ICIC Express Letters
Part B: Applications
Volume **11**, Number **4**, April **2020**

ICIC International ©2020 ISSN 2185-2766
pp. 355–362

TANGIBLE AR INTERACTION USING FINGERTIP INPUT BASED ON THE TRACKING OF COLORED DOTS

Sangjin Park[1], Dongwoo Seo[1], Hokyun Jung[2] and Hyungjun Park[2,*]

[1]M&S Center
Korea Institute of Science and Technology Information (KISTI)
245 Daehak-ro, Yuseong-gu, Daejeon 34141, Korea
{ sjpark; seodongwoo }@kisti.re.kr

[2]Department of Industrial Engineering
Chosun University
309 Pilmun-daero, Dong-gu, Gwangju 61452, Korea
kiesleaf@gmail.com; *Corresponding author: hzpark@chosun.ac.kr

Received October 2019; accepted January 2020

Abstract. *In this paper, we propose a method for creating fingertip input which is based on the tracking of colored dots in an augmented reality (AR) environment, and present its proper use for tangible AR interaction which requires only simple objects without any hardwired connections. In the proposed method, two dot stickers with good contrast to skin color are attached around to the user's index fingertip, and tracked in or near the hand regions of each image frame of the AR environment.*

Keywords: Augmented reality, Fingertip input, Tangible interaction, Object tracking, Ellipse estimation, Colored dots

1. **Introduction.** Augmented reality (AR) is a technology that overlays computer-generated information onto a real scene and thereby provides users with information-enhanced environments which seamlessly connect real and virtual worlds [1]. Although it has been studied for more than two decades, there are research issues that need further investigation including tangible and natural interaction techniques for various AR applications.

Tangible AR interaction, which is obtained by combining a tangible user interface (TUI) with an AR system, aims at removing the gap between the interaction with a natural environment and the interaction with a computer system and thereby making user interaction tangible and intuitive. In tangible AR, each virtual object is registered to a physical object and the user interacts with virtual objects by manipulating the corresponding physical objects [2]. Various tracking techniques can be used for object registration and tracking, including senor-based, marker-based, and markerless tracking [3,15,16]. Markerless tracking does not need any artificial markers to calculate the position and orientation of virtual objects in the real environment, but it is not so good in tracking accuracy and computational efficiency as fiducial marker tracking. Marker-based tracking is simpler to implement, and better in speed and accuracy than the others.

In this paper, we propose a method for creating fingertip input which is based on the tracking of the colored dots placed on the user's index finger in an AR environment. We also present that the fingertip input method can be used as a selection tool for tangible AR interaction in the virtual prototyping of digital handheld products. As the proposed method does not require the use of a pointing tool with markers for interaction, it is much simpler than marker-based input methods of placing AR markers around a pen-type object or the user's index finger. We have found that the tangible AR interaction

DOI: 10.24507/icicelb.11.04.355

based on the fingertip input method is useful to improve natural interaction and immersive visualization of virtual products in the AR environment.

2. **Related Work.** As the AR technology becomes more popular, there are increasing studies that physical objects are used as tangible objects for user interaction in AR environments. There has recently been a proliferation of works on tangible AR interaction in which physical objects are used as tangible objects to make user interaction tangible and intuitive in AR environments. In this section, we briefly describe only some notable ones related with our work.

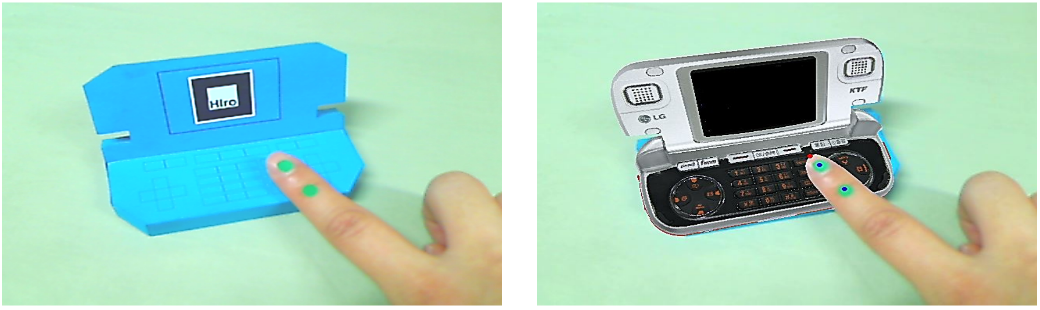
Verlinden et al. applied the concept of augmented prototyping which projects digital images on physical objects to perform the design review of handheld products like a voice recorder [4]. Lee and Park proposed augmented foam in which AR techniques are combined with physical blue foams [6]. The concept of integrating hardware and software in AR environments has been also studied [5,6]. It basically augments a virtual display onto the soft mockup of a product by incorporating various devices including micro switches or buttons, magnetic sensors, data gloves in order to support direct and tangible interfaces which are hardwired or wireless. Radio frequency identification (RFID) technology has been applied for user identification and interaction [17]. Klompmaker et al. combined R-FID technology with depth-sensing cameras to enable personalized authenticated tangible interactions on a tabletop [11]. Kanai et al. suggested the use of a glove-type wearable RFID R/W device and a physical mockup with small RFID tags attached to its faces for testing the usability of information appliances in an AR environment [7]. Aoyama and Kimishima proposed an AR-based system for evaluating designability and operability of digital appliances in which a video see-through HMD (Head Mounted Display), a data glove, and a physical object with a marker are combined with magnetic sensors to provide the pseudo feeling of touching product [8]. Takahashi and Kawashima proposed a system for viewing a digital product while grasping a tangible mockup and creating user inputs by fingertip touch [10]. To support fingertip touch and tracking, they combined a touch sensor attached to a fingertip with a small maker attached to the fingernail.

Physical mockups with hardwired connection can provide direct and accurate interfaces, but significant efforts are usually required to implement and build them. Moreover, it is not easy to make them available to many people who are located at different places. Park et al. [9,12-14,18] have studied a pioneering approach to AR-based user interaction which uses simple physical objects to provide tangible interaction without any hardwired connections in a simple and cheap AR environment. In their approach, the user creates input events by touching specified regions of the product-type object with the pointer-type object, and the virtual product reacts to the events by rendering its visual and auditory contents on the output device.

3. **Proposed Approach.** In this section, we describe a method for fingertip input based on the tracking of colored dots in an AR environment, and present a tangible AR interaction based on the method.

As shown in Figure 1, in order to facilitate the tangible interaction between the user and the product in the AR environment, we use a product-type tangible object and two dot stickers attached around to the user's index fingertip. The product-type object is used to acquire the position and orientation of the product, and the two dot stickers are used to recognize the position of a fingertip and create the fingertip input.

The product-type object and the dot stickers are colored in blue and green respectively in order to make them create good contrast not only to each other, but also to skin color. The product-type object representing a digital handheld product can be either a rapid prototyping (RP) mockup [4,7,9,12] or a paper model [13]. We can apply three dimensional (3D) printing or paper crafting to building the tangible object which is almost



(a) Product-type object and two dot stickers (b) Augmented view of the tangible objects

Figure 1. (color online) Tangible objects used in the proposed method

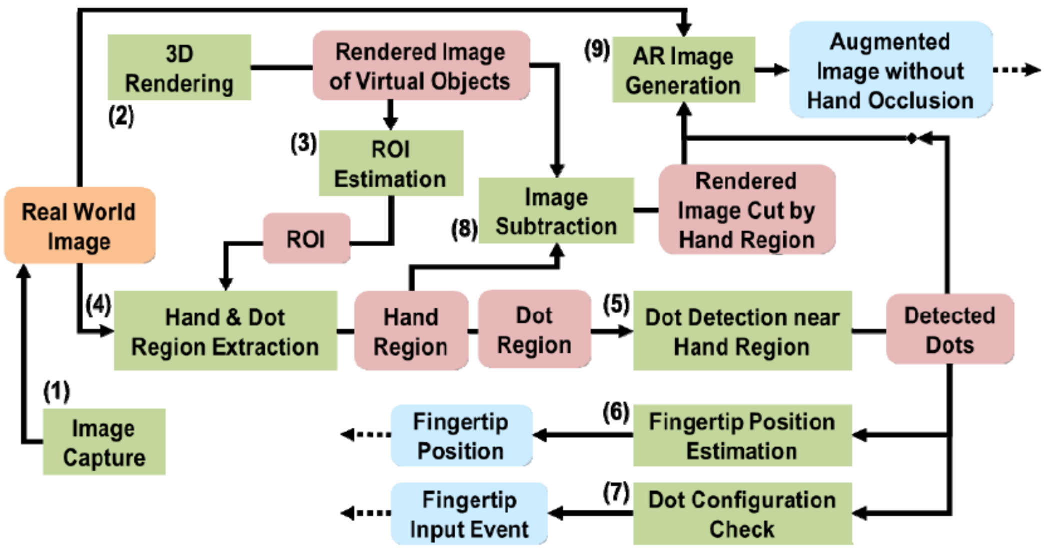


Figure 2. The procedure of the fingertip input method

the same in shape as its virtual object. Then, we can make the real image of the tangible object properly overlaid by the rendered region of the virtual object. In this paper, paper crafting is applied for making the paper model of a digital product. Shown in Figure 1 is the paper model of a game phone.

Figure 2 depicts the overall procedure of the proposed fingertip input method. The method is primarily based on the extraction of hand regions from each captured image of the real world and the detection of green dots in the hand regions. Additionally, it is incorporated into the hand occlusion solver which was proposed by Park and Moon [12,18]. The main steps of the method are summarized as follows.

(1) Capture a real world image **W** with a calibrated camera and acquire AR maker information from the image.

(2) By using the AR marker information, generate the rendered image **V** of 3D virtual objects with black background color.

(3) Estimate the region of interest (ROI) in which the tracking of colored dots and the occlusion of hand regions are considered.

(4) Extract hand and dot regions (**H** and **D**, respectively) from the real world image **W** by using color filtering.

(5) Detect dots from the dot regions (**D**) in or near the hand regions (**H**) by taking into account of contour area and distance between contours.

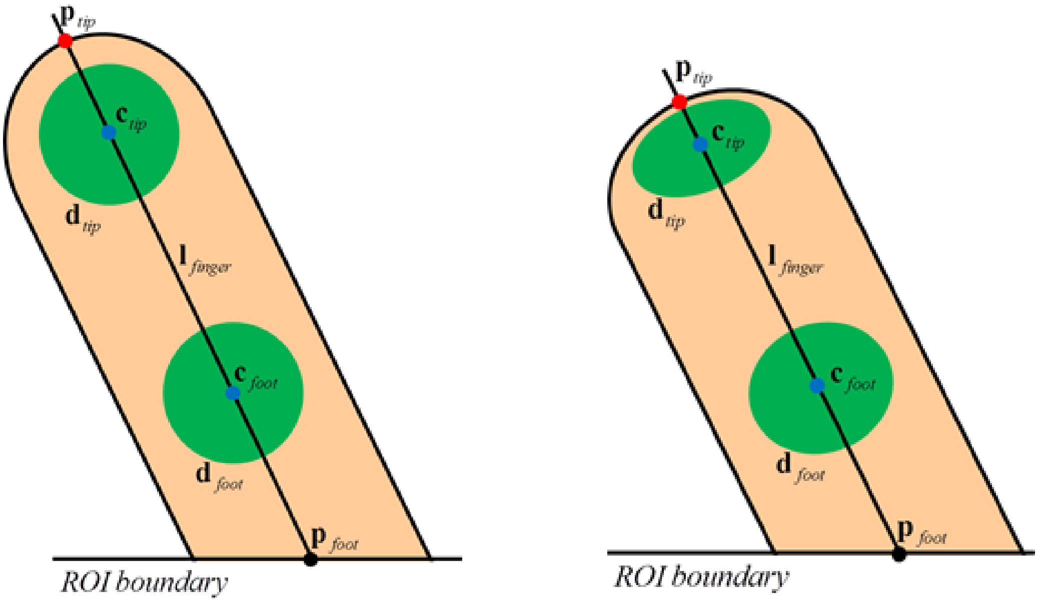
(6) Estimate the position of the fingertip based on the position of the detected dots.

(7) Based on the configuration of the detected dots, control the state of fingertip input.

(8) Refine the rendered image **V** by cutting out the hand and dot regions in the ROI out of it.

(9) Generate a final AR image **F** without hand occlusion by superimposing the refined rendered image **V** (free of hand occlusion) onto the real world image **W**. If necessary, display the centers of the detected dots in the AR image.

4. **Fingertip Input Mechanism.** In the AR environment considered in this work, there are at most two dots detected in the ROI. For each detected dot, we use its minimum-area rectangle (with its center $\mathbf{c}_{MAR}$, length l_{MAR}, and height h_{MAR}) to obtain its estimated ellipse whose center is $\mathbf{c}_{MAR}$, and the major and minor radii are l_{MAR} and h_{MAR}, respectively. Let $r = h_{MAR}/l_{MAR}$ be the shape ratio of the dot. If $r = 1$, its boundary becomes a circle. In this work, the dot is called fat if the ratio is greater than a ratio tolerance. Otherwise, it is called thin. In this work, the ratio tolerance is set to 0.7. The centers of the detected dots are used in computing the fingertip position, and the radii of their estimated ellipses are used in determining the state of the fingertip input. If there exist two dots which are detected in or near hand regions, we consider that fingertip input is active. In Step (6), we estimate the fingertip position by using the positions of the two detected dots. See Figure 3.



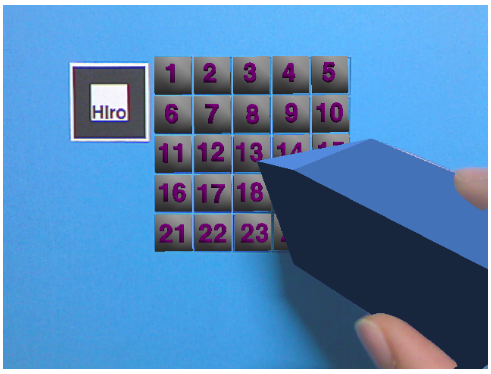
(a) Waiting state (b) Selection state

FIGURE 3. Computation of the fingertip position

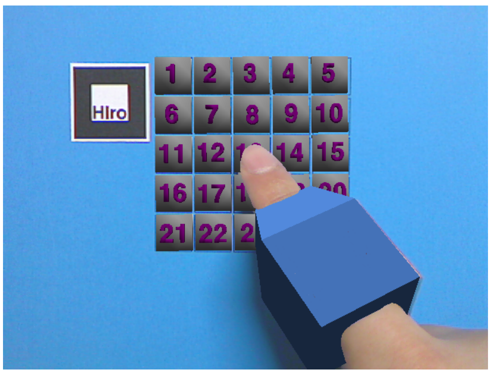
After generating a line $\mathbf{l}_{finger}$ passing the centers of the two dots, we compute the intersections between the line and the outmost contour of the hand region which is adjacent to the two dots. As the hand region touches the boundary of the ROI, one of the intersections is located on the ROI boundary, which becomes the foot point $\mathbf{p}_{foot}$. The fingertip position $\mathbf{p}_{tip}$ is defined as the other intersection which is not located on the ROI boundary. If two intersections are located on the ROI boundary, the fingertip position is not determined, which means that the fingertip input is not supported temporarily. Among the two dot centers, the dot center farther from the foot point $\mathbf{p}_{foot}$ becomes the one closer to the fingertip position. Let $\mathbf{c}_{foot}$ be the center of the dot $\mathbf{d}_{foot}$ closer to the foot point $\mathbf{p}_{foot}$ and let $\mathbf{c}_{tip}$ be the center of the dot $\mathbf{d}_{tip}$ farther from the foot point $\mathbf{p}_{foot}$. To define the state of the fingertip input, we use an indicator s_{tip} which indicates two kinds (fat or thin) of the shape of the dot $\mathbf{d}_{tip}$ closer to the fingertip. In Step (7), the state of the fingertip input is controlled as follows: If two dots are detected with $s_{tip} = fat$ regardless of the previous state, the current state becomes the *waiting* state as shown in Figure 3(a). If the previous state is the waiting state and the fat dot closer to the fingertip becomes *thin* as shown in Figure 3(b) (i.e., $s_{tip} = thin$), the state becomes the *selection* state, which triggers a fingertip input event. In other cases, the current state becomes the idle state.

5. **Experiments.** Let DOT denote the proposed approach using two dot stickers for tangible interaction. Let PEN denote the approach using a pointer-type object [9], and RING the approach using a ring-type object [13]. Let PATCH denote the approach using a small patch-type object for tangible interaction [14]. To compare the accuracy of the DOT approach with the three previous approaches, we conducted button selection experiments. In the experiments, a subject group consisting of 20 university students was voluntarily participated using a setup similar. In the setup, instead of a product-type object, a thin plate (size: 150×130 mm^2) is positioned in a fixed location. Square buttons with an AR maker (size: 28×28 mm^2) are put on the plate. The subjects consist of 16 males and 4 females and their ages range from 21 to 26 years (mean = 24). They all have basic knowledge of 3D geometric modeling and computer aided design (CAD), and most of them are familiar with the concept of AR.

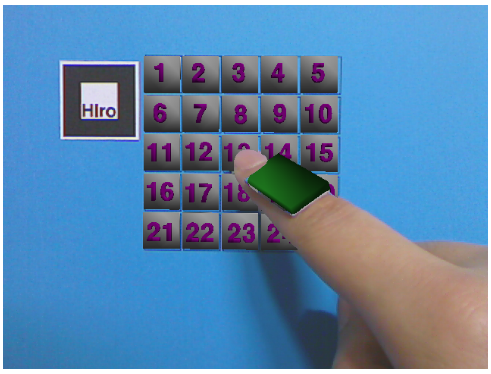
As shown in Figure 4, each subject was informed how to do button selection tasks with the four approaches, and he or she was asked to complete a set of button selection tasks. The side length of the buttons decreases from 20, 16, 12, 8, 6, 5, 4 mm. Given a sequence of 4 random numbers in each selection task, the user has to touch the center regions of the given numbers with his index finger holding the ring-type object. When using the ring-type object, every subject is allowed to see the small red sphere whose center is at



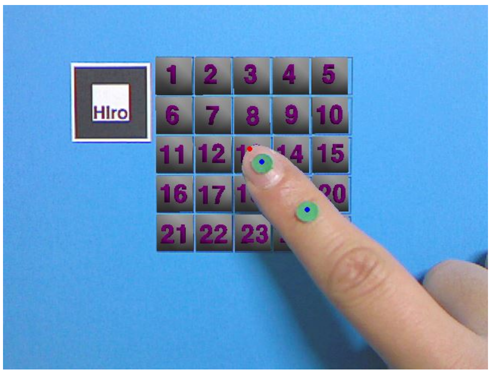
(a) Using the pen-type object



(b) Using the ring-type object



(c) Using the patch-type object

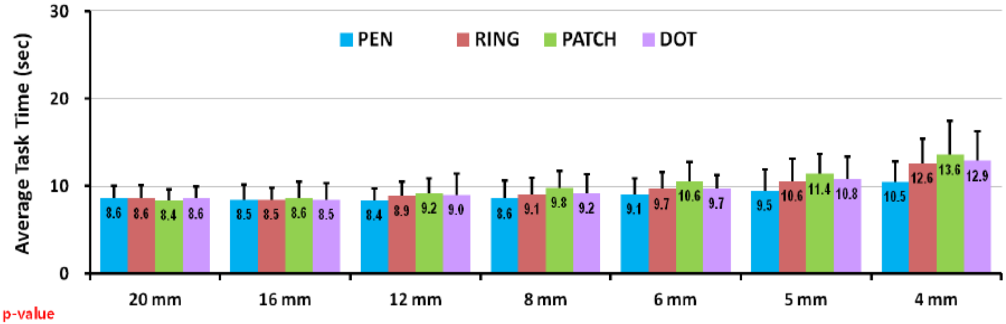


(d) Using the dot stickers

FIGURE 4. Button selection test (12 mm in size)

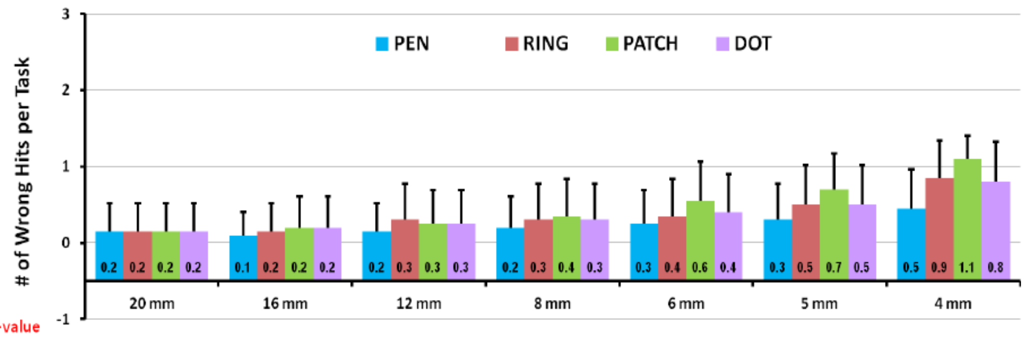
the reference location in the augmented view. Different sounds are given to differentiate between right and wrong selections of each number.

Figure 5 shows the accuracy assessment of the three approaches in the time required and the number of wrong selection per each task. In the figure, average values with standard deviations are plotted and the p-values for paired t-test [19] for sample means are included. Asterisks are put to indicate p-values less than a significance level of 0.05, indicating that there are statistically significant differences between two approaches at a significance level of 0.05.



	20 mm	16 mm	12 mm	8 mm	6 mm	5 mm	4 mm
(PEN, RING)	0.980	1.000	0.253	0.492	0.320	0.178	0.011*
(PEN, PATCH)	0.601	0.830	0.095	0.056	0.017*	0.043*	0.011*
(PEN, DOT)	0.967	0.986	0.357	0.440	0.329	0.091	0.013*
(RING, PATCH)	0.628	0.832	0.579	0.318	0.217	0.237	0.402
(RING, DOT)	0.981	0.987	0.863	0.880	0.993	0.793	0.760
(PATCH, DOT)	0.670	0.813	0.749	0.364	0.079	0.441	0.570

(a) Time required for each button selection task



	20 mm	16 mm	12 mm	8 mm	6 mm	5 mm	4 mm
(PEN, RING)	0.980	1.000	0.253	0.492	0.320	0.178	0.011*
(PEN, PATCH)	0.601	0.830	0.095	0.056	0.017*	0.043*	0.011*
(PEN, DOT)	0.967	0.986	0.357	0.440	0.329	0.091	0.013*
(RING, PATCH)	0.628	0.832	0.579	0.318	0.217	0.237	0.402
(RING, DOT)	0.981	0.987	0.863	0.880	0.993	0.793	0.760
(PATCH, DOT)	0.670	0.813	0.749	0.364	0.079	0.441	0.570

(b) Number of mistakes for each button selection task

FIGURE 5. Accuracy of tangible AR interaction

The gap between the fingertip and its computed location is usually bigger than the gap between the pen tip and its computed location, implying that the PEN approach is more accurate than the three (RING, PATCH, DOT) approaches. We found a slight tendency that the overall ranking of the accuracy is $PEN > RING \approx DOT > PATCH$ where $>$ means "better than" and $\approx$ means "comparable to". There is a tendency that the PATCH approach is less accurate in button selection tasks when the button size is not greater than 6 mm. It may result from the small size of the marker placed in the sticker-type object, which tends to deteriorate tracking performance. However, there are no statistically significant differences between any pair of approaches both in the average task time and in the number of wrong selections for button sizes greater than 6 mm.

Although the size of buttons of digital handheld products is very diverse, it is mostly not less than an adult's little fingernail whose size is usually bigger than 5×5 mm^2. Small keypad buttons of cellular or smart phones are not less than 5×5 mm^2, and users often make mistakes in typing tasks using such small keypad buttons. Based on this observation and by analyzing the experimental results, we found that the time and the error frequency are tolerable if the button size is not less than 5 mm. From the experimental results, the proposed tangible interaction approach can be applied to a wide variety of digital handheld products.

6. **Conclusions.** In this paper, we have described the fingertip input method based on the tacking of the colored dots attached on the user's index finger, and its proper use for tangible interaction in a simple and cheap AR environment setup. As the fingertip input method does not require the use of a marker-based pointing tool for interaction, it is simpler and more efficient than marker-based input methods of placing AR markers around a pen-type object or the user's index finger. Moreover, the tangible AR interaction based on the fingertip input method is available at very low cost without hardwired connections and useful to improve natural interaction and immersive visualization of virtual products. From the experimental results with the feedbacks from a group of users, we found that the proposed user interaction is accurate enough to be applied to virtual design evaluation of digital handheld products. However, the selection state can sometimes occur even when the actual finger does not touch the button of the mockup during the interaction, and improvement is required.

We expect that the proposed approach can be used for many applications including AR-based user experience of virtual prototypes and AR-based contents for education and training in which user interfaces are required to be tangible and natural. As future research work, we plan to enhance the approach by applying user hand depth information and machine learning methods.

Acknowledgment. This research was supported by Korea Institute of Science and Technology Information (KISTI).

REFERENCES

[1] R. T. Azuma, A survey of augmented reality, *Presence: Teleoperators and Virtual Environments*, vol.6, no.4, pp.355-385, 1997.
[2] M. Billinghurt, H. Kato and I. Poupyrev, Collaboration with tangible augmented reality interfaces, *Proc. of International Conference on Human-Computer Interaction (HCI)*, pp.234-241, 2001.
[3] V. Teichrieb et al., A survey of online monocular markerless augmented reality, *International Journal of Modeling and Simulation for the Petroleum Industry*, vol.1, no.1, pp.1-7, 2007.
[4] J. Verlinden, W. Van den Esker, L. Wind and I. Horvath, Qualitative comparison of virtual and augmented prototyping of handheld products, *Proc. of International Design Conference (DESIGN)*, pp.533-538, 2004.
[5] T. J. Nam, Sketch-based rapid prototyping platform for hardware-software integrated interactive products, *Proc. of Conference on Human Factors in Computing Systems (CHI)*, pp.1689-1692, 2005.
[6] W. Lee and J. Park, Augmented foam: Touchable and graspable augmented reality for product design simulation, *Bulletin of Japanese Society for the Design Science*, pp.17-26, 2006.
[7] S. Kanai, S. Horiuchi, Y. Kikuta, A. Yokoyama and Y. Shiroma, An integrated environment for testing and assessing the usability of information appliances using digital and physical mock-ups, *Lecture Notes in Computer Science*, pp.478-487, 2007.
[8] H. Aoyama and Y. Kimishima, Mixed reality system for evaluating design ability and operability of information appliances, *International Journal of Interactive Design and Manufacturing*, vol.3, no.3, pp.157-164, 2009.
[9] H. Park, H. C. Moon and J. Y. Lee, Tangible augmented prototyping of digital handheld products, *Computers in Industry*, vol.60, no.2, pp.114-125, 2009.

[10] H. Takahashi and T. Kawashima, Touch-sensitive augmented reality system for development of handheld information appliances, *International Journal on Interactive Design and Manufacturing*, vol.4, no.1, pp.25-34, 2010.

[11] F. Klompmaker, H. Fischer and H. Jung, Authenticated tangible interaction using RFID and depth-sensing cameras – Supporting collaboration on interactive tabletops, *Proc. of the 5th International Conference on Advances in Computer-Human Interactions (ACHI)*, pp.141-144, 2012.

[12] H. Park and H. C. Moon, Design evaluation of information appliances using augmented reality-based tangible interaction, *Computers in Industry*, vol.64, no.7, pp.854-868, 2013.

[13] H. Park, S. J. Park and H. K. Jung, Note on tangible interaction using paper models for AR-based design evaluation, *Journal of Advanced Mechanical Design, Systems, and Manufacturing*, vol.7, no.5, pp.827-835, 2013.

[14] H. Park, H. K. Jung and S. J. Park, Tangible AR interaction based on fingertip touch using small-sized non-square markers, *Journal of Computational Design and Engineering*, vol.1, no.4, pp.289-297, 2014.

[15] K. Oka, Y. Sato and H. Koike, Real-time fingertip tracking and gesture recognition, *IEEE Computer Graphics and Applications*, vol.22, no.6, pp.64-70, 2002.

[16] D. Lee and S. G. Lee, Vision-based finger action recognition by angle detection and contour analysis, *ETRI Journal*, vol.33, no.2, pp.415-422, 2011.

[17] J. Bravo, R. Hervas, G. Chavira, S. W. Nava and V. Villarreal, From implicit to touching interaction: RFID and NFC approaches, *Proc. of Conference on Human System Interactions (HSI)*, pp.743-748, 2008.

[18] H. C. Moon and H. Park, Resolving hand region occlusion in tangible augmented reality environments, *Transactions of the Society of CAD/CAM Engineers*, vol.16, no.4, pp.277-284, 2011.

[19] M. Kutner, C. Nachtsheim, J. Neter and W. Li, *Applied Linear Statistical Models*, McGraw-Hill Irwin, Boston, 2004.