MARS: Multi-Touch Augmented Reality System and Methods of Interaction with It

Matej Novotný – Ján Lacko – Martin Samuelčík

Abstract

Augmented reality and multi-touch interaction are two exciting and rapidly developing areas of technology. So far, they have existed and have been developed mostly separately. The presented work explores the attractive qualities of either and demonstrates how they can be merged together to create an immersive and versatile presentation/exploration tool. We present novel ways of interaction that stem from interleaving augmented reality with multi-touch devices and demonstrate them on our own prototype. Our solution creates an interactive environment for presentation and exploration of objects (and data in general) in two linked views with seamless attention transition between the views.

Keywords:

augmented reality, multi-touch, interaction, multiple displays, linked views

1. Introduction

One of the main qualities of a computer application is its user interface and many would argue it is *the* main quality of an application. Especially, when the tasks performed in the application are not trivial and the limited resources of user attention must be spared for the actual task performed in the application. In such case, additional strain put on the user by a cumbersome user interface might obstruct the effective use of the program.

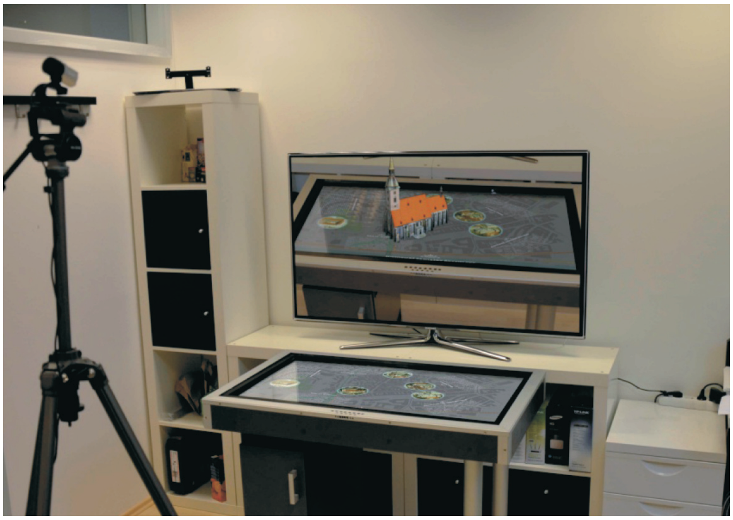
For many years now, user interfaces have been centered round the WIMP paradigm (Windows, Icons, Menu, Pointers) [1]. User input in a WIMP-oriented application is largely based on reading menus and deciphering icons. Both of these actions impose a relatively significant cognitive load on the user that can't be neglected. Of course, WIMP used to be the best choice considering the hardware options in the past: a keyboard, a mouse and a monitor.

However, the development of new user input technologies – (multi-)touch screens, gesture or voice recognition or kinetic sensors – provide wider options for creating better and more natural user interfaces. Unlike reading or writing, manipulating things physically is a natural action we are trained for since early childhood. It's no surprise touch-based interfaces are so intuitive. A modern application should try to take advantage of these benefits in order to help improve its user interface.

We believe touch-based input (and especially multi-touch) creates a simple and intuitive platform for conducting many tasks [9]. Our focus is on navigation and interaction in a spatial environment: e.g. virtual reality or augmented reality [12]. These environments create information spaces defined by the three-dimensional world; virtual world in case of virtual reality and real world in case of augmented reality. Navigation and interaction tasks performed in spatial environments frequently includeoperations such as pick, pan, zoom or rotate. We decided to combine the effortless spatial navigation of a multi-touch input device with the immersive presentation qualities of an augmented reality display. We took the best of both worlds and created a novel way of exploring visual information spaces.

This paper describes the combination of a multi-touch user interface with an augmented reality system and focuses on both of the two components. We describe the basic concepts of each in the next two sections. Then, we present the related work in combining the two and describe methods of interaction with the combined system. In Section 3, we describe our setup which creates a smooth dual-view environment for object/data visualization. We also propose several usage scenarios for this multi-touch augmented reality system.

Figure 1. *An example application of augmented reality running on the MARS: Bratislava sights are mixed with the interactive map of the old town.*



1.1 Multi-touch interaction

Touch-based interaction with computers has experienced a significant increase in popularity among the general public throughout the recent years. Fed mostly by the mobile devices industry – smart phones, tablets and formerly PDAs – the technology found its way from the laboratories and industrial applications to all consumers. It is thanks to the natural behaviour of touch interaction that the interfaces employing technology are becoming more and more common. The trend is well set and will most probably continue in future as well. As a matter of fact, often the first contact with a computer interface is with a touch-sensitive device (a mobile phone or a tablet) instead of the more traditional keyboard or mouse. Children and the elderly [2] find it more natural to interact via touch. The more natural the application interface behaves, the more cognitive resources remain available for understanding the application content.

In this work, we adhere to the already established principles of multi-touch interaction. We attempt to use it to create a user interface that does not require extensive attention dedicated to the interaction itself, thus leaving more attention resources for the augmented reality display and the displayed data in general.

1.2 Augmented reality

The core concept of augmented reality is mixing computer generated content with real world context. As opposed to CGI special effects in films, which also often combine computer generated imagery with real shots, the final product is interactive. The interactivity and the real world environment create a realistic experience and the content is perceived to be more "real" than when presented in its original virtual environment. Moreover, semantic bonds can be created between the virtual objects and the real environment, e.g. placing a digital price tag next to a camera shot of real house. We use the semantic bonds to help the user transfer his/hers attention between different data views and also to relieve the user's attention from the interaction controls. (See Section 3.1 for more details.)

Similar to touch interfaces, augmented reality is becoming popular thanks to the development of consumer electronics. The concept has been around for several years now, e.g. occasionally employed in medicine [3] or engineering. Recently, it has become a popular marketing presentation tool and a way of increasing content attractiveness among the audience. It can be presumed that, even though being a rather sophisticated concept, augmented reality will become a standard means of visualization, exploration and presentation of data.

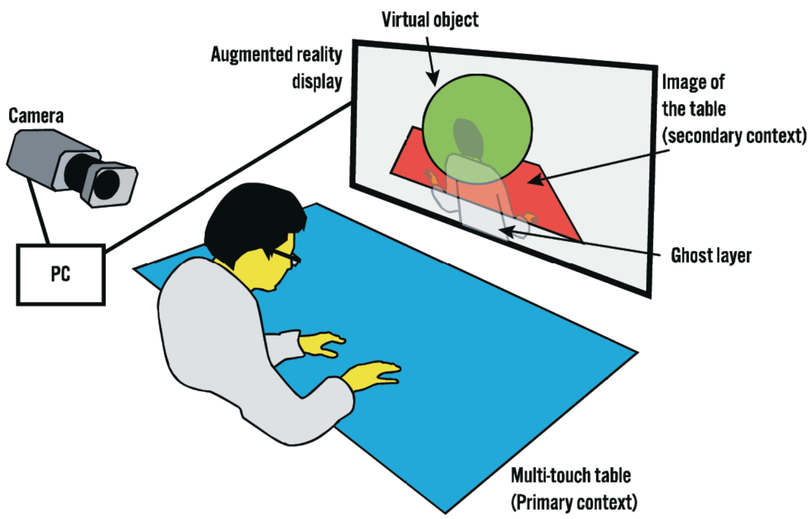
2. related work

The combination of augmented reality and multi-touch interaction has been experimented with previously. Benko et al. [4, 5] used to create an augmented reality system and explore the options of using gestures and touch to interact with the system. Head-mounted displays were also used by Dedual et al. [6], focusing mostly on the hardware/software solution. Collaborative cooperation on a multi-touch table top extended by mobile augmented reality was presented by Na et al. [7], again exploring mostly the technology of the devices. Wei et al. [8] created a system that uses a second display as a supplement to the table top touch display. Their solution is the closest to our, however, Wei et al. focus on a specific interior design application. At the same time, we rather explore the theoretical concept of our multi-touch augmented reality set-up and its immersion potential.

3. MARS: Multi-touch augmented reality system

Our proposed system – MARS – consists of two displays and a camera. One of the displays is a multi-touch table top display which creates the basic frame for user interaction. The camera records the user and the table he/she works with. Similarly to Krueger's original Videoplace [11], the video feed is then mixed in a computer with the virtual object(s) and it is projected on the augmented reality display. (This may be a projection screen or a large flat-screen display.)

Figure 2: *The plan of the MARS set-up.*

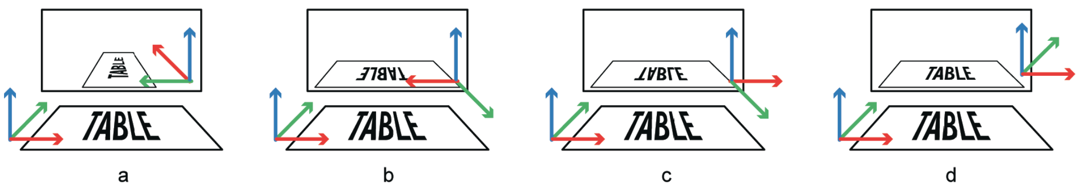


3.1 Camera position

There are several possible angles from which the scene can be recorded. Our goal was to maximize the seamlessness of the transition for the user when he/she switches attention from the table display to the augmented reality display. While there are no established measures of the transition seamlessness, our preliminary experiments show that the alignment of the spatial orientation of the two displays is crucial for building a mental model of the displayed data and for performing navigation and interaction tasks with ease.

Recording from left or right proved hard to comprehend for the users and thus the transition between the two displays was not smooth. Recording the scene from user's front half-space helps the users identify themselves with the video they see on screen. However, we found that this kind of projection strongly conflicts with the users learned expectation of a mirror image. The final effect is a confusion of left/right hands, similar to when trying to put on make-up or shave oneself using a web camera. Flipping the video horizontally helps, however, placing the camera behind the user's back turned out to be the most immersive because it aligns the augmented reality space with the real space in all three dimensions.

Figure 3: *Relative orientation of the real world coordinate system and the coordinate system of the augmented reality display. Depending on the position of the camera: a) to the left of the user, b) in front of the user, c) in front of the user, horizontally mirrored image, d) behind the user*



If the camera is placed behind the user, the projected image is an aligned extension of the 2D table space. Our experiments show that the users were able to operate the table top easily in this set-up even when looking at the augmented reality display instead of the table display. We argue that this improves the immersion and reduces the confusion when switching attention from one display to another.

3.2 Spatial registration

The registration of the camera position relative to the scene is the vital part of the augmented reality system. The camera in MARS is placed in a static position and does not move. Neither does the virtual base, which in our case is the table display. This eliminates the need for real-time spatial registration as we know if from augmented reality on mobile devices or when the augmented reality marker is being moved in front of the camera.

The only spatial calibration takes place when the camera is first placed (or when it's moved to a different position). The calibration consists of manually selecting the four corner points of the table display within the image space of the camera.

3.3 Augmentation

The recorded video feed is augmented by the rendered virtual object(s) which are aligned in 3D space to a virtual extrusion of the 2D space of the table display. The one disadvantage of recording the scene from behind the user's back is the user obstructing the view of the table display. However, we devised a way to fix this by first recording a still image of the scene without the user present and then blending it with the recorded video.

The final video contains three layers: the recorded still image serves as the background layer. Then, the virtual object is rendered in the second layer. And finally, the recorded video containing the user is alpha-blended on top of them. This creates a ghost image of the user and the virtual object is visible at the same time as the users see themselves interacting with the table.

We plan using Kinect [10] in next version of the system to completely replace the ghost layer by virtual hands or a virtual actor incorporated into the second layer. Thus, the occlusion will be eliminated completely and the user will still see his/her actions on the screen.

3.4 Two contexts

The dual-monitor set-up gives us an opportunity to present the virtual object in another context than what is displayed on the table top display. Figure 2 shows the original context in blue. The second context is displayed in the augmented reality display as a quadrilateral texture mapped to the space occupied by the original context. Figure 2 shows it in red.

This helps in situation where two contexts are necessary at the same time. E.g. in urban planning scenario, where the primary context (displayed on the touch table) shows cadastral plans and the secondary context shows satellite map as a ground for the planned buildings displayed in augmented reality.

4. Perception and Presentation

The goal of the MARS system is to create an immersive presentation/exploration platform. To achieve this, we eliminated the confusion when moving attention from horizontal table display to the vertical augmented display. The augmented display – aligned with the table display – creates an intuitive extension of the table display thus removing the mentioned confusion and providing for seamless transition of attention between the two displays.

Head-mounted displays (HMD) used in similar projects do not suffer from this confusion since they use only a single display (the table display) and the augmented displays are placed in the head gear. The user wearing the HMD does not need to move attention from one to another display as he sees both as one. While this behaviour is certainly a positive side to using HMDs, combining the table display with the augmented display into one has its drawbacks. The table display is inevitably obstructed by the virtual object and the user can see the data on the table display only when hiding the image.

Our solution offers the user the option to either look down on the table display and observe the 2D data without any occlusion by the virtual object, or look up and see the virtual object in the augmented reality display. These two views are interactively linked – the actions performed on the table display are immediately affecting the augmented display, thus the user builds a mental model of a three-dimensional space that is the extension of the two-dimensional table display.

We believe this type of presentation can be beneficial for many usage scenarios including architectural or urban design, engineering, education or medicine. The secondary context and the dual-display configuration can be put to good use in exploratory data visualization, where multiple linked views are common.

5. Example Applications

The vertical augmented reality display in combination with the horizontal multi-touch display creates a virtual 3D box where the bottom plane is aligned with the table and the front plane is aligned with the augmented reality display. This predetermines the usage of MARS for scenarios that involve exploration and visualization of 3D environment. We propose several example applications, some of which either have been tested in our company or are currently in development.

5.1 Architectural visualization and urban planning

Joint presentation of 2D and 3D data is a frequent situation in architecture, urban development or real estate business. Our experience shows that building a mental model which merges two-dimensional floor plans and three-dimensional visualizations is not an easy task for the average user. The interactivity and the feel of direct manipulation improve this situation. We propose a usage scenario where the table holds the neighbourhood map, floor plans or geometry plans from the land register. The focused 3D model of the building is added to the augmented view. Basic user interaction such as panning and zooming can easily be performed in the table space.

This scenario can also be adapted for similar needs where the ground plans are combined with a 3D object geo-located within the 2D environment. These include: archaeology, geology or forensics.

All of them can also benefit from the secondary context where a different view of the ground plane can be displayed: satellite map, weather map, etc...

5.2 Product design and engineering

Engineering or product design usually do not embed the focus object in a spatial context but the table display can be used to display blueprints or 2D CAD drawings of an object whereas the augmented displays shows the 3D rendering of the same object.

5.3 Volume data visualization

Medical visualization or other applications of volume visualization can benefit from the 2D/3D duality of MARS when putting a segmented 3D volume subarea into the spatial context of the whole body displayed in 2D table space. For instance, the table display shows 2D image obtained from a cutting plane placed in the volume space. At the same time, the augmented display shows a specific 3D organ in the context of the cutting plane.

5.4 Exploratory data visualization

What seems to be an underestimated usage of augmented reality and multi-touch is the exploration and analysis of multidimensional data. Information visualization and exploratory data visualization in general have used multiple linked views for a long time now. With addition of two possible contexts and an augmented 3D view, the MARS seems to be an interesting choice for experiments with visualization.

6. Extensions and Future Work

In addition to the proposed MARS configuration, the set-up can be further extended in terms of user collaboration and interaction. Tablet displays can be registered relative to the table space and additional augmented views become available to additional users. This opens the floor to collaborative interaction. Further technical improvements include Kinect gestures for manipulation within the virtual 3D box above the table or using 3D display to improve the spatial immersion of the user.

We plan to further improve the technical configuration of MARS and perform additional user testing in different usage scenarios.

7. Conclusion and Acknowledgments

We developed and presented a multi-touch augmented reality system that uses two displays – one with multi-touch capability and second with augmented reality display. We especially devised the combination of these two visual spaces in a way that is seamless for the user when shifting the attention between them.

This work was supported by the Slovak Research and Development Agency under the contract No. VMSP-II-0035-09

References

- [1] Hinckley, K., Wigdor, D., Input Technologies and Techniques. Chapter 9 in The Human-Computer Interaction Handbook – Fundamentals, Evolving Technologies and Emerging Applications, Third Edition
- [2] Leonardi, K., Albertini, A., Pianesi, F. and Zancanaro, M., An exploratory study of a touch-based gestural interface for elderly, Proceedings of the 6th Nordic Conference on Human-Computer Interaction: Extending Boundaries, Reykjavik, Iceland: ACM, 2010, pp. 845-850.
- [3] Liao, H., Edwards, P.J., Pan, X., Fan Y. and Yang, G., Medical Imaging and Augmented Reality: 5th International Workshop, MIAR 2010, Beijing, China, September 19-20, 2010
- [4] Benko, H., Ishak, E.W., Feiner, S., Cross-dimensional gestural interaction techniques for hybrid immersive environments, Proceedings of IEEE Virtual Reality VR 2005, March 2005, pp. 209-216
- [5] Benko, H., Ishak, E.W., Feiner, S., Collaborative Mixed Reality Visualization of an Archaeological Excavation, Proceedings of the 3rd IEEE/ACM International Symposium on Mixed and Augmented Reality ISMAR ’04, 2004, pp. 132-140
- [6] Dedual, N., Oda, O., Feiner, S., Creating Hybrid User Interfaces with a 2D Multi-touch Tabletop and a 3D See-Through Head-Worn Display, Proceedings of the 2011 10th IEEE International Symposium on Mixed and Augmented Reality, ISMAR ‘11, pp. 231-232
- [7] Na, S., Billinghurst, M., Woo, W., TMAR: Extension of a Tabletop Interface Using Mobile Augmented Reality,Transactions on Edutainment I, Springer-Verlag Berlin, 2008, pp. 96-106
- [8] Wei, D., Zhou, S.Z., Xie, D., “MTMR: A conceptual interior design framework integrating Mixed Reality with the Multi-Touch tabletop interface”, Proceedings of the 2010 9th IEEE International Symposium on Mixed and Augmented Reality, ISMAR ‘10, pp.279-280.
- [9] Chang, R., Wang, F., You, P., A Survey on the Development of Multi-touch Technology. APWCS ‘10 Proceedings of the 2010 Asia-Pacific Conference on Wearable Computing Systems. IEEE Computer Society,USA,2010, pp. 363-366
- [10] Microsoft Corp. Redmond WA. Kinect for Xbox 360. http://www.xbox.com/en-US/KINECT, page accessed 11/2012
- [11] Krueger, M.W., Gionfriddo, T., Hinrichsen, K., VIDEOPLACE – an artificial reality. CHI ‘85 Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, USA, 1985, pp. 35-40.
- [12] Bimber, O. Raskar, R., Spatial augmented reality : merging real and virtual worlds, A.K.Peters, 2005

Matej Novotný,
VIS GRAVIS, s.r.o., E-mail: novotny@visgravis.sk

Ján Lacko,
Faculty of Informatics, Pan-European University, Bratislava, VIS GRAVIS, s.r.o.,
E-mail: lacko@visgravis.sk

Martin Samuelčík,
VIS GRAVIS, s.r.o., E-mail: samuelcik@visgravis.sk