The ARTable: An AR-Based Tangible User Interface System

Youngmin Park and Woontack Woo

GIST U-VR Lab., Gwangju 500-712, S.Korea
{ypark, wwoo}@gist.ac.kr
http://uvr.gist.ac.kr

Abstract. Augmented Reality (AR) and Tangible User Interface (TUI) have been proven to provide intuition to human computer interface with richness of a tactile sense. Recent implementations in that field, however, have inherited a 2D Graphical User Interface (GUI) scheme and work as isolated systems. Consequently, the systems are limited in supporting intuitive interfaces and various applications. This paper presents an AR-based tangible interaction system for table-top interaction environments. A projector displays information or graphical images on the table through back-projection. Two cameras on and under the table track the tangible objects which are attached with ARToolKit markers and that are used as interaction tools. Furthermore, a Augmentation Display beside the table shows graphical objects combined with the real table. The system shares the object tracking information of the object as context through the vr-UCAM [8]. Moreover, we also present a scheme of tracking smoothing based on a tracking prediction using Kalman Filter. As an application, the ARTable is integrated with Responsive Multimedia System (RMS), a storytelling system [3][4]. The proposed system is applicable to a wide range of usages such as virtual reality and augmented reality applications.

1   Introduction

Vision-based Augmented Reality extends users' perception of the real world by augmenting virtual objects or information that users cannot directly recognize [1]. Tangible User Interfaces support intuitive interaction between human and computer by exploiting daily-life physical objects as a means of an interaction tool [2]. The purpose of this concept is to enhance enhancing the users' ability to complete their work by reducing the attention on the interface itself. Recently, researchers try to combine both concepts to design an intuitive interface which can provide rich tactile sensation simultaneously [6].

Ishii et al. proposed Tangible Bits which exploits daily-life objects as interaction tools to manipulate digital contents [2]. Rekimoto et al. developed an AR system which utilizes markers for tracking objects and a projector for displaying information on a table or wall. The system also supports information sharing among multiple devices [5]. In ARGroove, Poupyrev et al. controlled music by

manipulating real records. It augments virtual objects on the records to display the control state [6]. TMCS is a TUI system which exploits the user context to personalize services and controls multimedia contents with tangible objects [7]. Unfortunately, previous systems were not designed to support multiple applications as a common interface system. Therefore, the systems are isolated [2][5][6]. And, without proper feedbacks for the users' activity except of the final services, there is no way to find what is going wrong in case of an interface failure [7]. Direct inheritance of GUI-based interfaces is not intuitive to novice users [5].

The ARTable is motivated from the limitations of the previous AR and TUI systems. The ARTable aims to support intuitive interaction and to be adopted for various applications as a common interface. Our daily-life physical objects are used as tangible interaction tools for the ARTable. Moreover, it augments images on the table and graphical models on the objects. The augmentation simulates a virtual interface space which is assumed to be extracted on the real world. The ARTable uses vr-UCAM, which enables context sharing between various applications [8]. Therefore, the objects tracking information is shared with other applications in the form of context. Moreover, we present a scheme of tracking smoothing based on a tracking prediction using Kalman Filter.

The proposed system has following advantages. First, it simulates a virtual interaction space on the table-top environment and allows the user to manipulate tangible objects in that space. Thus, it supports intuitive interaction by combining the Augmented Reality and Tangible User Interface concept. Secondly, the system can support various applications by changing the augmentation and by sharing tracking information with other applications. Lastly, the tracking smoothing mechanism based on prediction generates denser tracking information.

This paper is organized as follows. In Section 2, we introduce the system configuration of the ARTable and the composition of each component including detailed specifications. In Section 3, we present the result of our experiments with the usability test of the ARTable applied as a navigational interface for RMS [3][4]. Finally, we give a conclusion and remaining works, in Section 4.

2   ARTable

Fig. 1 illustrates software components of the ARTable which consists of 4 key components; Calibration, Tracking, Table Display, and Augmentation Display. The Calibration is only for off-line step for a system setup and calculates transformation between two cameras on and under the table. The Tracking detects ARToolKit markers and calculates those poses into the simulated enhanced resolution based on a prediction. The Table Display and the Augmentation Display are in charge of providing information to help users' interactions. Each component, except for Calibration, extends vr-Sensor and vr-Service of vr-UCAM. Therefore they share user context and objects tracking information with other applications.

Fig. 2 shows the practical system configuration. Under the table there is a projector for a table Display and a camera for tracking objects. The camera
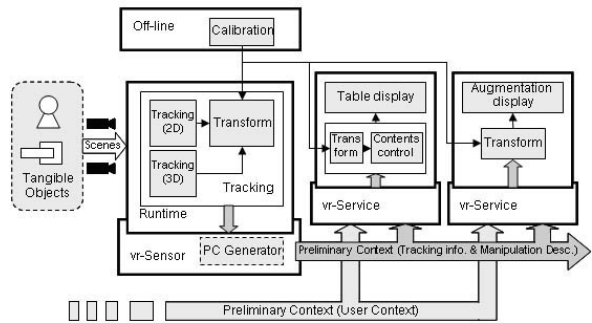
Fig. 1. Software components

under the table is only used to track object's 2D position and 1D orientation. Under the table, a mirror is equipped to extend the line of sight of the projector and the camera. Beside the table the Augmentation Display shows virtual graphical models augmented on the scene of the real table captured from the camera above the table. The upper camera is also used to track the 3D position and orientation of tangible objects.
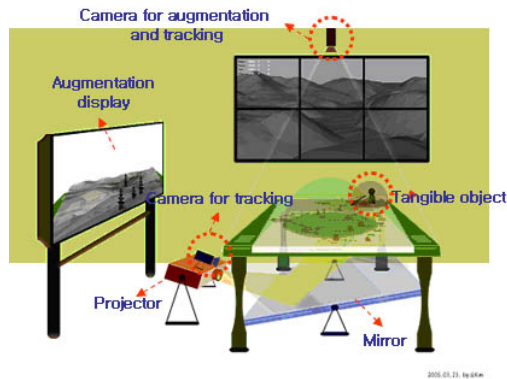

Fig. 2. System configuration

2.1   Calibration

The Calibration defines positional and rotational transformation matrices between two cameras and between each camera and the table. Since the scale of object tracking information is unknown, it is also required to define a normalization matrix to share the tracking information with other applications. All of these processes are conducted during off-line.

The calibration process is as follows. First of all, the calibration needs two double-sided markers that have same but horizontally-flipped patterns on each side. Then, we place one of the markers (marker C_1) to be fit on the left-upper

corner of the table surface. Another marker (marker C_2) is placed on the right-bottom corner of the table. In this way, the marker C_1 determines the table coordinate system. And, the distance between marker C_1 and C_2 is the table size used for normalization.

Once the markers are placed on the table and both of them are visible through two cameras, we can calculate transformations between cameras and the table as a function of ARToolKit [10]. Through a function of ARToolKit, the two markers are detected and the transformations between cameras and markers are available. is the transformation from the upper camera to marker C1. Likewise, is from the lower camera to marker C_2.

Then, T_{AO}, a transformation between the upper camera and a tangible object, is transformed into the table coordinate system by the equation 1.

$$T_{AC}^{-1}T_{AO} = T'_{AO} \tag{1}$$

In the same way, the transformation from the lower camera to a tangible object T_{BO} is transformed into the table coordinate system. The only difference of T'_{AO} and T'_{BO} is their opposite direction of the z-axis. Because the z-value of T'_{BO} is not used, they are considered to be same.

$$T_{BC}^{-1}T_{BO} = T'_{BO} \tag{2}$$

Consequently, T'_{AO} obtains the position and the orientation of a tangible object in the table coordinate system. Since the table size is calculated from the distance between marker C_1 and marker C_2, the pose values can be normalized.

2.2   Tracking

Tangible objects are attached with ARToolKit markers to be tracked from two cameras; one above the table, the other under the table. Fig. 3 illustrates examples of tangible objects. In Fig. 3(a), a marker is only attached on the bottom of a tangible object. In this case, only the lower camera can track the object, which limits the tracking information to be 2D position and 1D rotation. Nevertheless, the marker is not shown to the user and tracking does not fail even when the user occludes the object by hands. Therefore, it is possible to freely design the appearance of tangible objects. In case of 3(b) which attaches two markers on both sides of the object, the tracking is performed in a 3D space. In that case, however, the marker is revealed to the user and it restricts the objects to be partially planar. Tracking information, the pose, of the object is filtered to have denser resolution before it is shared with other applications using vr-UCAM. Principally, the resolution of tracking information is limited depending on the resolution of the image taken from a camera. Contrarily, the requirement of applications using the tracking information is not known beforehand. Therefore, when the applications require dense resolution of tracking information, insufficiently dense continuous poses might be appeared as errors obstructing proper services.

(a)   (b)

Fig. 3. Examples of tangible objects - the top and bottom sides (a) an object for 2D tracking having free form (b) an object for 3D tracking

Our scheme for increasing the resolution of tracking information is based on an interpolation between current tracking information (at time t) and the predicted one (at time $t+1$) of the following frame. Before a prediction, in order to remove a jittering, the extracted position is thresholded by a fixed value. Thus, if a new position of the marker is not different more than a threshold T from the previous, then it is considered as a jittering. Next step is to predict marker's movement using Kalman Filter [9]. By assuming a uniform acceleration of the markers' movement, the state of Kalman Filter is defined as equation 3.

$$x_t = \left[\tau_t \ \rho_t \ \frac{d\tau_t}{dt_t} \ \frac{d\rho_t}{dt_t} \ \frac{\frac{d\tau_t}{dt_t} - \frac{d\tau_{t-1}}{dt_{t-1}}}{dt_t} \ \frac{\frac{d\rho_t}{dt_t} - \frac{d\rho_{t-1}}{dt_{t-1}}}{dt_t} \right] \tag{3}$$

Where t is the time when the scene was captured by a camera. τ and ρ are the 2D position and 1D rotation of the object respectively. From the state, the measurement contains detected object pose as

$$z_t = \left[\tau_t \ \rho_t \right] \tag{4}$$

Thus, the measurement model which extracts the state of the measurement removes the velocity and acceleration terms.

$$\begin{bmatrix} \tau_t \\ \rho_t \end{bmatrix} = \left[I_{3\times 3} \ O_{3\times 3} \right] x_t \tag{5}$$

The process model to predict the next state of object pose is shown in equation 6. In that equation, the current velocity influences the next pose, and the current acceleration influences the next velocity, respectively.

$$\hat{x}_{t+dt}^- = \begin{bmatrix} I_{3\times 3} & I_{3\times 3} & O_{3\times 3} \\ O_{3\times 3} & I_{3\times 3} & I_{3\times 3} \\ O_{3\times 3} & O_{3\times 3} & I_{3\times 3} \end{bmatrix} \tag{6}$$

On each frame, the marker is detected and next position and rotation of the marker is predicted. Until the camera captures the next frame, the detected and predicted values are used to generate N tracking information between them. The two values are linearly interpolated as Equation 7. As an effect, the tracking information is increased N times than its resolution.

$$\begin{aligned} \tau_{t+\frac{n}{N}dt} &= \tau_t + (\tau_{t+dt} - \tau_t)\frac{n}{N}dt \\ \rho_{t+\frac{n}{N}dt} &= \rho_t + (\rho_{t+dt} - \rho_t)\frac{n}{N}dt \end{aligned} \tag{7}$$

2.3   Table Display and Augmentation Display

The Table Display presents the interface range in which the user can manipulate object and display the state of interaction space. As an example, when the ARTable is used for virtual world navigation, a 2D map of the virtual world is displayed on the table surface through the projector. It helps the user to be aware of the user's viewpoint in the whole world. And the display can also show dynamic visual effects according to user's object manipulation.
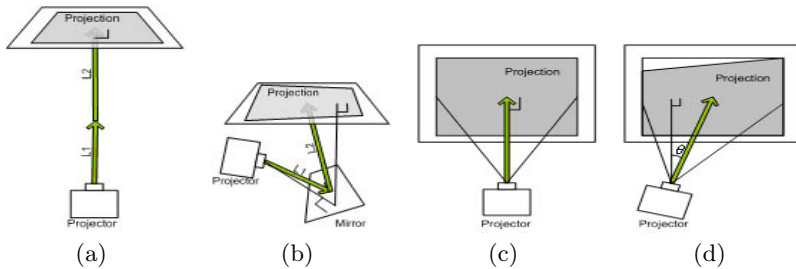
Fig. 4. Projection setup and keystone effects: (a) requires long projection distance (b) by using a mirror the distance is decreased but keystone effects occur (c) perpendicular setup doesn't cause keystone effect (d) keystone effect occurs due to misalignment between the projector and the projection surface
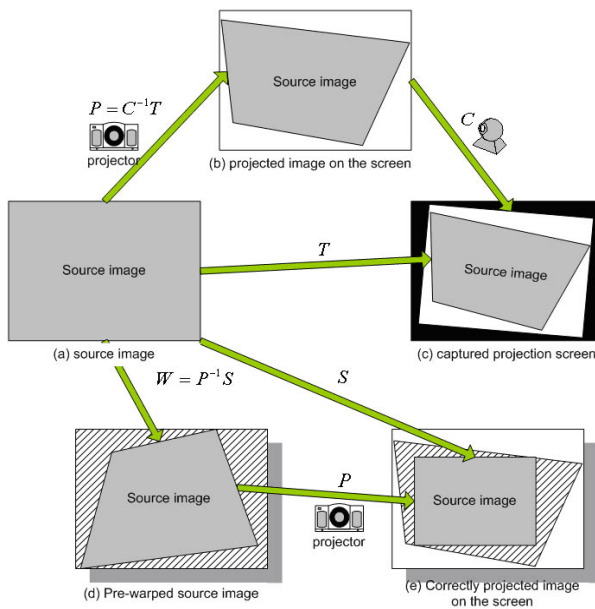

Fig. 5. Homography for keystone correction

The use of a mirror enables the system to decrease the height of the table while maintaining the scale of projection (Fig. 4(a)). However it causes keystone effects due to geometrical misalignments between the projection surface and the projector as shown in Fig. 4(b)(c)(d). In order not to cause keystone effects, the projector should be perpendicular to the projection plane. In our system setup, the misalignments occur twice in between projector-mirror and mirror-surface. In order to make the projection rectangular, keystone correction is applied.

The keystone correction directly follows the homography relations described in [11]. As shown in the Fig. 5, by capturing the projection screen (the table surface in our system), on which a full-screen source image is displayed, we could calculate P, a transformation influenced by keystone effect. In the sense that P should contain only translational and scaling factor if there is no effect of keystone, the pre-warp transformation W is calculated. Fig. 5(d) is before they are projected on the table surface. That results in correctly projected images without keystone.

The table surface is semi-transparent so that the lower camera can see through the surface and the projector can display information on the surface simultaneously. Therefore, it is possible to detect markers attached on the bottom of tangible objects. Fig. 6(a)(b) shows the table surface with a projector screen and a marker captured from the lower camera.

The Augmentation Display overlays the whole interaction space which is assumed to exist virtually on the real table. The upper camera captures the table surface with tangible objects and user's hands. Then, graphical objects are overlaid on the scene to show what happens in the interaction space. The
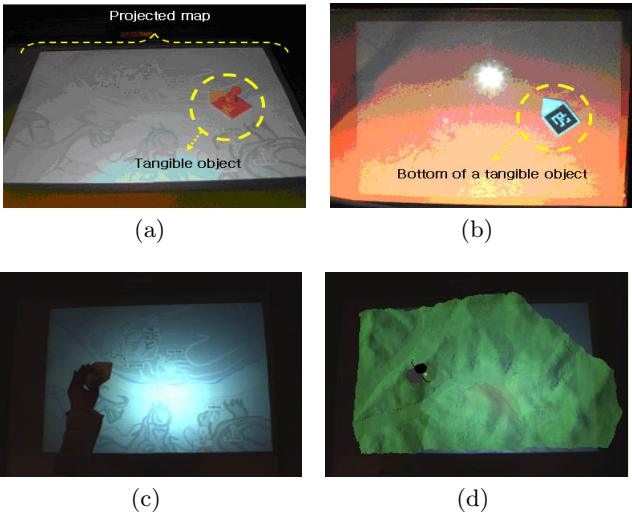
Fig. 6. Table Display and Augmentation Display (a) table surface on the table (b) the bottom the a tangible object seen through the lower (c) a scene of the upper camera (d) augmented graphical objects on the original scene; a terrain and a character camera

augmentation enables users to notice the objects' role in interaction. For instance, when we use the ARTable for a virtual world navigation application, the terrain of the virtual environment is augmented on the table. Fig. 6(c) shows the original scene of the upper camera and 6(d) shows graphical objects augmented on it.

3   Experiments

We have conducted an experiment to verify the effectiveness of the ARTable by integrating with RMS, a storytelling system developed in our lab. In the integration, the ARTable was simplified to leave out the Augmentation Display. The Tracking module which extends the vr-Sensor generates a preliminary context with an object type, object tracking information, object manipulation information. Thus, the context is shared with other applications through vr-UCAM. Fig. 7 illustrates the experimental setup; the ARTable and large back-projection screens for a virtual world.
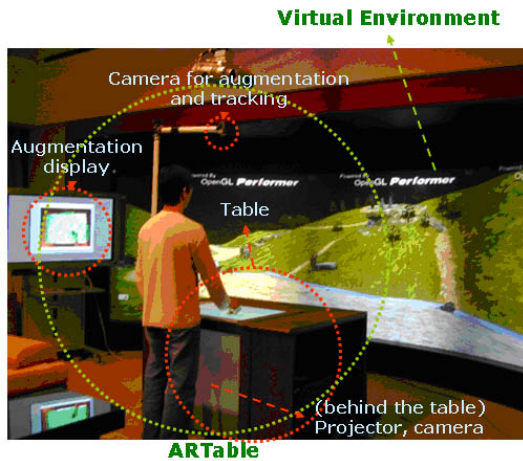

Fig. 7. Integrated system

We carried out a comparison between the ARTable and a tracker-embedded joystick, which are used for navigation in the virtual world. Ten people attended the experiment, and they are only the ones who have seen virtual environment applications but have not been tried it before. Therefore, they have no experience in both of interfaces but frequently used a keyboard and a mouse. In order to evaluate the intuition of our system, we measured the time took for the participants to learn the usage of the ARTable and a joystick. Then, the participants are ordered to move to a specific position of virtual world, and it is repeated 5 times. At last, we asked the participants about overall satisfaction level. Table 1 shows the result of the usability test.

Table 1. Comparison of ARTable and Joystick used for VE navigation - mean (standard deviation)

	ARTable	Joystick
Learning	32.4sec (8.8sec)	84.7sec (24.7sec)
Movement	10.4sec (7.1sec)	35sec (18.6sec)
Satisfaction	76% (19.1%)	54% (11.4%)

Experimental results reveal that all of the participants could use the ARTable without a prior knowledge and explanations. In case of a joystick, in contrary, they seem to have difficulties in learning the usage of it. After understanding how to handle them, the participants are ordered to move to a specific position in the virtual world. By using the ARTable, the participants could easily find the destination from the map projected on Table Display. Considering a joystick, even though they knew how to control it, they could not move freely as they wanted. Furthermore, the map on the table surface acted as an indicator which prevents the users from being lost. However, when they were using a joystick, they were lost in 12 trials of 50 times experiments, which caused the time to exceed over 120 seconds. Hence, we excluded the cases which exceed 120 seconds from the statistics. On the other hand, the ARTable has a defect, which couldn't recognize slight movements of objects. Some participants also pointed out that the current version support only 2D movement.

4   Conclusion

This paper proposed an AR-based tangible interaction system, named ARTable, for a table-top interaction environment. The ARTable provides intuitive interfaces by using daily-life physical objects as interaction tools and augmenting information for assist-ing the user's interaction. Furthermore, it can be used for interfaces of various applica-tions by adopting proper augmentation and by sharing the object tracking information with environmental applications through vr-UCAM. In our experiments, we developed a navigational interface on the ARTable for a storytelling system. From the usability test, we verified that users easily perceive how to use the proposed system and can enjoy this system through an easy interface. Further research issues include an investi-gation of a system configuration for a more convenient interface and an accurate ob-ject tracking technique with a stable and robust result. Designing tangible objects for enhancing intuition is one of additional issues. In addition, we should develop various applications to make full use of the proposed system.

Acknowledgements

References

1. R. Azuma, A Survey of Augmented Reality, Presence: Teleoperators and Virtual Environments 6, 4, pp.355 385, 1997
2. H. Ishii, B. Ullmer, ”Tangible Bits: Towards Seamless Interfaces between People, Bits and Atoms”, in Proceedings of ACM CHI ’97, pp. 234-241, Mar. 1997
3. Youngoh Lee, Sejin Oh, Woontack Woo, ”A Context-based Storytelling with Responsive Multimedia System (RMS),” LNCS(ICVS), 3805, pp. 12-21, 2005
4. Youngho Lee, Dahee Kim, Youngil Lim, Kyuhyung Kim, Haesun Kim, and Woontack Woo, ”Dream of Mee-luck: Aspiration for a New Dawn,” LNCS (ICVS),3805, pp. 282-285, 2005
5. J. Rekimoto and M. Saitoh, ”Augmented Surfaces: A Spatially Continuous Work Space for Hybrid Computing Environments”, In Proceedings of CHI 1999, ACM Press, pp. 378-385, 1999
6. I. Poupyrev, Augmented Groove : Collaborative Jamming in Augmented Reality, ACM SIGGRAPH 2000 Conference Abstracts and Applications, pp. 77, 2000
7. Sj.Oh and W.Woo, Manipulating Multimedia Contents with Tangible Media Control System, International Conference on Entertainment Computin, Vol.3166, pp. 57-67, 2004
8. Y.Lee, Sj.Oh, Y.Suh and W.Woo, Developing vr-UCAM for Interaction on Distributed Virtual Environments, KHCI2005, pp. 507-512, 2005
9. G. Welch and G. Bishop, An introduction to the kalman filter. CS TR 95-041, UNC, 1995
10. ARToolKit (http://www.hitl.washington.edu/research/shared_space/download)
11. Li, B., Sezan, I., ”Automatic keystone correction for smart projectors with embedded camera”, ICIP ’04. 2004 International Conference on Image Processing 2004 Volume 4, pp. 2829-2832, Oct. 2004