Virtual Object Manipulation on a Table-Top AR Environment

H. Kato[1], M. Billinghurst[2], I. Poupyrev[3], K. Imamoto[1], K. Tachibana[1]

[1]*Faculty of Information Sciences, Hiroshima City University*
3-4-1, Ozuka-higashi, Asaminami-ku, Hiroshima, 731-3194 JAPAN
kato@sys.im.hiroshima-cu.ac.jp

[2]*HIT Laboratory, University of Washington*
Box 352-142, Seattle, WA 98195, USA
grof@hitl.washington.edu

[3]*ATR MIC Laboratories, ATR International,*
2-2 Hikaridai, Seika-cho, Soraku-gun, Kyoto, Japan
poup@mic.atr.co.jp

Abstract

In this paper we address the problems of virtual object interaction and user tracking in a table-top Augmented Reality (AR) interface. In this setting there is a need for very accurate tracking and registration techniques and an intuitive and useful interface. This is especially true in AR interfaces for supporting face to face collaboration where users need to be able to easily cooperate with each other. We describe an accurate vision-based tracking method for table-top AR environments and tangible user interface (TUI) techniques based on this method that allow users to manipulate virtual objects in a natural and intuitive manner. Our approach is robust, allowing users to cover some of the tracking markers while still returning camera viewpoint information, overcoming one of the limitations of traditional computer vision based systems. After describing this technique we describe it's use in a prototype AR applications.

1. Introduction

In the design session of the future several architects sit around a table examining plans and pictures of a building they are about to construct. Mid-way through the design session they don light-weight see-through head mounted displays (HMDs). Through the displays they can still see each other and their real plans and drawings. However in the midst of the table they can now see a three-dimensional virtual image of their building. This image is exactly aligned over the real world so the architects are free to move around the table and examine it from any viewpoint. Each person has their own viewpoint into the model, just as if they were seeing a real object. Since it is virtual they are also free to interact with the model in real time, adding or deleting parts to the building or scaling portions of it to examine it in greater detail. While interacting with the virtual model they can also see each other and the real world, ensuring a very natural collaboration and flow of communication.

While this may seem to be a far-off vision of the future there are a number of researchers that have already developed table-top AR systems for supporting face-to-face collaboration. In Kiyokawa's work two users are able to collaboratively design virtual scenes in an AR interface and then fly inside those scenes and experience them immersively [Kiyokawa 98]. The AR2 Hockey system of Ohshima et. al. [Ohshima 98] allows two users to play virtual air hockey against each other, while the Shared Space interface supports several users around a table playing a collaborative AR card matching game [Billinghurst 99]. Finally the Emmie system of Butz et. al. [Butz 99] combines virtual three-dimensional AR information with conventional two-dismensional displays in a table-top system that supports face-to-face collaboration.

There are collaborative AR environments that do not rely on a table-top setting, such as Studierstube [Schmalsteig 96], however it is clear that this is an important category of AR interface. This is due to a number of reasons:

- In face-to-face meetings, people typically gather around a table.
- A table provides a location for placing material relative to meeting content.
- A table provides a working surface for content creation.

In creating an AR interface that allows users to manipulate 3D virtual objects in a real table-top there are a number of problems that need to be overcome. From a technical viewpoint we need to consider tracking and registration accuracy, robustness and the overall system configuration
From a usability viewpoint we need to create a natural and intuitive interface and address the problem of allowing real objects to occlude virtual images.

In this paper we describe some computer vision based techniques that can be used to overcome these problems. These techniques have been designed to support a

Tangible Augmented Reality (TAR) approach in which lessons from Tangible User Interface (TUI) design are applied to the design of AR interfaces. In the next section we describe the idea of Tangible AR interfaces in more detail and in section 3 some results from early prototypes of our Table-top AR interfaces. In section 4 our current registration and interaction techniques are described. Finally in section 5 we present our most recent prototype system based on our method and we conclude in section 6.

2. Tangible Augmented Reality

Although there have been many different virtual object manipulation techniques proposed for immersive virtual reality environments, there has been less work conducted on AR interaction techniques. One particularly promising area of research that can be applied is the area of Tangible User Interfaces. The goal of Tangible User Interface research is to turn real objects into input and output devices for computer interfaces [Tangible 2000].

Tangible interfaces are powerful because the physical objects used in them have properties and physical constraints that restrict how they can be manipulated and so are easy to use. However there are limitations as well. It can be difficult to change these physical properties, making it impossible to tell from looking at a physical object what is the state of the digital data associated with that object. In some interfaces there is also often a disconnect between the task space and display space. For example, in the Gorbet's Triangles work, physical triangles are assembled to tell stories, but the visual representations of the stories are shown on a separate monitor distinct from the physical interface [Gorbet 98].

The visual cues conveyed by tangible interfaces are also sparse and may be inadequate for some applications. The ToonTown remote conferencing interface uses real dolls as physical surrogates of remote people [Singer 99]. However the non-verbal and visual cues that these objects can convey is limited compared to what is possible in a traditional videoconference. Showing three-dimensional imagery in a tangible setting can also be problematic because it is dependent on a physical display surface.

Many of these limitations can be overcome through the use of Augmented Reality. We define Tangible Augmented Reality as AR interfaces based upon Tangible User Interface design principles. In these interfaces the intuitiveness of the physical input devices can be combined with the enhanced display possibilities provided by virtual image overlays. Head mounted display (HMD) based AR provides the ability to support independent public and private views of the information space, and has no dependence on physical display surfaces. Similarly, AR techniques can be used to seamlessly merge the display and task space.

Research in immersive virtual reality point to the performance benefits that can result from a Tangible Augmented Reality approach. The physical properties of the tangible interface can be used to suggest ways in which the attached virtual objects might interact and enhance the virtual interaction. For example, Lindeman finds that physical constraints provided by a real object can significantly improve performance in an immersive virtual manipulation task [Lindeman 99]. Similarly Hoffman finds adding real objects that can be touched to immersive Virtual Environments enhances the feeling of Presence in those environments [Hoffman 98]. While in Poupyrev's virtual tablet work, the presence of a real tablet and pen enbale users to easily enter virtual handwritten commands and annotations [Poupyrev 98].

Interfaces that combine Reality and Virtuality are not new. However, Ishii summarizes the state of AR research when he says that AR researchers are primarily concerned with ".. considering purely visual augmentations" rather than the form of the physical objects those visual augmentations are attached to [Ishii 97]. If we are to create more usable AR interfaces then researchers must have a better understanding of design principles based on form as well as function.

In our augmented reality work we advocate designing the form of physical objects in the interface using established Tangible User Interface design methods. Some of the tangible design principles include:

- Object affordances should match the physical constraints of the object to the requirements of the task.
- The ability to support parallel activity where multiple objects or interface elements are being manipulated at once.
- Support for physically based interaction techniques (such as using object proximity or spatial relations).
- The form of objects should encourage and support spatial manipulation
- Support for multi-handed interaction.

Physical interface attributes are particularly important in interfaces designed to support face-to-face collaboration. In this case people commonly use the resources of the physical world to establish a socially shared meaning [Gav 97]. Physical objects support collaboration both by their appearance, the physical affordances they have, their use as semantic representations, their spatial relationships, and their ability to help focus attention. In an AR interface the physical objects can further be enhanced in ways not normally possible such as providing

dynamic information overlay, private and public data display, context sensitive visual appearance, and physically based interactions.

In the next section we describe how the Tangible Augmented Reality approach was applied in an early collaborative table-top AR experience.

3. Case Study: Shared Space Siggraph 99

The Shared Space Siggraph 99 application was designed to explore how augmented reality could be used to enhance face to face collaboration in a table-top setting. In order to do this we aimed to develop a compelling collaborative AR experience that could be used by novices with no training or computer experience. We based this experience on a simple child's card matching game. In our variant three people around a table wear Olympus HMDs with cameras attached (figure 1).

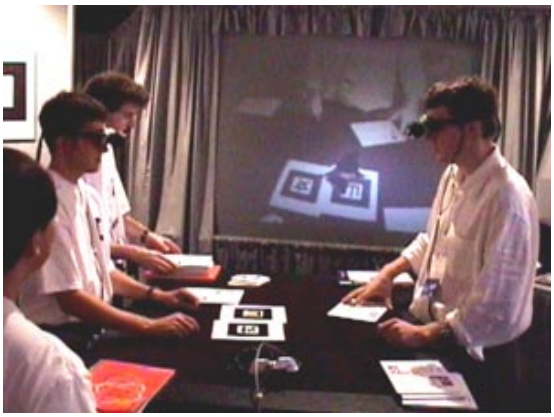

Fig. 1: Users Around the Playing Table

On the table there are large cards with Japanese Kanji characters on them. When the users turn over the cards they see different three-dimensional virtual objects appearing on top of the cards (figure 2).

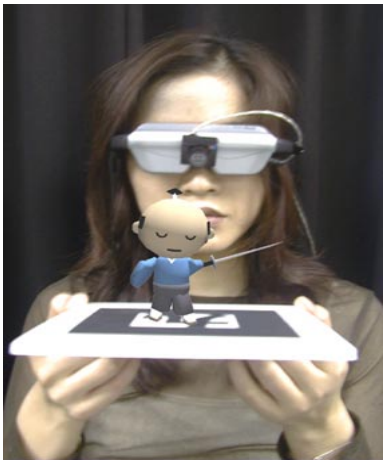

Fig. 2: A Virtual Object on a Card

The goal of the game is to collaboratively match objects that logically belong together. When cards containing correct matches are placed side by side an animation is triggered involving the objects (figure 3a,3b). For example, when the card with the UFO on it is placed next to the card with the alien on it the alien appears to jump into the UFO and start to fly around the Earth. Since the players are all co-located they can easily all see each other and the virtual objects that are being exposed.

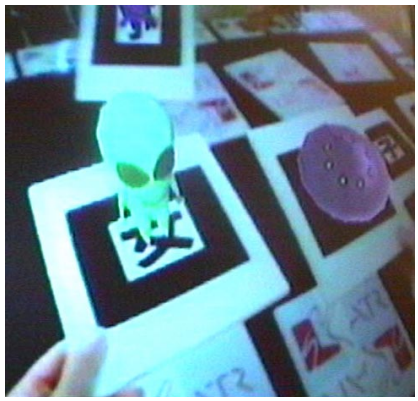

Fig. 3a: Two Matching Objects Being Brought Together

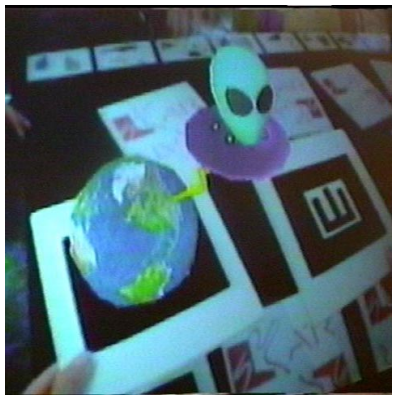

Fig. 3b: The Virtual Object Interaction

The HMD and camera are connected to an SGI O2 computer that performs image processing on the video input and composites computer graphics onto the image for display in the HMD. The users experience a video see-through augmented reality, seeing the real world through the video camera. The real cards are all labeled with square tracking markers. When users look at these cards, computer vision techniques are used to find the tracking mark and determine the exact pose of the head mounted camera relative to it [Kato 99a]. Once the position of the real camera is known, a virtual image can then be exactly overlaid on the card. Figure 4 overleaf summarizes the tracking process.

Although this is a very simple application it provides a good test of the usefulness of the tangible interface metaphor for manipulating virtual models. The Kanji

characters are used as tracking symbols by the computer vision software and were mounted on flat cards to mimic the physical attributes people were familiar with in normal card games. This was to encourage people to manipulate them the same way they would use normal playing cards. However, the tracking patterns needed to be placed in such a way that people would not cover them with their hands when picking the cards up, and they needed to be large enough to be seen from across the table. So there was a design trade-off between making the cards large enough to be useful for the tracking software and too large that they could not easily be handled. The physically based interaction techniques were also chosen based on natural actions people perform with playing cards, such as turning them over, rotating them, holding them in the hands, passing them to each other and placing them next to each other.

3.1 User Experiences

The Shared Space demonstration has been shown at the SIGGRAPH 99 and Imagina 2000 conferences and the Heniz-Nixdorf museum in Germany. Over 3,500 people have tried the software and given us feedback.

Users had no difficulty with the interface. They found it natural to pick up and manipulate the physical cards to view the virtual objects from every angle. Once they held a card in view and could see a virtual object, players typically only made small head motions. However it was common to see people rotating the cards at all angles to see the virtual objects from different viewpoints. Since the matches were not obvious some users needed help from other collaborators at the table and players would often spontaneously collaborate with strangers who had the matching card they needed. They would pass cards between each other, and collaboratively view objects and completed animations. They almost always expressed surprise and enjoyment when they matched virtual objects and we found that even young children could play and enjoy the game. Users did not need to learn any complicated computer interface or command set. The only instructions people needed to be given to play the game was to turn the cards over, not cover the tracking patterns and to find objects that matched each other.

At the Imagina 2000 conference 157 people filled out a short user survey. They were asked to answer the following questions on a scale of one to seven (*1= very easily/real* and *7* = not very easily/real):

1: How easily could you play with other people ?
2: How real did the virtual objects seem to you?
3: How easily could you interact with the virtual objects?

Table 1 summarizes the results. As can be seen, users felt that they could very easily play with the other people (5.64) and interact with the virtual objects (5.62). Both of these are significantly higher than the neutral value of 3.5; the t-test value row showing the results from a one-tailed t-test. It is also interesting that even though the virtual object were not real, on average people rated them as being midway between not very real and very real. When asked to fill what they enjoyed most about the system the top three responses were: the interactivity (25), the ease of use (18), and how fun it was (15).

	Q1 (n=132)	Q2 (n=157)	Q3 (n=157)
Average	5.64	4.01	5.62
Std Dev.	1.19	1.20	1.20
t-test val.	237.09	66.70	278.74

Table 1: Shared Space Survey Results

These results illustrate that by applying a tangible interface metaphor we are very able to create a compelling table-top AR experience in which the technology was transparent. In the next section we describe in more detail our current tracking and interaction techniques which overcome some of the limitations of the Shared Space Siggraph 99 application, including occlusion of virtual images by real objects, robust tracking, and a limited range of tangible interaction methods.

4. An Improved Method

In the previous section we described our Shared Space

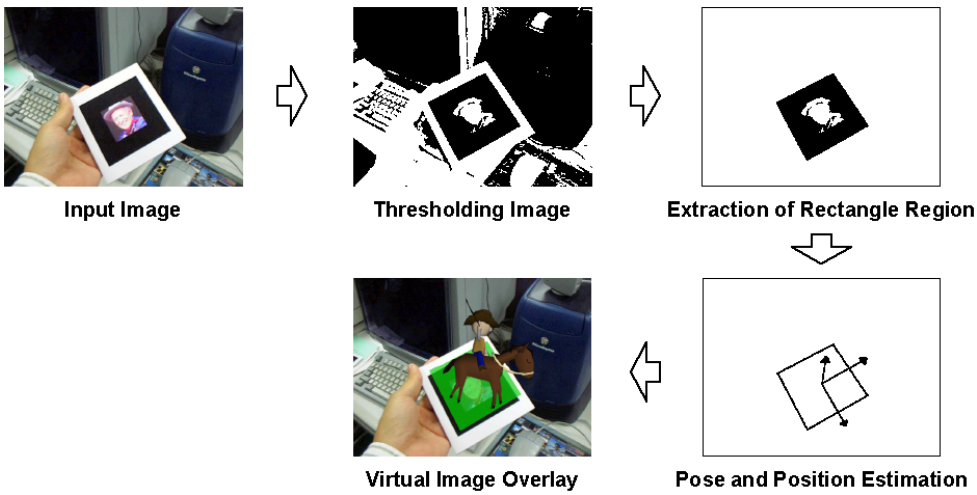

Figure 4: The Vision-Based AR Tracking Process

Siggraph 99 collaborative AR application which was based on our computer vision tracking technique and a TUI design method. Although users found this a successful Tangible AR interface and were able to collaborate easily with each other, there were a number of shortcomings. First the tracking method only provided user head position relative to each of the cards in view, not to any global world coordinate system. This makes it difficult to implement certain types of Tangible Interaction techniques. Secondly, since the vision-based tracking used single large markers the system failed when a tracking marker was partially covered by a user's hand or other object. Finally, we didn't solve the problem of the real cards not being able to occlude the virtual models on other cards, causing foreground/background confusion. In this section we describe a new approach to table-top AR that overcomes these limitations.

4.1 Implementing Global Coordinate Tracking

In order to track user and object position we modified the table-top AR environment by attaching tracking fiducials to the table top surface. Figure 5 shows the new system configuration.

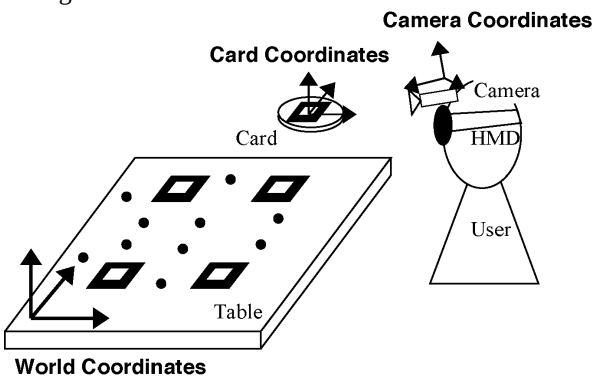

Figure 5 Table-top Configuration.

The table-top fiducials consist of a mixture of square tracking patterns with small circular blobs between them. We define the world coordinates frame as a set of coordinate axes aligned with the table surface. The camera attached to the HMD detects the self-pose and position in the world coordinates by looking at multiple fiducials on the table. In section 4.2 we describe the vision-based tracking method used for head tracking from multiple fiducials. Our method is robust to partial occlusion, so users can move their hands across the table-top and the camera position is still reliably tracked. Finding the user head position in world coordinates means that 3D virtual objects can also be represented in the world coordinates and the user can see them appearing on the on the real table.

The user can also still pick up an object on which a fiducial is drawn, and our previous method can be used to calculate the relationship between the object and camera coordinates. However because the camera pose in world coordinates is known, we can now find the object pose in the world coordinate frame. Using this information we can use new manipulation methods based on object pose and movement. These are described in section 4.4.

Since this configuration uses only one camera as a sensor, it is compact and could be portable. Even if there are multiple people around the table, the systems for each user do not interfere so our global tracking approach scales to any number of users. In fact, information from several users could be integrated to increase the accuracy or robustness, although this still needs to be done.

4.2 Tracking of Multiple Fiducials

Our previous tracking method provides satisfactory accuracy for a table-top AR environment, however it uses a single relatively large square marker as a fiducial. So if a hand or other object to even partially overlapped the fiducial the tracking was lost. This decreased the robustness of tracking under the conditions where a hand could overlap the fiducials. Also if there is some distance between tracked fiducials and displayed virtual objects, tracking errors strongly influence the registration accuracy. That is, using a single fiducial decreases the accuracy of registration under the conditions where virtual objects need to be displayed around on the table.

We have developed a new tracking method in which multiple large square and blobs are used as fiducials and pose and position are estimated from all of the detected fiducial marks. This means that many of the fiducial can be covered up without losing tracking. Many tracking methods using multiple markers have been proposed at such conferences as IWAR99 or ISMR99. However there are few methods that use combination of different types of tracking markers.

The square marker used previously has the characteristic that 3D pose and position can be estimated from a single marker. The same results can be achieved by using a set of circular blobs. Since circular blobs are relatively small and can be spread over a wider area, it is more difficult to cover them all. However the disadvantage is that three blobs are required for pose and position estimation and identification of each blob is difficult from visible features. Therefore another method for identification of each blob has to be adopted. Our tracking method uses the features of both the square and blob markers. As shown in figure 6, multiple squares and blobs lie on the table spread over a wide area. The relationships among all markers are known and are described in world coordinates.

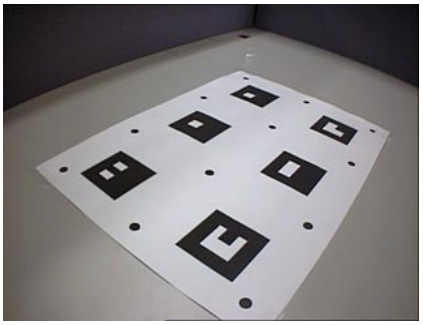

Figure 6 An Example of Fiducials.

Considering just the square markers, there are two situations that might occur in the captured video image:

1) One or more square markers are visible.
2) No square markers are visible.

In the rest of this section we explain how we can achieve robust pose tracking in each of these circumstances.

1) One or More Squares are Visible

If there is a square marker in the image, it is possible to estimate 3D pose and position using our earlier method [Kato 99a]. However if there is more than one square visible we can achieve more robust tracking if we estimate pose from all of available features. In order to do this we adopt following procedures:

step 1) The biggest square marker is selected in the image. 3D pose and position are initially estimated from it using our earlier method. This information is represented as the following transformation function from marker coordinates to camera coordinates:

$$(x_c, y_c, z_c) = trans(x_w, y_w, z_w) \qquad \text{(eq.1)}$$

where (x_w, y_w, z_w) is a position in world coordinates and (x_c, y_c, z_c) is the same position in camera coordinates.

step 2) The positions of all the circular blobs are estimated in screen coordinates by using the above transformation function, a projective function and the 3D positions of blobs in the world coordinates:

$$(x_s, y_s) = perspect(\ trans(x_w, y_w, z_w)\ ) \qquad \text{(eq.2)}$$

where the function *perspect* is a projective function. This function consists of perspective projection parameters and image distortion parameters [Kato 99b].

step 3) The actual screen coordinates of the detected blobs are compared to the estimated positions. Using the positions of all successfully matched blob markers and the 4 vertices of all extracted square markers, the 3D pose and position are re-estimated. For this calculation, the initial transformation function is used and modified as the amount of errors between the actual feature positions in the image and the estimated positions goes to minimum using a hill-climbing method.

2) No Square Markers are Visible

In this case, we assume that some of the circular blobs are visible so a procedure for robust identification of blob markers is needed. If we assume that the video capture rate is sufficiently fast then there is little difference in blob position between frames. So we can use the blobs positions that are estimated at last frame containing a square marker and then track these over subsequent frame. The blob positions in the frame with the square marker are found using the above method.

This method of tracking blobs from frame to frame works well when head motion is not too fast and a hand moves to overlap some of the square markers. As we discovered in the Shared Space Siggraph 99 application, rapid hand motion is more likely than rapid head motion. However if the head moves quickly in condition where only dot markers can be seen the tracking will fail. In order to decrease this possibility the layout of fiducials is also important.

Figure 7 shows an example of the tracking. In figure 7a both square and blob markers are visible, while in figure 7b some square markers are covered by a hand. In this case, we can see that virtual objects are still displayed on the correct position. However, we can also we can see the incorrect occlusion between the virtual objects and the hand. In the next section we describe how to address this problem.

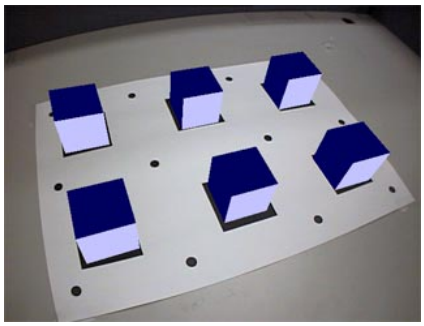

Figure 7a: Virtual Objects on Multiple Markers

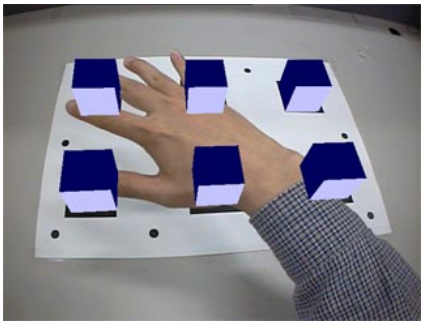

Figure 7b: Markers Covered by a Hand

4.3 The Occlusion Problem

When integrating real and virtual objects, if depth information is not available, problems with incorrect occlusion can result. That is, a virtual object which should be far from the user sometimes occludes a real object that is nearer to the user. This problem prevents a user from recognizing depth information and decreases usability. Yokoya proposed a method that overcomes this problem by getting depth information from stereo cameras [Yokoya 99]. This could be achieved by two cameras and fast computer.

With regard to table-top virtual object manipulation this problem mostly arises between a hand which manipulates virtual objects and the virtual objects on the table. As the person moves their hand above the table the virtual objects on the table surface incorrectly appear in front of the hand (see figure7b). Considering this problem we arrived at the following solutions.

1) We restrict users to interacting with virtual images with physical objects they hold in their hands. These objects can have a fiducial marker on them so the position and pose can be detected. Also the shape of the object is known. Thus using virtual models of the hand-held real objects we can correctly occlude the virtual models. That is, far-off virtual objects might cover the user's hand but the real object manipulating the virtual objects correctly occludes them. We hypothesize that this will affect usability less than a total absence of occlusion support.

2) Since there are no virtual objects in the naturally occurring in the real world, we think that user's will not find it unnatural that virtual objects have transparency. Therefore we hypothesize that a user will not object if virtual objects cannot completely occlude real objects. This is especially the case in optical-see through AR where every virtual object is at least a little transparent making it is difficult for them to cover a real object perfectly.

These can be realized by using Alpha-buffer and Z-buffer information when rendering. Figure 8a shows a physical object correctly occluding virtual objects. In this figure, we can see all depth information is correctly represented except for the hand.

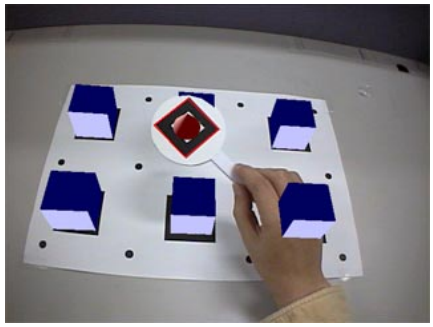

Figure 8a: correct overlay of a physical object

Figure 8b shows virtual objects with a little transparency. In this case, even if the depth information of the hand is still incorrect, we can see the hand because of the transparency, reducing the visual discrepancy.

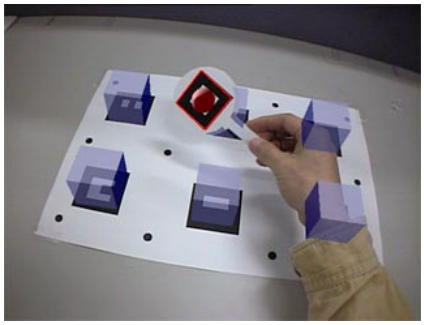

Figure 8b: transparent virtual objects

4.4 Implementing Natural and Intuitive Manipulation

In the Shared Space Siggraph 99 application users were able to easily interact with the application because the physically based interaction techniques matched the affordances of the real cards. However because the cards were not tracked relative to global coordinates there were only a limited number of manipulation methods that could be implemented.

If the virtual objects are attached to a card, or manipulated by a card there are a number of other possible manipulation methods that could be explored:

- *Inclining:* If the card the virtual object is on is tilted, the object should slide across the card surface.
- *Pushing down:* When a card pushes down a virtual object on the table, it should disappear into the table.
- *Picking & pulling:* When a card picks a virtual object on the table from above it, it should

appear to be connected with a card by short virtual string. Pulling the string can then move it.

- *Shaking:* When shaking a card, an object could appear on the card or change to another object.

Some of these commands simulate physical phenomena in the real world and other simulate table magic. In all these cases we establish a cause-and-effect relationship between physical manipulation of the tangible interface object and the behavior of the virtual images.

These behaviors can be implemented using knowledge about the real object position and orientation in world coordinates. There are two classes of physical interaction techniques. One in which behaviors can be determined purely from knowing the relationship between card coordinates and camera coordinates. Card *shaking* belongs to this class. The other is a class in which behaviors can be determined by using two relationships: between card and camera coordinates and between world and camera coordinates. Behaviors such as *inclining*, *picking* and *pushing* belong to this class. In the remainder of this section we show how to recognize examples of these behaviors.

Detecting Type A Behaviors: *Shaking*

A series of detected transformation matrices from the card to camera coordinate frames are stored over time. Observing rotation and translation components from these matrices, the user behavior can be determined. For the *shaking* behavior,

1) The pose and position at t[sec] before the current time are almost same as current pose and position.
2) There is little changes in the card rotation period.
3) There is a time when the card is moved farther than y [mm] in surface plane of the card.
4) There is little movement in the surface normal direction of the card.

When all the above conditions are satisfied, it is assumed that the user is *shaking* the physical card and the corresponding *shaking* command is executed.

Detecting Type B Behaviors: *Inclining and Pushing*

When the camera pose and position and a card pose and position are detected, a transformation matrix between the card coordinate frame and world coordinate frame can be calculated. Observing the rotation and translation components of this transformation matrix, behaviors such as card tilting and pushing can be determined. At this time, the pose, position and size of virtual objects on the table are also be used to determine the user interaction.

5. Prototype System

We are currently developing a prototype table-top AR system for virtual interior design using the interaction and tracking techniques described above. Figure 9 shows the current version of this prototype. As can be seen users are able to user a real paddle to move around virtual objects in the AR interface. There is correct occlusion between the paddle and the virtual objects and transparency cues are use to minimize the hand occlusion problem. Multiple users can gather around the table-top and simultaneously interact with the virtual scene. Using this system, we plan to conduct user studies to explore the effects of Tangible AR interfaces on face to face collaboration.

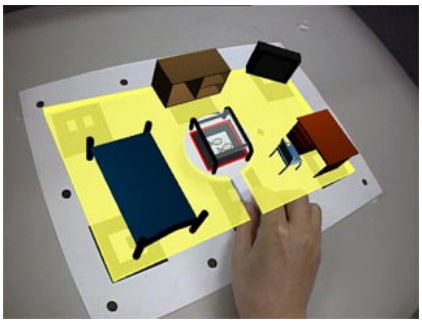

Figure 9 A Prototype of an Interior Design Application

6. Conclusions

In this paper we addressed the problems of virtual object interaction and user tracking in a table-top Augmented Reality (AR) interface. We first described an approach to AR interface design based on Tangible User Interface design principles. Next we showed how using these design principles we were able to create a compelling table-top AR experience which could be used by novices with no computer experience. Coupling a tangible interface with AR imagery achieved a technology transparency that enhanced face to face collaboration. However there were problems with the tracking approach and the limited types of interaction method support in the Shared Space Siggraph 99 experience.

In the second half of the paper we address these issues. We presented a more accurate and robust vision-based tracking method for table-top AR environments that finds pose information from multiple fiducial marks. This tracking technique also allows us to track users and card in world coordinates. Tangible user interface (TUI) techniques based on this method that allow users to manipulate virtual objects in a natural and intuitive manner. We are currently developing a virtual interior design application so we can further explore the effect of AR tangible user interface in table-top collaboration.

References

[Billinghurst 99] Billinghurst, M., Kato, H., Kraus, E., May, R. Shared Space: Collaborative Augmented Reality.

In *Visual Proceedings, SIGGRAPH 99*, August 7-12th, Los Angeles, CA, ACM Press, 1999.

[Butz 99] A. Butz, T. Höllerer, S. Feiner, B. MacIntyre, C. Beshers, Enveloping Users and Computers in a Collaborative 3D Augmented Reality, *In Proc. IWAR '99*, San Francisco, CA, October 20-21, 1999, pp. 35-44.

[Gav 97] Gav, G., Lentini, M. Use of Communication Resources in a Networked Collaborative Design Environment. http://www.osu.edu/units/jcmc/IMG_JCMC/ResourceUse.html

[Gorbet 98] Gorbet, M., Orth, M., Ishii, H. Triangles: Tangible Interface for Manipulation and Exploration of Digital Information Topography. *In Proceedings of CHI 98*, Los Angeles, CA, 1998.

[Hoffman 98] Hoffman, H. Physically Touching Virtual Objects Using Tactile Augmentation Enhances the Realism of Virtual Environments. In *Proceedings of Virtual Reality Annual International Symposium (VRAIS '98)*, 1998, pp. 59-63.

[Ishii 97] Ihsii, H., Ullmer, B. Tangible Bits: Towards Seamless Interfaces between People, Bits and Atoms. In *Proceedings of CHI 97*, Atlanta, Georgia, USA, ACM Press, 1997, pp. 234-241.

[Kato 99a] H. Kato, M. Billinghurst: Marker Tracking and HMD Calibration for a Video-based Augmented Reality Conferencing System, *In Proc. IWAR '99*, San Francisco, CA, October 20-21, 1999, pp.85-94.

[Kato 99b] H. Kato, M. Billinghurst, K. Asano, K. Tachibana, An Augmented Reality System and its Calibration based on Marker Tracking, *Transactions of the Virtual Reality Society of Japan*, Vol.4, No.4, pp.607-616, 1999 (in Japanese).

[Kiyokawa 98a] Kiyokawa, K., Iwasa, H., Takemura, H., Yokoya, N. Collaborative Immersive Workspace through a Shared Augmented Environment, In *Proceedings of the International Society for Optical Engineering '98 (SPIE '98)*, Vol.3517, pp.2-13, Boston, 1998.

[Lindeman 99] Lindeman, R., Sibert, J., Hahn, J. Towards Usable VR: An Empirical Study of User Interfaces for Immersive Virtual Environments. In *Proceedings of CHI 99*, 15th-20th May, Pittsburgh, PA, 1999, pp. 64-71.

[Ohshima 98] Ohshima, T., Sato, K., Yamamoto, H., Tamura, H. AR^{2}Hockey: A case study of collaborative augmented reality, In *Proceedings of VRAIS'98*, 1998, IEEE Press: Los Alamitos, pp.268-295.

[Poupyrev 98] Poupyrev, I., Tomokazu, N., Weghorst, S., Virtual Notepad: Handwriting in Immersive VR. In *Proceedings of IEEE VRAIS'98*, 1998, pp.126-132.

[Schmalsteig 96] Schmalsteig, D., Fuhrmann, A., Szalavari, Z., Gervautz, M., Studierstube - An Environment for Collaboration in Augmented Reality. In *CVE '96 Workshop Proceedings*, 19-20th September 1996, Nottingham, Great Britain.

[Singer 99] Singer, A., Hindus, D., Stifelman, L., White, S. Tangible Progress: Less is More in Somewire Audio Spaces. In *Proceedings of CHI 99*, 15th-20th May, Pittsburgh, PA, 1999, pp. 104 – 111.

[Tangible 2000] MIT Media Lab, Tangible Media Group http://tangible.www.media.mit.edu/groups/tangible/

[Yokoya 99] N. Yokoya, H. Takemura, T. Okuma, M Kanbara, Stereo Vision Based Video See-through Mixed Reality, *Mixed Reality (Proc. Of ISMR99)*, Springer-Verlag, 1999, pp.131-145.