PocketHouseAR - An approach to use a Pocket PC as interaction tool for Augmented Reality

Gerald Binder*

Media Technology and Design
Upper Austria University of Applied Sciences
Hagenberg College of Information Technology
Hagenberg / Austria

Abstract

Today Augmented Reality is one of the most promising technologies in the field of computer graphics. The enrichment of the real world with virtual content can be useful in a large number of applications.

Depending on the particular application, the user is more or less immersed in the virtual world. Therefore the best method of user interaction may be different in every case. Obviously there is still a lot of research which has to be done to increase the usability of Augmented Reality applications.

At the same time, new ways of human-computer interaction are becoming common, for example the Pocket PC. The interaction metaphors are quite the same as on a desktop PC. This helps the user to understand quickly how to interact with the application.

Our approach was to combine the possibilities of Augmented Reality with the easy-to-use interface of a Pocket PC. For this purpose, a prototypical application in the field of architecture was implemented. It allows the user to modify the computer-generated three-dimensional model of a simple house. All interaction can be achieved with a standard Pocket PC connected wirelessly to the Augmented Reality server.

Keywords: Augmented Reality, Pocket PC, Human-Computer Interaction

1 Introduction

Interaction is one of the key issues in AR environments. It can be roughly categorized into three main aims:

- Navigation (e.g. control of the viewpoint),
- Selection and manipulation of the virtual content and
- Configuration and control of the AR application.

The importance of each of these categories may vary depending on the application. Nevertheless, it is important to consider their impact on the design of an AR interaction tool. As an Augmented Reality application integrates both reality and virtuality, the interaction metaphor can also be a combination of them. The interaction tools presented later in this paper will demonstrate this.

As stated by Milgram, the interaction can also be classified by the congruency of interaction and the effect the interaction causes [3]: The more direct the interaction works, the more congruent the interaction metaphor becomes. For example, using a real steering wheel to drive a car in a computer game is more direct than using keys on the keyboard. But also the spacial distance between interaction tool and display is important. Finally, the way the AR scene is displayed (monitor, projection or Head Mounted Display) is another factor to consider. As a conclusion the interaction metaphor has to meet the special requirements of a particular AR application.

The Pocket PC provides several features which are very useful in AR environments. Its wireless network connection offers great mobility and it can also be extended by different devices, such as a camera. And finally, the use of a Pocket PC is very intuitive for people who already have experience with desktop computers.

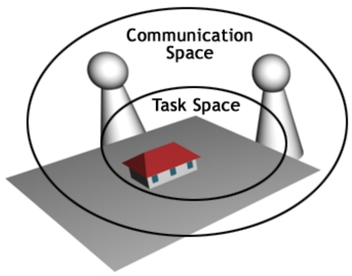
Especially for Collaborative AR, the Pocket PC could close the gap which arises with the separation of the communication space and the task space (see figure 1). With the use of a Pocket PC, the task space is a sub-set of the communication space. As then the users are collaborating in a face-to-face situation, natural non-verbal communication is still possible. Furthermore every participant has his own private user interface with all the necessary control and interaction tools. Thus each user can have other tools depending on his special needs.

In the field of Ubiquitous Computing [7], the Pocket PC could be used as both an output device for common Ubiquitous Computing applications and as an interaction tool for applications extended with AR functionality.

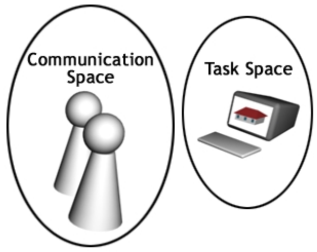
2 System Overview

The aim of this study was to prototypically design a whole system which processes the user input from a Pocket PC

*gerald.binder@fh-hagenberg.at



(a)



(b)

Figure 1: (a) Face-to-face collaboration. (b) Collaboration in front of a computer monitor. (after [1])

and delivers it to the AR server, which finally renders the three-dimensional AR scene.

There should be clear interfaces between the user input, the application logic and the AR functionality. Thus the implemented system consists of two major components: *PocketHouseAR* and *SmartSpace*. *SmartSpace* itself is a framework which enables network access for a special version of the ARToolKit [2]. Figure 2 shows an overview of the system.

The front end application is encapsuled within *PocketHouseAR*. The *Client-Software* on every user's device sends the user's input to the *PocketHouseAR Middleware*. The *PocketHouseAR Middleware* synchronizes the *Client* user interfaces and calculates the transformation commands.

After that the *PocketHouseAR Middleware* sends these commands to the *TMsgServer*, which acts just as a message server. It distributes the message to any application registered to this service. In this case, only *TView* receives these messages. *TView* contains the core AR functionality and is based on the *ARToolKit* [2]. Here the AR scene is calculated and rendered depending on the user's input.

As all the application logic is processed only within these two components, it is easy to change the application itself whereas the *SmartSpace* component does not have to be changed. It is also thinkable that in a real multi-user environment every user receives his own video output into his Head Mounted Display.

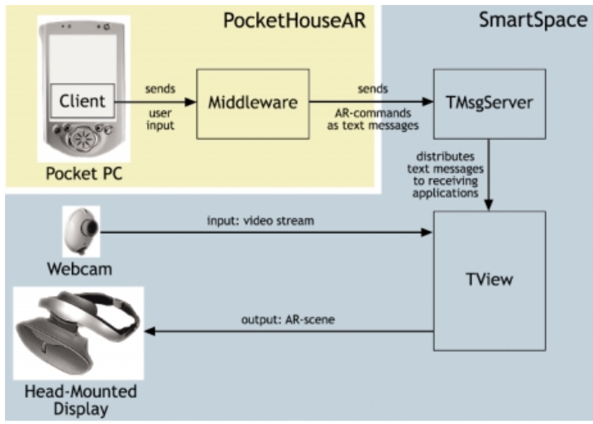


Figure 2: *PocketHouseAR* and *SmartSpace* are the major components of the system.

In the implemented prototype, it is possible to use one Pocket PC as a client. But it is also possible to use a standard desktop PC, notebook or even a Smartphone. The middleware itself also contains all the client functionality and thus may be used by a second user.

3 PocketHouseAR

As mentioned above, *PocketHouseAR* itself consists of two components: the middleware and the client software. The client software consists of the graphical user interface and its major function is to handle the user input and send it to the middleware.

Here the current properties of the 3D objects are stored. Incoming user commands are modified based on this data and then sent to *SmartSpace*. For example, imagine that a user has changed the roof transparency to 50 percent. Another user wants to decrease the transparency by 10 percent. If the middleware had not stored the current value of 50 percent, the second user input would result in an overall transparency of 80 percent. This outlines again that the client only handles the user input. The application logic has to be implemented within the middleware as it has to process the input of every user.

This results in the middleware also being responsible for the synchronisation of the user interfaces. As every user can independently interact with the 3D objects, the client software has to know the current properties of these objects to display the correct position of the transparency slider, for example.

3.1 Functionality

The use case for *PocketHouseAR* could be described as follows: Several architects are gathered around a table on which the virtual object of a building is embedded into a

Figure 3: Using *PocketHouseAR*. On the screen the user's view of the AR scene and the user interface of the Pocket PC can be seen.

physical model of the object's surrounding. Every participant is equipped with a Head Mounted Display and thus has his own view of the AR environment. He can move freely around the table and interact with the virtual objects. The AR scene can also be projected onto the wall, in which case the viewpoint would be fixed to a certain position with a certain orientation.

Since system design was the major focus of this study, there are just a few basic functions implemented. *PocketHouseAR* provides the user with the following interaction possibilities:

- Change roof texture
- Change wall texture
- Change windows' appearance
- Rotation, translation and scaling of the 3D model

All interaction can be done with the stylus or the hardware keys on the Pocket PC. This is especially important in cases where the user wears a Head Mounted Display, as this impairs his view of the user interface. Figure 3 shows the user equipped with a Pocket PC and a HMD with mounted camera.

3.2 User Interface

As mentioned earlier, the big advantage of using a Pocket PC as interaction device is that well known interaction metaphors can be used. Typical GUI elements such as sliders and buttons do not have to be explained seperately to the user. Thus, the interface design can focus on the logical aspect. The main issue here is to determine when a certain function should be possible. The user interface should be as intuitive and easy to handle as possible. This assures that the user can concentrate on the AR scene. Additionally, the importance of the logical aspect increases the more functionality has to be integrated.

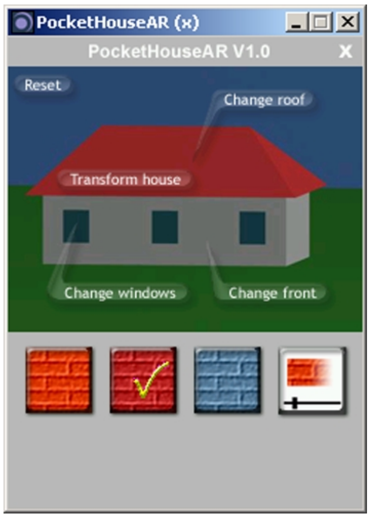


Figure 4: The user interface of *PocketHouseAR* consists of two areas: The selection area to choose the functionality and the interaction area to change properties.

In the case of *PocketHouseAR*, a quite simple approach could be taken. The GUI mainly consists of two areas: The selection area to choose the functionality and the interaction area below to change properties of certain 3D objects (see figure 4).

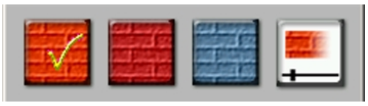
The selection area always remains visible while the interaction area has to change depending on the selection functionality. As seen in figure 4, the interaction areas always contain simple buttons and sliders which are as big as possible to facilitate quick user input.

4 Related Work

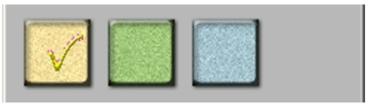
Many different approaches have been taken in the research of 'Pen and Tablet'-based interaction metaphors. The project *AR Pad* [4] follows the work of Rekimoto, who introduced the concept of a 'Magic Lens' with NaviCam [5] and TransVision [6]. It combines both the interaction tool and the display in one device (see figure 6). It is also possible to fix the relation between a certain virtual object and the *AR Pad*. Therefore it is possible to use the *AR Pad* as a sort of cursor, as every movement of the *AR Pad* results in a movement of the virtual object as well.

In the project *AR-PDA*, the augmentation of the video stream delivered by the device has to be done by a remote server. The aim of this project is to develop a product for the mass market. An interesting aspect of this system is the approach to do the analysis of the captured visual information via a remote server. This extends the possible amount of related data (stored in large databases) enormously. A highly advanced image recognition algorithm had to be implemented to really make this system suitable for the mass market.

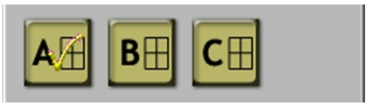
In contrast, an optimized ARToolKit [2] version to run



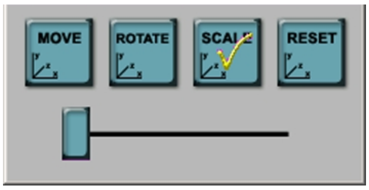
(a)



(b)



(c)



(d)

Figure 5: Several different interaction areas. (a) Changing the roof texture and transparency. (b) Changing the wall texture. (c) Changing the windows' appearance. (d) Transforming the whole 3D house.

Figure 6: AR Pad: Interaction tool and display combined in one device. (after [4])

directly on a Pocket PC has recently been released [11]. As the computing power and memory of Pocket PCs are supposed to increase, this could be an interesting alternative.

The 'Studierstube' [8] uses the *Personal Interaction Panel* [9] as a hybrid interaction tool. While the interaction tool itself is physical (tracked pen and tablet), the user interface is computer-generated and two- or three-dimensional. Thus the system provides haptical feedback through the pen touching the panel as well as great possibilities in user interface design. It is even possible to interact with 3D objects directly.

The *Pinch Glove* [10] was developed to support hands free interaction. It consists of a touch screen mounted on one wrist and several markers attached to both of the user's hands. Therefore the hands can be used as '3D cursors'. The video output is displayed via a HMD.

5 Taxonomy of the presented interaction tools

As described above there are several different approaches to implement an interaction tool for Augmented Reality. To examine the potential usefulness of one of them for a specific application they can be classified by asking three questions (after [3]):

- Is the interaction more 'real' or 'virtual'?
- How congruent is the interaction and the displayed AR scene?
- Is the interaction done from an egocentric or an exocentric perspective?

After analyzing the tools presented above, the result can be visualized like in figure 7. Every axis of this cube stands for one of the questions above.

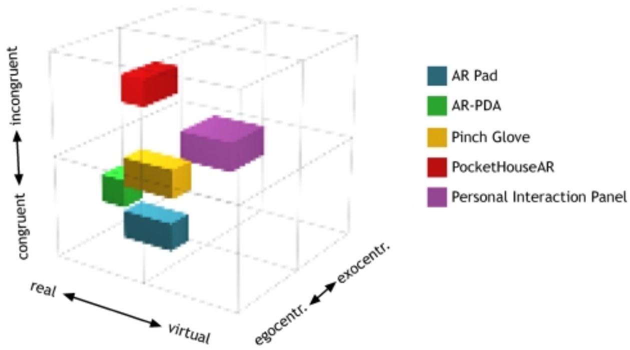


Figure 7: Classification of interaction concepts based on three different parameters.

A clearer overview can be achieved by only comparing two aspects in a two-dimensional diagram. In figure 8, the degree of reality/virtuality on the one hand and the congruency of interaction and graphical display on the other hand is shown.

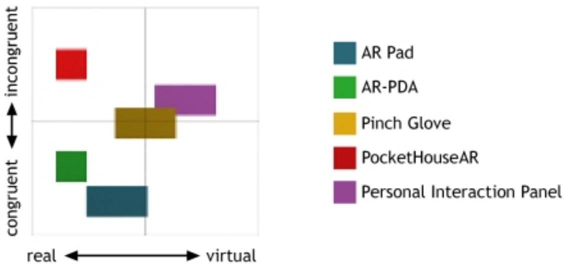


Figure 8: Classification of interaction concepts based on their degree of reality/virtuality and the congruency of interaction and graphical display.

Reality/Virtuality - *AR-PDA* and *PocketHouseAR* are the most real interaction concepts, as no virtual elements are used to interact. The 3D cursor of the *AR Pad* and the relatively virtual *Pinch Glove* (the hand grabs a virtual object) bring these concepts to the middle of the reality/virtuality continuum. The *Personal Interaction Panel* is the most virtual interaction tool.

Congruency - *AR Pad* offers the most direct user interaction. In addition, the spacial distance between interaction and graphical display of the AR scene is relatively short. The interaction is particularly abstracted by the remote server. Finally, the tracking of the interaction medium, which is necessary for the *Pinch Glove* and the *Personal Interaction Panel*, is the reason for their centered position in the congruency spectrum. *PocketHouseAR* offers a relatively low congruency as the user has to move his view from the AR scene to the Pocket PC to interact.

Perspective - Basically each of the presented interaction tools uses an egocentric perspective. Theoretically, *PocketHouseAR* and the *Personal Interaction Panel* could also be implemented to offer an exocentric perspective as well.

6 Conclusion

Basically, the use of a Pocket PC as an interaction medium in AR environments offers the best of two worlds. First of all, it supports all the forms of interaction necessary for Augmented Reality applications: navigation, modification and configuration can be implemented according to specific needs. Secondly, it is an easy-to-use interface as it still uses the WIMP (Windows, Icons, Menus, Pointers) paradigm. Of course, this paradigm is not suitable for every AR application and has to be adapted if required.

Another advantage of the Pocket PC (or a similar handheld computer) is that a user can instantly participate in a running application with his own personal device, presuming he has the client application installed on it. This provides great flexibility to the application itself as no special hardware except for the HMD's is needed.

Even though the combination of real and virtual contents is not the main aspect of using this interaction mechanism, the use of it could be preferable to pure virtual reality. The reasons may be the 'more natural' appearance of the world seen through the HMD. Navigation might be easier, especially for new users who are not familiar with navigation in a pure virtual environment.

The use of a Pocket PC and a Head Mounted Display at the same time can be problematic if the user interface cannot be seen properly through a HMD. The user then has to look beneath the HMD. This causes a sort of media change (view through HMD vs. natural view) which could be uncomfortable for some users. This is obviously a main disadvantage of the presented system *PocketHouseAR*. But with the further development of HMD's this problem could become less prevalent.

7 Potential Applications

In addition to the presented sample application, *PocketHouseAR*, there are a lot of other possible fields of use for this interaction concept.

For example, the use in education could offer teaching and learning on 3D models of items which normally would be too big, too heavy or too expensive to work with.

Another example would be the work with a large three-dimensional map. Several people who are gathered around a table can then explore the map and point with their fingers to specific places. The main advantage compared to a conventional two-dimensional map is that information which contains 3D data can be displayed so that it can be understood more easily. Of course, it would also be feasible to change the viewpoint or the appearance of the map instantly.

8 Future Work

In a future version of *PocketHouseAR*, real multi-user support could be implemented. Every user will then have his own view of the AR scene. Especially in the field of architecture, the integration of real physical objects would be a useful feature. At a building site, *PocketHouseAR* could give an impression of the new building within its natural environment. For this use of the system other tracking methods have to be implemented, as the marker-based optical tracking used in this system would not be suitable.

Undoubtedly, the user acceptance of a specific interaction tool is the basis for a successful spread of AR applications. As the development of AR is tightly related to Ubiquitous Computing interaction should work in an uncomplicated and familiar way. Therefore, a lot of research on interaction techniques for AR still has to be done.

References

[1] M. Billinghurst, D. Belcher, A. Gupta, and K. Kiyokawa. Communication behaviors in co-located collaborative ar interfaces. To appear.

[2] Human Interface Technology Laboratory. Shared space/artoolkit download page, February 2002. http://www.hitl.washington.edu/research /shared_space/download/.

[3] Paul Milgram and H. Colquhoun Jr. *A Taxonomy of Real and Virtual World Display Integration*, pages 5–30. Springer Verlag, Tokyo, Japan, 1999.

[4] D. Mogilev, M. Billinghurst, K. Kiyokawa, and J. Pair. The ar pad, October 2001. Demonstration at the 2001 IEEE and ACM International Symposium on Augmented Reality (ISAR).

[5] Jun Rekimoto. The magnifying glass approach to augmented reality systems. In *International Conference on Artificial Reality and Tele-Existence '95 / Conference on Virtual Reality Software and Technology '95 (ICAT/VRST'95) Proceedings*, pages 123–132, 1995.

[6] Jun Rekimoto. Transvision: A hand-held augmented reality system for collaborative design. In *Virtual Systems and Multi-Media (VSMM)'96*, 1996.

[7] Bradley J. Rhodes, Nelson Minar, and Josh Weaver. Wearable computing meets ubiquitous computing: Reaping the best of both worlds. In *The Proceedings of The Third International Symposium on Wearable Computers (ISWC '99)*, pages 141–149, October 1999.

[8] D. Schmalstieg, A. Fuhrmann, Z. Szalavari, and M. Gervautz. Studierstube - an environment for collaboration in augmented reality. In *CVE '96 Workshop Proceedings*, September 1996.

[9] Z. Szalavri and M. Gervautz. The personal interaction panel – a two-handed interface for augmented reality. In *Proceedings of EUROGRAPHICS'97*, volume 16, pages 335–346, September 1997.

[10] S. Veigl, A. Kaltenbach, Ledermann F., G. Reitmayr, and D. Schmalstieg. Two-handed direct interaction with artoolkit. Technical Report TR-188-2-2002-11, Vienna University of Technology, 2002.

[11] Daniel Wagner. Handheld ar, June 2003. http://www.ims.tuwien.ac.at/research/handheld_ar/.