Journal of Korea Multimedia Society Vol. 21, No. 1, January 2018(pp. 69-76)
https://doi.org/10.9717/kmms.2018.21.1.069

Development of Infants Music Education Application Using Augmented Reality

Seunguk Yeon†, Sukyong Seo††

ABSTRACT

Augmented Reality (AR) technology has rapidly been applied to various application areas including e-learning and e-education. Focusing on the design and development of android tablet application, this study targeted to develop infant music education using AR technology. We used a tablet instead of personal computer because it is more easily accessible and more convenient. Our system allows infant users to play with teaching aids like blocks or puzzles to mimic musical play like game. The user sets the puzzle piece on the playground in front of the tablet and presses the play button. Then, the system extracts a region of interest among the images acquired by internal camera and separates the foreground image from the background image. The block recognition software analyzes, recognizes and shows the result using AR technology. In order to have reasonably working recognition ratio, we did experiments with more than 5,000 frames of actual playing scenarios. We found that the recognition rate can be secured up to 95%, when the threshold values are selected well using various condition parameters.

Key words: Augmented Reality, E-education, Android App, Object Recognition.

1. INTRODUCTION

The smartphone and tablet PC craze, which has been sweeping the world for several years, is changing people's consumption pattern and life-style. Nowadays, IT devices have a close relationship with our daily lives. In the meantime, anyone can easily use new technologies that appeared in the movies. One of the typical new technologies is 'augmented reality (AR)', which shows virtually related information about real objects. AR is now popularized widely by the general public. For example, map retrieval via the Internet, location retrieval, etc. are also included in AR in a broad sense. Because computers are difficult to use while on the move, AR technologies for portable devices such as smartphones and tablet PCs have begun to be used more. There are several requirements to apply AR with GPS technology [1,2]. They are GPS devices that transmit and receive geographic information, gravity sensors, location information systems and AR application that receives detailed information and displays it on a realistic background, and finally an IT device like smart phone, tablet PC that outputs it to the display.

Such AR technology is being applied recently in the field of infant education [3-5]. Instead of creating a new teaching aid for AR, we use an image recognition algorithm to recycle the existing ones. Young children can learn best through direct experience, activity-based, and play-oriented activities, not through formal teaching. In this paper, the target infants are children from aged of around 5 years. Infants at this age often lose concentration

※ Corresponding Author : Sukyong Seo, Address: (13135) 553 Sansung-daero, Sungnam-si, Gyonggi-do, Korea, TEL : +82-31-740-7128, FAX : +82-31-740-7359, E-mail : sue.seo@gmail.com
Receipt date : Dec. 8, 2017, Revision date : Jan. 4, 2018
Approval date : Jan. 5, 2018

† Dept. of Information Security, Seoul Women's University (E-mail : seunguk@swu.ac.kr)
†† Dept. of Nursing, Eulji University
※ This work was supported by a research grant from Seoul Women's University (2017).

and are likely to get bored when teaching aid's response is slow.

This paper focused on the design and development of android tablet application targeted for young child and infant education using AR technology. Our software recognizes physical block using AR technic without any physical connection between the block and the tablet. The recognition algorithm uses different threshold value that shows the best recognition rate per block.

The composition of this paper after chapter 1 is as follows. Chapter 2 summarizes existing related studies. Chapters 3 and 4 provide an overview, design, and implementation of the system. Chapter 5 describes the results of our experiment performed by real teaching aids, followed by conclusions in Chapter 6.

2. RELATED WORKS

2.1 Augmented Reality

Augmented Reality (AR) can be described as an application that unifies the real world and the virtual world. The AR overwrites the actual image with a virtual image with the camera built into the system. Fig. 1 shows AR example. The purpose of AR is to mount two-dimensional or three-dimensional objects on an image of the real world. This is done in two different ways: the first is to find the marker on the image and the second is to use a special camera.

On the contrary, it is also an application of AR to mount an object recognized in the real world image on a virtual screen. This paper can be regarded as an implementation of AR in this direction. AR applications also use GPS and web browser technology [1,2]. AR can be applied to various fields such as education, security, health, and social media. AR raises the quality of education [3-5].

Fig. 1. AR example - Driving information displayed on the windshield.

In relation to object recognition, there are four kinds of object recognition methods based on image recognition area. First, there is a method using a color histogram [6,17]. This method extracts the color information value of the object region from the input image, processes it, and recognizes the object by comparing the similarity with the color information value of the object to be searched. Second, there is a method using the directional histogram of the edge [7]. In this method the background image of the object to be recognized is removed and the binary image with 36 features through an edge directional histogram is created, and a similar object is searched by a method of clustering objects using Principal Component Analysis. Third, there are couple of methods using feature points of objects such as SIFT [8], SURF [9,18], ORB [10], and BRISK [11]. Although the feature point methods provide high performance, they are not suitable for real-time object recognition due to high computational complexity [8,9] and in case of an object with weak internal features, object recognition is difficult. This method also includes studies that speed up existing ORB algorithm with parallel processing through comparative data sorting and core-level classification [12]. Fourth, there is a method using histogram-based correlation, which is a method of recognizing an object by a relative positional comparison method between the pixels of the reference image and the pixels of the input image [13]. The above methods are not effective in terms of recognition rate and real-time response rate for the teaching aid objects covered by this paper.

2.2 E-Education

Several approaches exist, such as Scratch, which teaches children by manipulating objects and blocks using a graphical user interface on a personal computer [14]. In some environments, a personal computer or desktop computer is less accessible and less convenient than a tablet PC, so a simple tablet system is preferred. There has been research on systems that allow similar experiences to be programmed using operation cards on desktop computers as well as on tablet devices using the Android operating system [15]. There is a model developed by Unity using SDK such as Vuforia for developing AR application program [15,16]. However, the model is not suitable for this paper, so we implemented our system using OpenCV and Unity from scratch.

3. SYSTEM DESIGN

3.1 Hardware Requirements

The main hardware components of the system developed in this study are 1) tablet for calculation and graphical user interface, storage of game scenarios, teaching aid objects' feature database, and 2) mirror for internal camera and 3) tablet stand. This is illustrated in Fig. 2.

3.2 Software Requirements

The software system to be developed is an application composed of the following parts, a part for extracting a region of interest(ROI) in an image acquired by the tablet's internal camera, a part for separating the foreground image from the background image, a part for recognizing the object by analyzing it, a part of gaming based on recognition result, and a part that showing the result on the tablet screen as AR.

Fig. 2. Hardware picture.

3.3 System Design

From an application point of view, the system is basically an educational game program in which infants learn music while playing with teaching aids, which are blocks or puzzles. It shows the relationship between scales and numbers through animations that play in order of scale. They use a colored water cup. After listening to the song, the goal is to adjust the number of water cups in such a way that if the number is less than the number of colored water cups, water is added to the water cups. In addition, if it is larger than the number of colored water cups, water is removed from the water cups.

Blocks(puzzles) shown in Fig. 3 used to add or remove water from the cups. The total number of blocks is 11, and image recognition algorithm is used to recognize a block. The number and size of a block is adjusted for effective image recognition. In addition, the game scenario is designed to effectively recognize the image of the blocks.

Fig. 4 shows the overall system configuration. The user sets the puzzle piece on the playground in front of the tablet and presses the play button. At this time, the image is captured and delivered to the built-in camera through a removable mirror attached to the tablet, and the puzzle pieces are

Fig. 3. Pictures of sample blocks(puzzles).

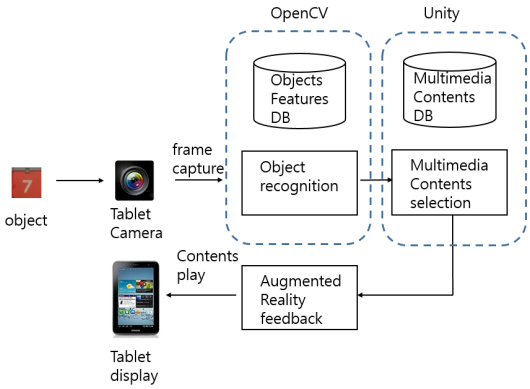

Fig. 4. Overall System diagram.

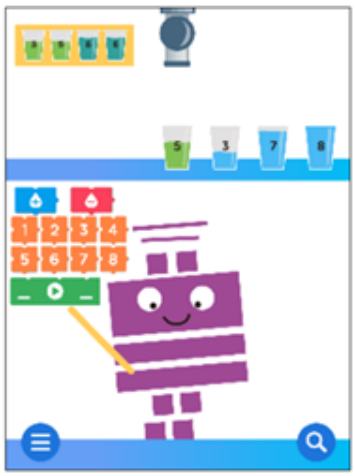

Fig. 5. Sample screen of playing.

recognized by analyzing the input image in real time, and the corresponding multimedia contents are played as an AR on the tablet.

Fig. 5 is an example of the actual multimedia content screen.

4. SYSTEM IMPLEMENTATION

In order to develop our system efficiently, object recognition part and user interaction part are divided and implemented independently each.

4.1 Object Recognition

Open CV, an open source image recognition library, was used to recognize blocks. The block recognition algorithm used in this paper is shown in Fig. 6.

The object recognition process is explained like below. First, the distortion of the region of interest is corrected by the preprocessing process, and the matching candidate is determined using the classification algorithm for the region of interest. After that, it is checked whether the play button is pressed and the code of the finally determined block is returned. Correction of the distortion of the ROI in the preprocessing process is performed in two steps: first step is to correct the geometric distortion of the image due to the slope of the tablet, and second step is to adjust the rotation of the region of interest. After the region of interest is extracted, object recognition is performed. At the beginning of the program, a feature database of blocks is constructed for the reference. The classifier obtains the size (horizontal and vertical length) and type information of the input image and then bitwise-XORs the object block with the reference block at a pixel level and classifies puzzles that are higher than a certain threshold value. When the play button is pressed, it is necessary to recognize the point that the play button is pressed because at this point the blocks are recognized as target blocks selected by the player. To do this, two white buttons are added to the left and right of the play button, and a change in the color of the button is detected by checking a threshold value. Fig. 7 shows this process.

4.2 User Interaction

The system was developed using Unity, a popular game development platform, because it was

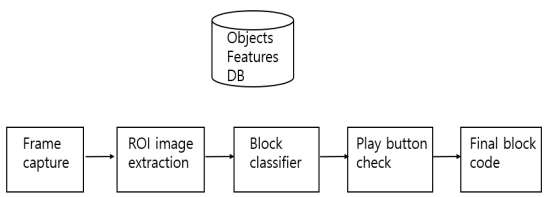

Fig. 6. Object Recognition Diagram.

Fig. 7. Detecting play button.

considered effective for infant to interact with the blocks in a game like way. It shows the relationship between scale and number through an animation that plays in order of scale. After the child watches this animation, the system displays the scale in the upper left corner of the screen, displaying the number and the amount of water in the colored cup. Below the faucet are the infant's water cups which are moving from right to left.

The goal of this game is to make the number of water cups the same as the number of colored water cups. Fig. 8 shows this water matching process. To do it, infants need to select how much water to add to the water cup or how less water to remove from it. They need to use one block to add water and another block to remove water and each number block like to match the water. The process of selecting a physical block and pressing the play button is done through the object recognition module. The process of displaying user interaction via multimedia content is done through Unity modules.

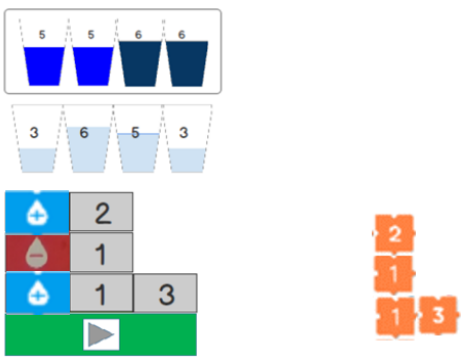

Fig. 8. Water matching process.

5. EXPERIMENTS AND RESULTS

5.1 Test Parameters

In order to evaluate the recognition rate of the recognition algorithms presented in Chapter 4, we performed an experiment investigating how the block recognition rate changes by adjusting parameters. The block recognition algorithm is divided into two parts: a part for recognizing each numbers and the other for recognizing the play button. For the first part, the recognition rate varies according to the recognized threshold value for each number (see Fig. 9).

For example, the number 1 graph indicates that if the threshold value is from 87 to 80 then the recognition rate is best. For the number 2, if the threshold value is from 89 to 82 then the recognition rate is best. On the other hand, the number 5 has best recognition rate when threshold value ranges between 95 and 91. The rest of the numbers except for 5 are recognized well with a value of 85. The number 5 is similar to the number 8 or 3, so it is not easy to recognize it. We found that there is no significant correlation with the change of illumination.

In recognizing a block of numbers, we performed a bitwise-XOR comparison with the reference block at a pixel level to obtain a puzzle with a threshold value higher than a certain threshold value. Because the threshold value shows the best recognition rate for each number, the algorithm is adjusted accordingly. For example, we set the threshold value as 85 for the numbers other than 5, and for the number 5, we increased the value and tried again.

When the play button is recognized, the bitwise-XOR is compared according to the change of the shadow, and an experiment is performed to find an appropriate threshold value to judge play when it is higher than a predetermined threshold value. However, in the experiment result, the recognition rate differs according to the brightness of the illu-

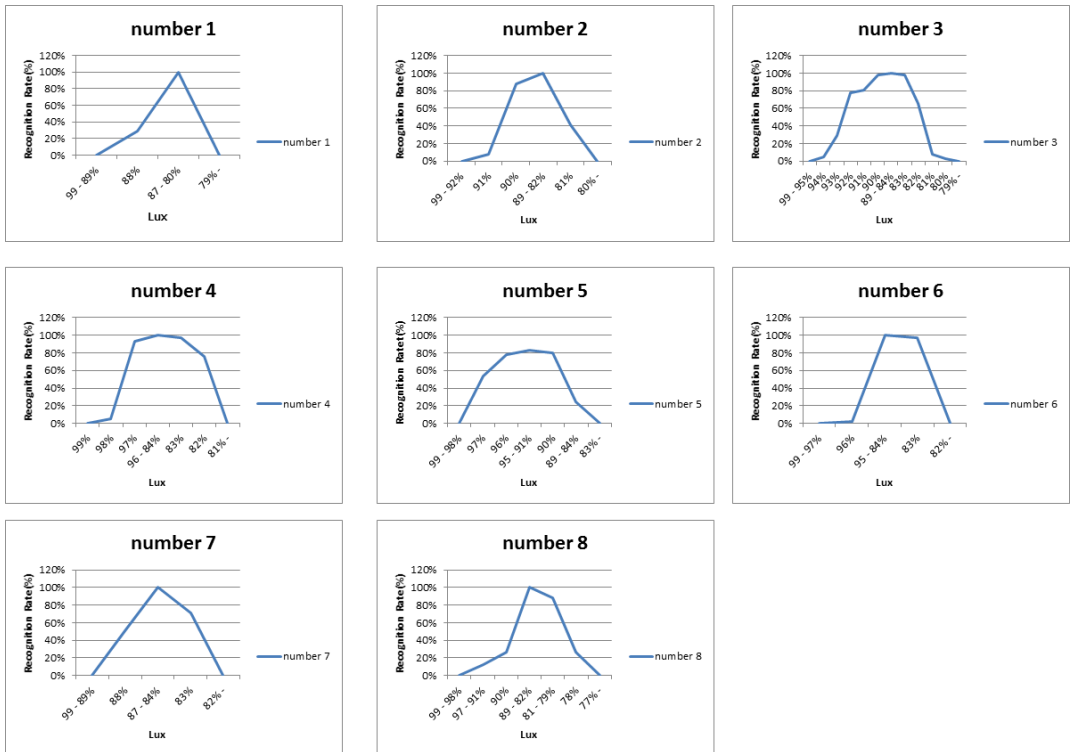

Fig. 9. Number Graphs.

mination rather than the threshold value change. Fig. 10 shows this result.

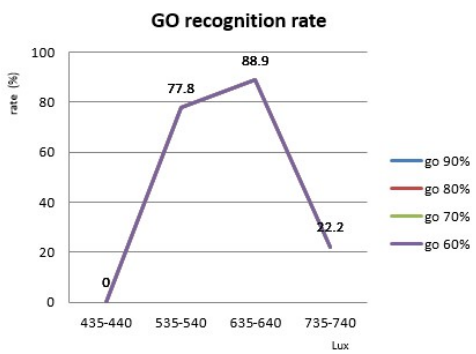

Fig. 10. Graph of play button recognition threshold change.

6. CONCLUSIONS

This paper presented the design and development of an Android tablet application for infant education using AR technology. Our software has the advantage of recognizing physical blocks only by AR based imaging technology without any physical connection between the block and the tablet. It extracts a region of interest(ROI) in an image acquired by the tablet's internal camera and separates the foreground image from the background image and recognize the object using the algorithm and shows the result on the tablet screen as AR. The recognition algorithm used different thresholds showing the best recognition rate per block. This block recognition algorithm was divided into two parts: the first part for recognizing numbers and the second for recognizing play button. Since the number recognition algorithm had different threshold value that shows the best recognition rate per number, the algorithm was adjusted. Play button recognition was sensitive to the change of illumination.

In order to have reasonably working recognition ratio, we did experiments with more than 5,000

frames of actual playing scenarios. We found that the recognition rate can be secured up to 95%, when the threshold values are selected well using various condition parameters.

Future research will be focused on the design and development of algorithm that recognizes non-numeric alphabet blocks. Also we will develop strong algorithm that recognize Play button regardless of illumination changes.

REFERENCE

[1] M. Ulas, S.M. Tasci, and S. Ulas, "Augmented Reality and Application Sample on Education," *International Journal of Advances in Software Engineering and Research Methodology*, Vol. 2, No. 1, pp. 44-46, 2015.

[2] M. Ulas and S. Tasci, "A Review About Augmented Reality Tools And Developing A Virtual Reality Application Based On Education," *Academic Journal of Science*, Vol. 2, No. 3, pp. 139-146, 2014.

[3] D.J. Johnson and H.R. Myklebust, *Learning Disabilities: Educational Principles and Practices*, *Grune and Stratton*, New York, 1967.

[4] R.A. Duschl and J. Osborne, "Supporting and Promoting Argumentation Discourse in Science Education," *Studies in Science Education*, Vol. 38, No. 38, p. 39-72, 2002.

[5] M. Felder and L. Silverman, "Learning and Teaching Styles in Engineering Education," *Proceeding of Frontiers in Education Conference*, Vol. 78, No. 7, pp. 674-681, 1988.

[6] J. Luo and D. Crandall, "Color Object Detection Using Spatial-Color Joint Probability Functions," *Institute of Electrical and Electronics Engineers Transactions on Image Processing*, Vol. 15, No. 6, pp. 1443-1453, 2006.

[7] K.J. Lee, "A Study on Gesture Recognition Using Edge Orientation Histogram and HMM," *Korea Institute of Information and Communication Engineering*, Vol. 15, No. 12, pp. 2648-2654, 2011.

[8] D.G. Lowe, "Distinctive Image Features From Scale-Invariant Keypoints," *International Journal of Computer Vision,* Vol. 60, No. 2, pp. 91-110, 2004

[9] H. Bay, E. Andreas, T. Tuytelaars, and L.V. Gool, "Speeded-up Robust Features," *Computer Vision and Image Understanding*, Vol 110, No. 3, pp 346-359, 2008.

[10] E. Rublee, V. Rabaud, and K. Konolige, "ORB: An Efficient Alternative to SIFT or SURF," *Proceeding of Computer Vision and 2011 Institute of Electrical and Electronics Engineers International Conference*, pp. 2564-2571, 2011.

[11] L. Stefan, M. Chli, and R.Y. Siegwart, "BRISK: Binary Robust Invariant Scalable Keypoints," *Proceedings of Institute of Electrical and Electronics Engineers International Conference on Computer Vision*, pp. 2548-2555, 2011.

[12] T.R. Park, "Implementation of Real Time Based Multi-Object Recognition Algorithm," *Journal of Institute of Korean Electrical and Electronics Engineers*, Vol. 17, No. 1, pp. 51-56, 2013.

[13] Y.H. Cho, "Shape Image Recognition by Using Histogram-Based Correlation," *Korean Institute of Intelligent Systems*, Vol. 20, No. 4, pp. 548-553, 2010.

[14] M. Resnick, J. Maloney, A. Monroy-hernadez, N. Rusk, E. Eastmond, K. Brennan, et al, "Scratch: Programming for All," *Communication of Association for Computing Machinery*, Vol. 52, No. 11, pp. 60-67, 2009.

[15] S. Meeklai, K. Obara, K. Sripraprutchai, Y. Hirai, and K. Kanekoa, "Development Of A Programming Experience System By Using Operation Cards," *Proceedings of ISERD International Conference*, pp. 31-37, 2016.

[16] A. Boyaci and M. Ulas, "Augmented Reality

Based Expert Learning Assistant Model," *Proceedings of 66th The International Institute of Engineers and Researchers International Conference*, pp. 4-7, 2016.

[17] M. Kim, "Image Retrieval Using Histogram Refinement Based on Local Color Difference," *Journal of Korea Multimedia Society,* Vol. 18, No. 12, pp. 1453-1461, 2015.

[18] S.H. Ko, M.H. Yang, S.H. Oh, S.H. Hong, and S.S. Hwang, "Image Matching-based Mobile Application Providing Indoor Place Information," *Journal of Korea Multimedia Society*, Vol. 19, No. 8, pp. 1424-1432, 2016.

Seunguk Yeon

Aug. 1992: Bachelor's degree of Computer Engineering from Seoul Nat'l Univ., Seoul, Korea

Aug. 1994: Master's degree of Computer Engineering from Yonsei Univ., Seoul, Korea

Oct. 2016~current : Assistant professor of Information Security, Seoul Women's Univ., Seoul, Korea

Research Interests : Massive Server Programming, Cloud Computing, Database, IoT

Sukyong Seo

Feb. 1998: Master's degree of Public Health from Seoul Nat'l Univ., Seoul, Korea

May 2007 : PhD from School of Publid Health, UC Berkeley, CA, USA

Sept. 2014~current: Assistant professor of Community Public Health Nursing, Eulji Univ., Sungnam-si, Korea

Research Interests: Health Education, Chronic Disease Management, Health Economics