AUGMENTED REALITY INTERACTION TECHNIQUES
Design and Evaluation of the Flip-Flop Menu

Mickael Naud, Paul Richard and Jean-Louis Ferrier

LISA Laboratory, University of Angers, 62 Avenue Notre-Dame du Lac, Angers, France

Keywords: Augmented reality, 3D interaction techniques, 3D menus, Human performance.

Abstract: We present the evaluation of a bimanual augmented reality (AR) interaction technique, and focus on the effect on viewpoint and image reversal on human performance The interaction technique (called flip-flop) allows the user to interact with a 3D object model by using a V-shaped AR menu placed on a desk in front of her/him. The menu is made of two complementary submenus. Both submenus (master and slave) are made of four Artag fiducial markers. The functionalities of the slave submenu are the following : (1) increase/decrease the size or rotate/stop the 3D object, (2) apply a color (one over four) or (3) a 2D texture (one over four) on the 3D object and (4) apply predefined material parameters. Each event is triggered by a masking of a Artag marker by the the user's right or left hand. 40 participants were instructed to perform actions such as rotate the object, apply a texture or a color on it, etc. The results revealed some difficulties due to the inversion of the image on the screen. Finally, although the proposed interaction technique is currently used for product design, it may also be applied to other fields such as edutainment, cognitive/motor rehabilitation, etc. Moreover, other tasks than the ones tested in the experiment may be archived using the menu.

1 INTRODUCTION

Virtual Reality (VR) allows a user to experience both immersion and real-time interactions that may involve visual feedback, 3D sound, haptics, smell and taste.

Instead, Augmented Reality (AR) proposes to immerse virtual objects in the real world. In this sense, AR seems to be more interesting for product design because it allows the designer or the end-user to both visualize and interact with the virtual mockup of the designed product in the real world (its working environment). There is therefore no need to build a whole 3D world (kitchen, bedroom, etc.) to estimate the esthetic properties of the product in a given context.

Although several compelling AR systems have been demonstrated (see (Azuma, 1997) for comprehensive survey), many serve merely as information browsers, allowing users to see or hear virtual data embedded in the physical world (Ishii and Ullmer, 1997).

In order to make AR fully effective, simple and efficient interaction techniques need to be developed and validated through user studies. Moreover, inherent drawbacks of AR techniques such as real/virtual image discrepancy, system calibration (free marker

AR), or ergonomic problems related to head-mounted displays have to be overcome.

In most of AR techniques designed for control application, each fiducial marker is used to trigger a single event or action. This may become a drawback when a large number of event / action are required.

In this paper, we present an evaluation of a bimanual augmented reality (AR) interaction technique based on a V-shaped menu that uses 8 fiducial markers. The menu is made of two complementary submenus. Both submenus (master and slave) are made of four ARtag fiducial markers (Fiala, 2004).

The paper is organized as follows. In section 2, we give an overview of the related work. In section 3 we describe and provide finite state modeling of the "Flip-flop" interaction technique. In section 4 we present an experimental study that was carried out in order to evaluate the "Flip-flop" interaction technique in different experimental desk-top configurations. The paper ends by a conclusion.

2 RELATED WORK

Most AR application examples of design work developed up to present are limited to large-scale objects that the designers are not able to grasp and move. In these examples, the information provided was limited to visual information, lacking physical interactions between the observer and the object. For example Fata Morgana project (Klinker et al., 2002), designers were able to walk around a newly designed virtual car for inspection and comparisons with others.

In this context, Lee et al. proposed to bring the users hand into the VE using a Mixed Reality platform (Lee et al., 2004). The hand region was separated and inserted into Virtual Environment (VE) to enhance reality.

Regenbrecht et al. developed a Magic Pen that uses ray-casting techniques for selection and manipulation of virtual objects (Regenbrecht et al., 2001a). Camera-based tracking was performed using ARToolkit, a software library that supports tangible user interaction with fiducial markers (Kato and Billinghurst, 1999). Two or more light emitting diodes (LEDs) were mounted on a real pen barrel. A camera was used to track the position of the pen, using the LEDs as position markers. Direct manipulation was performed with the end of the pen, or a virtual ray can be cast in the direction that the pen is pointing. This provides a cable-less interaction device for AR environments.

Slay et al. have developed an AR application called ARVIS (Augmented Reality Visualization) that allows users to view and interact with three-dimensional models (Slay et al., 2002). The key AR technology incorporated in the ARVIS is also AR-Toolkit. Based on ARVIS, three different interaction techniques which extend the use of ARToolkit fiducial markers were proposed.

In the first one *Array of Markers* fiducial markers are simply arranged on a surface. To change one of the attributes of the virtual object the associated fiducial marker was covered so it was no longer recognized by ARVIS. To turn the attribute back, another marker associated with the opposite of the task that was performed first was covered.

In a second interaction technique called "Menu System" the fiducial markers were used as a menuing system (Slay et al., 2002). Five small markers were created to represent five menu options. Each layer of the menuing system had a limit of five options. To select one of the options, a card has to be flipped over. The virtual representations of the cards would then change to show the next layer of the menuing system. This

was very restrictive on the design of the menu system. It was also found very cumbersome to reach tasks in the lower layers of the menu hierarchy, and because of the limited number of tasks to choose from, the technique seemed very heavy-handed.

In a third interaction technique called *Toggle Switches* the fiducial markers were used as toggle switches. This technique can be seen to stem from the *Array of Markers* interaction technique. The pattern had a fiducial marker on each side of the card. The two fiducial markers were associated with tasks that would provide opposite functions. This technique was used to connect the nodes of a directed graph together. The nodes were connected when the first side of a card was shown, and disconnected when it was flipped over. The system remembers the state that each option is in. ARVIS continually checks to see if any cards are present in the field of view. If any changes have been made it generates the appropriate VRML model and displays this as the virtual object.

The *toggle switch* system was found easy to use. Out of the three proposed interaction techniques, it was found to be the most easy and self-explanatory to use. Turning the cards over, in effect toggling the behavior exhibited by the virtual models seemed very natural.

Dias et al. have developed a series of tools based on ARToolkit : the *Paddle* and the *Magic Ring* (Dias et al., 2004). They both have a specific marker attached. As a visual aid, when a marker is recognized by the system, a virtual paddle or virtual blue circle is displayed on top of it. The Magic Ring (a small marker attached to a finger, by means of a ring) is used in several interaction tasks, such as object picking, moving and dropping, object scaling and rotation, menu manipulation, menu items browsing and selection, and for all various types of commands given to the system.

A similar AR/VR unified user interface was proposed by Piekarski and Thomas . This interface called the Tinmith-Hand is based on the flexible Tinmith-evo5 software system (Piekarski and Thomas, 2002). Using some modeling techniques, based on constructive solid geometry, tracked input gloves, image plane techniques, and a menu control system, it is possible to build applications that can be used to construct complex 3D models of objects in both indoor and outdoor settings.

In the context of collaborative design, Szalavri and Gervautz have introduced a multi-functional 3D user interface, the Personal Interaction Panel (PIP), which consists of a magnetic tracked simple clipboard and pen, containing augmented information presented to the user by see-through HMDs. The low technical level the panel and pen itself allows flexible design of

the interface and rapid prototyping. The natural two-handed interaction supported by this device makes the device fit to a rich variety of applications (Szalavri and Gervautz, 1997).

Buchmann et al have proposed a technique for natural, fingertip-based interaction with virtual objects in Augmented Reality (AR) environments (Buchmann et al., 2004). They use image processing software and finger- and hand-based fiducial markers to track gestures from the user, stencil buffering to enable the user to see their fingers at all times, and fingertip-based haptic feedback devices to enable the user to feel virtual objects. This approach allows users to interact with virtual content using natural hand gestures. It was successfully applied in an urban planning AR application.

Besides more traditional AR interaction techniques, like mouse raycast, MagicBook, and models-on-marker e.g. (Regenbrecht et al., 2001b) some techniques were proposed by Regenbrecht and Wagner (Regenbrecht and Wagner, 2002). The main one, called *cake platter* uses a turnable, plate-shaped device functions as the central location for shared 3D objects. The objects or models can be placed on the platter using different interaction techniques, e.g. by triggering the transfer from a 2D application or by using transfer devices brought close to the *cake platter*.
Another technique uses a Personal Digital Assistant (PalmPilot IIIc) as a catalogue of virtual models; the main form of interaction within the system being model selection and transfer to and from the cake platter. Still in the context of the *cake platter*, a clipping plane and lighting technique is used to see what is inside of a virtual object. The user holds a (transparent or opaque) real plane in his or her hand to clip through the model on the *cake platter*.

Others more recent applications, based on the same approach (use of Artag fiducial markers), are propose by Xin (Xin et al., 2008) and Henderson (Henderson and Feiner, 2008).

3 FLIP-FLOP INTERACTION TECHNIQUE

In most of AR techniques designed for control application, each fiducial marker is used to trigger a single event or action. This may become a drawback when a large number of event / action are required. The approach we propose here is based on a bimanual interaction and a V-shaped menu that allows to trigger many event/action with only 8 fiducial markers.

3.1 Description

The flip-flop interaction technique uses 9 ARtag fiducial markers arranged on a table, as illustrated in Figure 1. This specific number of markers was partially determined based on technical and ergonomic considerations. Firstly, the size of the markers has to be large enough to facilitate the interaction technique. Secondly, all the markers have to be always viewed by the camera and recognized. The number of markers was also determined by the number of menu options we needed to implement.

The main (central) ARtag fiducial marker is used to display the 3D model(a virtual mannequin). Side parts of the menu are made-up of 4 ARtag fiducial markers. The left part of the menu (the *master sub-menu*) allows to activate different functionalities. The right part of the menu (the *slave sub-menu*) is used to interact with the 3D model.

We called this interaction technique "flip-flop" because of the multiples back-and-forth movements that the user must do between the *master sub-menu* and the *slave sub-menu*.

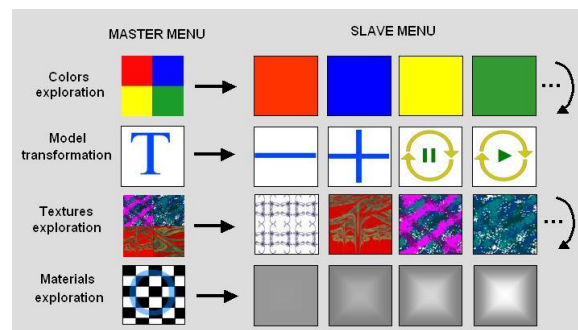Figure 1: Top view of the V-shaped fiducial markers arrangement.

Figure 2: Illustration of the functionalities of the slave sub-menu activated from the master sub-menu.

The figure 2 illustrates the functionalities of the developed application, activated by the covering of the fiducial markers of the *master sub-menu* by the user's hand. These functionalities are the following :

- **Colors exploration**: Exploration of the colors palettes (figure 3). The change of palette is done each 800ms. This value was tuned using preliminary testings involving few participant.

- **Model animation**: Activation of functionalities allowing to (1) reduce or (2) increase the size of the mannequin and (3) to make the mannequin rotate or (4) to stop it in a specific position.

- **Texture database exploration**: Exploration of the different preset texture sets. The display of a new texture sets is done automatically each 800ms.

- **Materials exploration**: Activation of functionalities allowing to change the material that simulate the fabric visual aspect.
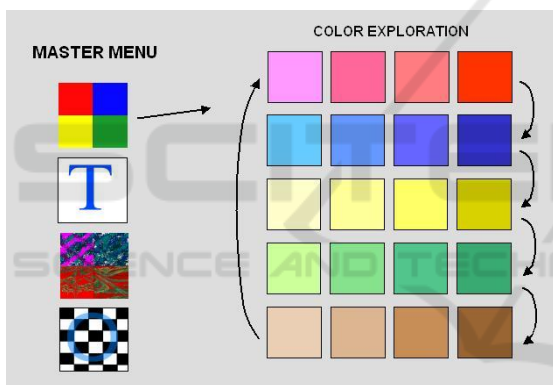

Figure 3: Exploration of the colors palettes displayed on the slave sub-menu fiducial markers from the master sub-menu.

3.2 State Automata Modelling

The automaton relative to the activation of the *slave sub-menu* functionalities from the *master sub-menu* is presented in figure 4. This automaton is composed of the following states:

- **C** : "color" (application of a color among 4)

- **I** : "image" (application of an image among 4)

- **T** : "transformation" (change of size and rotation)

- **M** : "materials" (application of materials)

The changes of state are activated by the covering of the *master sub-menu* fiducial markers by the user's hand. The following variables are involved:
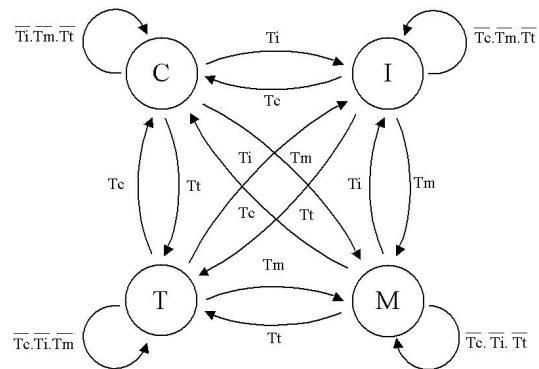
- T_c : activation of the "color" state,

Figure 4: Finite state automaton of the slave *sub-menu*.

- T_i : activation of the "image" state,

- T_t : activation of the "transformation" state,

- T_m : activation of the "materials" state.

Application of the colors, displayed on the *slave sub-menu* fiducial markers, on virtual clothes, is based on the same type of automaton as the precedent, in which states **C**, **I**, **T** and **M** are replaced by the colors C_1, C_2, C_3, and C_4, and the buttons T_c, T_i, T_t, and T_m, by the buttons T_1, T_2, T_3, and T_4.

Animation of the virtual mannequin is currently limited to a change of size (enlarging or reduction) and to a rotation. The automaton presented in figures 5 and 6 represent a modeling of the associated interaction.

The automaton relating to size modification of the model is composed of the following states:

- **A** : model size increases (by steps)

- **I** : same size
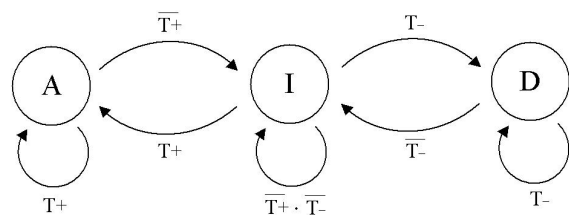
- **D** : model size decreases (by steps)

Figure 5: Finite state automaton controlling the size of 3D model.

As before, the changes of state of this automaton are activated by the covering of the *master sub-menu* fiducial markers by the user's hand. The following variables are involved:

- T^+ : increase of the size

- T^- : decrease of the size

Thus, as the user keeps covering the fiducial marker associated with the "**+**" symbol (fig. 2), the size of the model increases incrementally.

Similarly, when the user keeps covering the fiducial marker associated with the "**-**" symbol, the size of the model decreases incrementally.
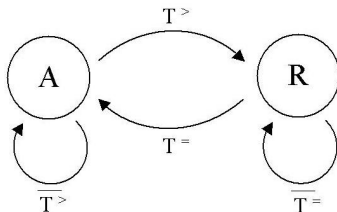

Figure 6: Finite state automaton controlling the rotation of the model.

The automaton handling the rotation of the 3D model is composed of the following states (fig. 6):

- **A** : static model
- **R** : model in rotation

Changes of state are snapped by the covering of the fiducial markers by the user's hand. *Slave sub-menu* functionalities relative to model animation are modeled by the following variables:

- $T^>$: activation of the rotation
- $T^=$: stopping the rotation

Thus, as the user keeps covering the fiducial marker associated with the $>$ symbol(fig. 2), the 3D model is rotated. Similarly, as the user keeps covering the fiducial marker associated with the $\|$ symbol, the 3D model stops.

4 EVALUATION

Augmented reality techniques may involve different configurations : (1) immersive configurations in which the users are equipped with either an optical see-through or a video see-through head-mounted display, (2) desk-top configurations that may be based on a large screen placed in front of the users, and (3) embedded configurations that involve mobile devices such as PDAs.

In our context, desk-top configurations may be interesting because the user could see himself/herself on the screen, manipulating a virtual product. This aspect is crucial in the context of fashion design where the product (clothe) are worn buy the user. In such a configuration, the video camera has to be placed in front of the user. Then the images on the screen have to be reversed.

4.1 Aim

The objective of this experiment is to study the usability, and the effectiveness of this technique for different configurations in which the video camera is placed in front or behind the user (figure 7).

4.2 Task

The task was split into 9 sub-tasks that the subjects have to perform in a sequential predefined order. These sub-tasks are the following:

1. Scroll all textures (5 sets of 4 textures)
2. Scroll colors and select the red set
3. Apply the clearest red color on the 3D model
4. Scroll colors and select the blue set
5. Apply the darkest blue color on the 3D model
6. Select transformations
7. Shrink the character (five steps)
8. Enlarge the character (three step)
9. Rotate the character and stop it (one turn)

4.3 Experimental Protocol

A total of 40 volunteers subjects (20 males and 20 females) right-handed, aged 15 to 35 years, participated in the experiment. They never used augmented reality techniques before. The subjects were divided in 4 groups (G1,G2,G3,G4). Each group has to perform the task in the following conditions :

- C1 : the webcam was placed in front of the user, and the image was not reversed
- C2 : the webcam was placed in front of the user, and the image was reversed
- C3 : the webcam was placed behind the user, and the image was reversed
- C4 : the webcam was placed behind the user, and the image was not reversed

Task has to be repeated 5 times by each subject. A 2 minutes rest period was imposed between each trial. Task completion time was measured for each single trial.
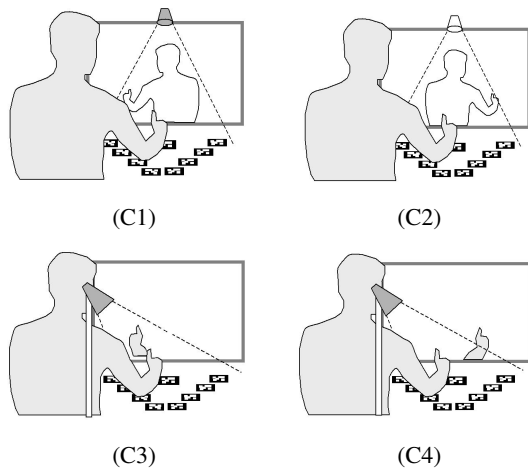
Figure 7: Experimental configurations for the evaluation of the flip-flop menu.

4.4 Results

The results illustrated in figure 8 are presented in the next three sub-sections. They were analyzed through a two-way ANOVA. We firstly examine the effect of camera position on subjects performance. Data from groups G1 and G2 were compared to the data from groups G3 et G4.

Then, the effect of image reversal on subjects performance was examined. Therefore, data from groups G1 and G3 were therefore compared to the data from groups G2 et G4. Finally we look at the joint effect of the studied parameters.
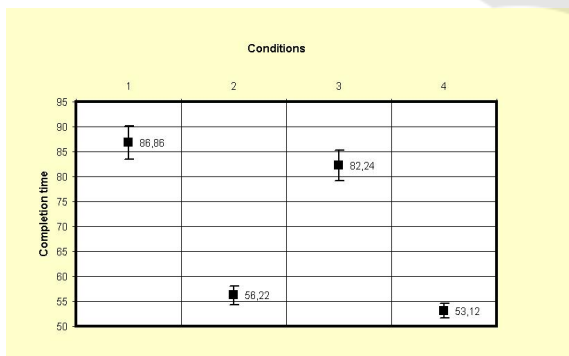

Figure 8: Mean task completion time vs. experimental conditions.

4.4.1 Effect of Camera Position

We observed that the effect of camera position on task completion time is statistically significant ($F(1,19) = 17.7, P < 0.0005$). Subjects that have the camera in front of them performed the task, on average in 71.5

seconds (std = 5.2) while those whose camera was placed behind achieved the task in 67.7 seconds (std = 4.6).

4.4.2 Effect of Image Reversal

We observed that the effect of image reversal on task completion time is very significant ($F(1,19) = 790.26, P < 0.0001$). Subjects who performed with the reversed image achieved the task, on average, in 84.55 seconds (std = 6.4) while those who performed with none reversed image achieved the task in 54.7 seconds (std = 3.4). We observe that image reversal has more significant effect on user performance than camera position.
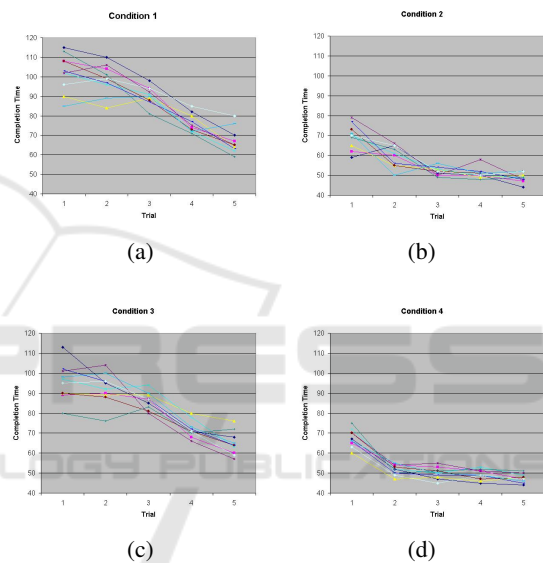

Figure 9: Task completion time over trial for condition 1 (a), condition 2 (b), condition 3 (c), and condition4 (C).

4.4.3 Joint Effect

The analysis of variance (ANOVA) reveals that the experimental conditions have a significant effect on users performance ($F(3,9) = 298.8, P < 0.0001$). The subjects of Group G1, G2, G3, and G4 have carried out the task respectively in 86.9 seconds (std = 6.6), 56.2 seconds (std = 3.8), 82.2 seconds (std = 6.2), and 53.1 seconds (std = 2.9).

Results show that the joint effect of the tested parameters is not much significant. Indeed, the subjects of the group G3 (reversed image) performed the task in 29.1 seconds faster than the subjects of group G4 (non-reversed image) when the camera was positioned behind, while the subjects of group G1 (non reversed image) performed the task in 30.7 seconds more than the subjects of group G2 (reversed image), when the camera was positioned in front.

4.4.4 Learning Process

Learning is defined here by the improvement of subjects performance during task repetitions. We asked subjects to repeat 5 times the previously defined task 4.2. The results show that the subjects of group G1 achieved the task in 102.5 seconds (std = 9.6) during the first trial and in 65.7 seconds (std = 6.5) during the last trial. Subjects of groups G2, G3 and G4, achieved the task respectively in 69.5 seconds (std = 6.2), 95.5 seconds (std = 9.0) and 67.4 seconds (std = 4.4) during the first trial and 49.1 seconds (std = 2.6), 64.8 seconds (std = 5.9) and 47.5 seconds (std = 2.2) during the last trial. This result in performance improvement percentage of 34%, 29%, 32% and 30% respectively for groups G1, G2, G3 and G4.

Figures 9(a), 9(b), 9(c) and 9(d) show learning curves of each subject. We observed that for configurations in which the image is reversed, learning is slower.

5 CONCLUSIONS

We presented the evaluation of a bimanual augmented reality (AR) interaction technique, and focus on the effect on viewpoint and image reversal on human performance The interaction technique (called flip-flop) allows the user to interact with a 3D object model by using a V-shaped AR menu placed on a desk in front of her/him. The menu is made of two complementary submenus. Both submenus (master and slave) are made of four Artag fiducial markers. The functionalities of the slave submenu are the following : (1) increase/decrease the size or rotate/stop the 3D object, (2) apply a color (one over four) or (3) a 2D texture (one over four) on the 3D object and (4) apply predefined material parameters. Each event is triggered by a masking of a Artag marker by the the user's right or left hand. 40 participants were instructed to perform actions such as rotate the object, apply a texture or a color on it, etc. The results revealed some difficulties due to the inversion of the image on the screen. Finally, although the proposed interaction technique is currently used for product design, it may also be applied to other fields such as edutainment, cognitive/motor rehabilitation, etc. Moreover, other tasks than the ones tested in the experiment may be archived using the menu.

REFERENCES

Azuma, R. (1997). A survey of augmented reality. In *Presence: Teleoperators and Virtual Environments*, pages 355–385.

Buchmann, V., Violich, S., Billinghurst, M., and Cockburn, A. (2004). Fingartips: gesture based direct manipulation in augmented reality. In *GRAPHITE '04: Proceedings of the 2nd international conference on Computer graphics and interactive techniques in Australasia and South East Asia*, pages 212–221, New York, NY, USA. ACM.

Dias, J. M. S., Santos, P., and Bastos, R. (2004). Gesturing with tangible interfaces for mixed reality. In Camurri, A. and (Eds.), G. V., editors, *GW 2003,LNAI 2915*, page 399408, Berlin Heidelberg.

Fiala, M. (2004). Artag, an improved marker system based on artoolkit. In *NRC Publication Number: NRC 47166*.

Henderson, S. and Feiner, S. (2008). Opportunistic controls : Leveraging natural affordances as tangible user interfaces for augmented reality. In *VRST 2008, ACM Symposium on Virtual Reality Software and Technology*, pages 211–218, Bordeaux, France.

Ishii, H. and Ullmer, B. (1997). Tangible bits towards seamless interfaces between people. In *Bits and Atoms. CHI 97.*, pages 234–241.

Kato, H. and Billinghurst, M. (1999). Marker tracking and hmd calibration for a video-based augmented reality conferencing system. In *In 2nd IEEE and ACM International Workshop on Augmented Reality*, pages 84–94, San Francisco, USA.

Klinker, G., Dutoit, A., Bauer, M., Bayer, J., Novak, V., and Matzke, D. (2002). Fata morgana - a presentation system for product design. In *ISMAR2002*, pages 76–85.

Lee, S., Chen, T., Kim, J., Kim, G. J., Han, S., and Z., P. (2004). Using virtual reality for affective properties of product design. In *Proc. of IEEE Intl. Conf. on Virtual Reality*, pages 207–214.

Piekarski, W. and Thomas, B. H. (2002). Unifying augmented reality and virtual reality user interfaces.

Regenbrecht, H., Baratoff, G., Poupyrev, I., and Billinghurst, M. (2001a). A cable-less interaction device for ar and vr environnement. In *Proc. ISMR*, pages 151–152.

Regenbrecht, H., Baratoff, G., and Wagner, M. (2001b). A tangible ar desktop environment. In *Computers and Graphics, Special Issue on Mixed Realities*.

Regenbrecht, H. T. and Wagner, M. T. (2002). Interaction in a collaborative augmented reality environment. In *CHI 2002*, Minneapolis, Minnesota, USA.

Slay, H., Thomas, B., and Vernik, R. (2002). Tangible user interaction using augmented reality. In *Third Australasian User Interfaces Conference (AUIC2002)*, Melbourne, Australia.

Szalavri, Z. and Gervautz, M. (1997). The personal interaction panel - a two-handed interface for augmented

reality. In Publishers, B., editor, *EUROGRAPHICS 97*.

Xin, M., Sharlin, E., and Costa Sousa, M. (2008). Napkin sketch - handheld mixed reality 3d sketching. In *VRST 2008, ACM Symposium on Virtual Reality Software and Technology*, pages 223–226, Bordeaux, France.