven further. In this section, we explain how this principle can be applied to prototyping future handheld AR devices by employing today's laptops.

According to Moore's Law, processing speed doubles every two years. For extrapolating the speed of future mobile devices, a useful tool is to look at the gap between the speed of mobile phones and laptops. This gap can be assumed to be about seven years[13]. Table 1 illustrates this point by comparing Apple's current mobile phone and their seven year old high-end laptop.

The fundamental constraint for AR is tracking. As described in Section 2, PTAM comes close to an ideal solution; but, there are also other viable approaches. For the sake of the argument, we will limit our discussion to PTAM. PTAM was first demonstrated in 2007. When can PTAM run on a mobile phone? A conservative estimate is depicted in Figure 1. By Moore's law alone, it can be expected to run on mobile phones by 2014. This estimate is conservative, because it does not account for developments in computer vision.

By using today's laptops, we can simulate future AR user interfaces and applications on mobile phones. Figure 2 depicts one possible application: AR x-ray vision. We have created a system[1] that creates an image-based reconstruction of a remote scene, based on captured images and models. That reconstruction is overlaid on the users view of the environment to provide x-ray vision. In order to provide correct depth cues, we highlight the edges of foreground objects. To perform the edge detection in realtime, we use GLSL shaders. It would not be possible to run it on a current mobile phone, since they lack shader support.

	Powerbook G4	iPhone 3G
Release Date	January 2001	July 2008
Hard drive	20 GB	16 GB
Processor	500 MHz G4	620MHz ARM
Memory	256 MB	512 MB

Table 1. Laptops vs. mobile phones.

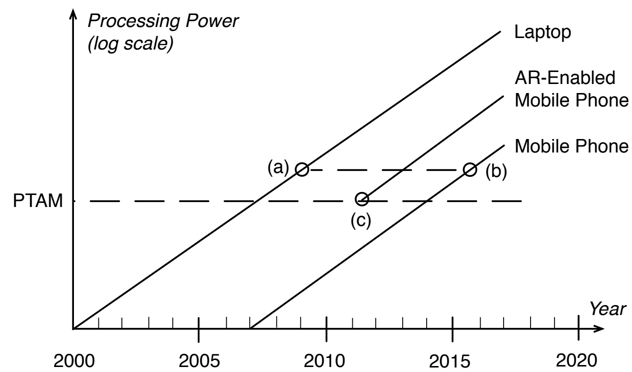

Figure 1. Using today's laptops (a) to prototype AR on future mobile phones (b). This could inspire mobile phone companies to create AR-enabled mobile phones (c).

Figure 2. Example of a high-fidelity user interface prototype[1]: an image-based reconstruction of a remote scene is overlaid on the users view of the environment to provide x-ray vision capabilities.

This application is an example of a high-fidelity prototype. Instead of focussing on building AR systems on

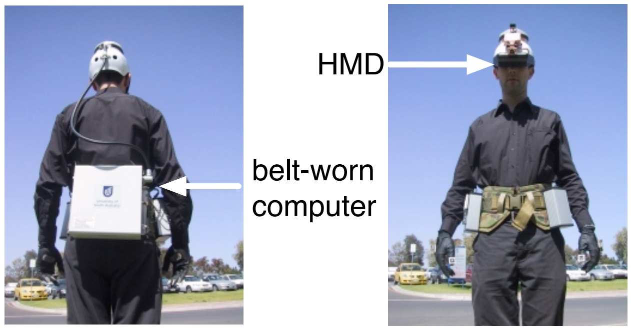

Figure 3. TINMITH: Our previous high-fidelity prototyping platform for AR with a head-mounted display (HMD).

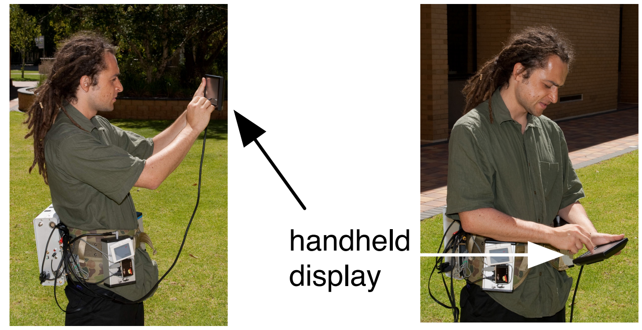

Figure 4. TINT: Our next generation high-fidelity prototyping platform for future mobile phones.

current mobile phones, we instead use laptops to simulate tomorrow's mobile phones. This strategy has the advantage that much higher fidelity user interfaces can be explored, because of better tracking and graphics capabilities. Through these high-fidelity prototypes, we hope to create compelling applications that inspire mobile phone companies to create AR-enabled mobile phones (for example, by creating a PTAM chip), which would speed-up mass-adoption of AR significantly (see Figure 1).

3.2. Implementation

This section describes how we intend to create prototypes of AR on future mobile phones. First, we describe the hardware/software platform that we are in the process of creating. Then, we describe how we intend to explore novel user interfaces and applications.

We base our implementation on the belt-worn computer of the TINMITH[11] system. TINMITH is a powerful mobile AR hardware/software platform (see Figure 3) that has been developed in our lab for the last 8 years. For example, the system presented in Figure 2 is running on TINMITH. However, to pursue our strategy of high-fidelity prototyping of future mobile phones, we intend to do major modifications to most components of TINMITH, thereby creating a new platform: TINT[1] (see Figure 4).

The current TINMITH system is based on a modified Toshiba Tecra M5 with a Pentium-M 2.0 GHz, 2 GB RAM, NVIDIA GeForce 6600 graphics chipset, 802.11 wireless networking, and Bluetooth. The computer is mounted on the back of the belt, with battery power for 2-3 hours. Headphones are fixed to the inside the helmet, so ambient sounds can still be heard. The Tinmith system employs video see-through technology. A 640x480 pixel 30fps Point Gray Firefly firewire video camera is attached on the front of the helmet. A GPS antenna and an InterSense InertiaCube3 inertial orientation sensor are mounted on the helmet. The GPS receiver has sub-50 cm accuracy and is used to track the position of the user's head outdoors. Additionally, the user has several ways of performing input with pinch gloves and by tracking his hands with ARToolkit.

For TINT, we intend to completely remove TINMITH's helmet and input facilities. We intend to keep the belt-worn computer and the sensors used for tracking (GPS, Gyroscope). The main novelties are going to be the addition of PTAM to improve tracking and a handheld display for displaying AR content and performing user input. The handheld display will have multi-touch input and tactile feedback, both on[12] and off[9] the screen.

Also, we are going to rewrite the software framework from scratch. TINMITH's software framework was geared towards efficiently performing constructive solid geometry in an outdoor setting to model buildings and other real-world objects. However, TINT is mainly geared towards exploring novel user interfaces and applications. Therefore, we intend to create a more flexible and lightweight framework.

Once we have completed TINT's basic hardware/software platform, we intend to create novel user interfaces and applications. Then, we intend to evaluate them with users who have never used AR before in a real-world setting such as a shopping mall. Most user studies for AR user interfaces neither have novice users as subjects, nor are these studies performed in a real-world setting.

[1]The acronym TINT (*T*his *i*s *n*ot *T*INMITH) is a pun on the acronym TINMITH (*T*his *i*s *n*ot *m*ap-*i*n-*t*he-*h*at). Map-in-the-hat was an early prototype by Prof. Bruce Thomas.

4. Conclusions

In this paper, we have presented how we intend to contribute towards the mass-adoption of AR by creating high-fidelity prototypes of future mobile phones, based on the combination of a powerful wearable computer with a handheld display. We believe that by creating high-fidelity prototypes, we can help to convince mobile phone companies to include specialized AR hardware into their phones.

For example, consider that several consumer cameras already contain specialized chips that perform complex computer vision tasks, such as detecting whether a person in a photo smiles. Similarly, a specialized chip for running PTAM could be constructed. However, in order for a company to develop this chip, convincing applications and user interfaces have to be demonstrated first.

References

[1] B. Avery, C. Sandor, and B. H. Thomas. Improving spatial perception for augmented reality x-ray vision. In *VR '09: Proceedings of the IEEE Virtual Reality Conference 2009*, Lafayette, Louisiana, USA, March 2009.

[2] B. Buxton. *Sketching User Experiences: Getting the Design Right and the Right Design*. Morgan Kaufmann, April 2007.

[3] R. O. Castle, G. Klein, and D. W. Murray. Video-rate localization in multiple maps for wearable augmented reality. In *ISWC '08: Proceedings of the 12th IEEE International Symposium on Wearable Computers*, pages 15–22, September 2008.

[4] S. Feiner, B. MacIntyre, T. Höllerer, and A. Webster. A touring machine: Prototyping 3d augmented reality systems for exploring the urban environment. *Personal and Ubiquitous Computing*, 1(3), 1997.

[5] C. Harris. Tracking with rigid models. *Active vision*, pages 59–73, 1993.

[6] H. Kato and M. Billinghurst. Marker tracking and hmd calibration for a video-based augmented reality conferencing system. In *IWAR '99: Proceedings of the 2nd IEEE and ACM International Workshop on Augmented Reality*, page 85, Washington, DC, USA, 1999.

[7] G. Klein and D. Murray. Parallel tracking and mapping for small AR workspaces. In *ISMAR '07: Proceedings of the Sixth IEEE and ACM International Symposium on Mixed and Augmented Reality*, Nara, Japan, November 2007.

[8] D. G. Lowe. Distinctive image features from scale-invariant keypoints. *International Journal of Computer Vision*, 60(2):91–110, 2004.

[9] J. Luk, J. Pasquero, S. Little, K. MacLean, V. Levesque, and V. Hayward. A role for haptics in mobile interaction: initial design using a handheld tactile display prototype. In *CHI '06: Proceedings of the SIGCHI conference on Human Factors in computing systems*, pages 171–180, New York, NY, USA, 2006.

[10] M. Ozuysal, P. Fua, and V. Lepetit. Fast keypoint recognition in ten lines of code. In *CVPR '07: IEEE Conference on Computer Vision and Pattern Recognition*, pages 1–8, 2007.

[11] W. Piekarski and B. Thomas. Tinmith-metro: new outdoor techniques for creating city models with an augmented reality wearable computer. In *ISWC '01: Proceedings of the 5th International Symposium on Wearable Computers*, pages 31–38, 2001.

[12] I. Poupyrev, J. Rekimoto, and S. Maruyama. Touchengine: a tactile display for handheld devices. In *CHI '02: CHI '02 Extended Abstracts on Human Factors in Computing Systems*, pages 644–645, New York, NY, USA, 2002.

[13] K. Pulli. (Nokia Research Center), personal communication, 2007.

[14] G. Reitmayr and T. W. Drummond. Going out: Robust tracking for outdoor augmented reality. In *ISMAR '06: Proceedings of the Fifth IEEE and ACM International Symposium on Mixed and Augmented Reality*, pages 109–118, Santa Barbara, CA, USA, October 2006.

[15] R. C. Smith and P. Cheeseman. On the representation and estimation of spatial uncertainly. *International Journal of Robotics Research*, 5(4):56–68, 1987.

[16] D. Wagner, G. Reitmayr, A. Mulloni, T. Drummond, and D. Schmalstieg. Pose tracking from natural features on mobile phones. In *ISMAR '08: Proceedings of the Seventh IEEE and ACM International Symposium on Mixed and Augmented Reality*, pages 125–134, September 2008.

[17] D. Wagner and D. Schmalstieg. First steps towards handheld augmented reality. In *ISWC '03: Proceedings of the 7th IEEE International Symposium on Wearable Computers*, page 127, Washington, DC, USA, 2003.