2018 International Conference on Computer, Communication and Network Technology (CCNT 2018)
ISBN: 978-1-60595-561-2

A Gesture-based Augmented Reality System at Mobile Terminal

Meng-ting XIAO[1,2], Zhi-quan FENG[1,2,*], Tao XU[1,2],
Xiao-hui YANG[1,2],Chang-sheng AI[3] and Wei XIE[4]

[1]School of information Science and Engineering, University of Jinan, 250022 Jinan, China

[2]Shandong Provincial Key Laboratory of Network Based Intelligent Computing, 250022 Jinan, China

[3]School of Mechanical Engineering, University of Jinan, 250022 Jinan, China

[4]School of Information and Electrial Engineering, Harbin Institute of Technology at Weihai, 264209 WeiHai, China

*Corresponding author

Keywords: Augmented Reality, Gesture recognition, Vuforia, Light interaction.

Abstract. The combination of Augmented Reality (AR) technology and gesture recognition has become increasingly important in the educational system. In the smart classroom application scenario, the virtual three-dimensional complex model is combined and interacted with real hands based on augmented reality technology. The main contribution of this paper is to determine the coordinates of the five-finger tip of the hand, and calculate the initial length from the center of the 3D model. When the distance between five fingertips and three-dimension mode is shorter than the initial length during the process of fingers grasp, the three-dimension mode will get smaller. Light interaction is achieved by establishing the relationship of the two spatial models with coordinates. AR is achieved under Vuforia, and the three-dimensional model gradually becomes smaller with the five fingers grasping. The proposed method is used in an Intelligent Teaching System and run at the mobile terminal. The experimental results show that the proposed method works well in ITS.

Introduction

Now we are entering the era of information, human-machine interaction is hindered by the forth putting of mouse, keyboard and other control devices. Gestures have replaced physical devices and become one of the main forms of human-computer interaction[1,2]. Therefore, the combination of gestures and AR enhances the immersion in the virtual space[3]. João Paulo Lima et al. [4]used AR in the automotive industry to achieve the functionality of the car through on-site modeling, system calibration, and tracking. There is also proposed a user behavior-driven augmented reality method to infer the user's current status by cross-checking their past behavior[5].

In recent years, most of Chinese AR technologies have been used in research and education to improve students' learning effectiveness in the teaching activities [6,7]. But these lack of immersion in the virtual scene. In order to make students better immerse in the virtual space, this article based on the visual method to realize the model tracking of the fusion of gesture and virtual models. The integration of gesture recognition and AR not only eases the burden on teachers, but also students can experience the joy of learning in entertainment and help students develop their thinking skills[8,9].

Related Work

AR methods are divided into hardware-based method and vision-based method. Hardware devices include AR glasses, Microsoft Hololens and other wearable device[10]. José Miguel Mota et al. [11] proposed a creative tool that they use AR as a learning tool to achieve augmented reality. Through the framework of learning analysis, educators and students can make assessments. The inadequacies lack the interactive nature of AR. This method is not combined with interactive devices such as Leap motion.

A vision-based method captures the characteristics of the image, visualizes the image and connects the three-dimensional model with two-dimensional image points. To set up and calculate the position of the camera, finally achieve the augmented reality model tracking. Zhimin He et al. [12] studied gestures and head movements as the main modes of interaction, using voice commands and touch interaction modes as auxiliary interaction modes. It is insufficient that we have to touches the screen with our hand. Yelda Turkan et al. [13] studied mark-based AR applications for registration of AR video virtual content on a tablet computer. The shortcoming of this method lies in the interactive mode of video and interface. AR technology researched by Chenxi Sun et al. [14] is the three-dimensional model of gestures and voice animation systems, gesture interactions, particle effects, real-time color mapping, and game interactions in recently. Aiming to increase children's imagination, it is better to realize the combination of virtual and reality. We propose a gesture recognition algorithm for AR, which combines gestures and three-dimensional virtual models to realize interactivity.

Overview of our Method

Gesture Recognition Algorithm.

Kinect, Data Glove and other sensing devices are not requested in proposed gesture recognition. In order to recognize the hand gesture by a monocular camera, this experiment achieves gesture recognition in OpenCV. The gesture recognition based on neural network transforms the detected hand preprocessed data into a fuzzy gesture feature model. It tests behavioral models and fuzzy reasoning to determine actual gestures[15].

We process gestures feature in OpenCV. First of all, the gestures were collected. The experiment uses two gestures to interact with the virtual model. The gestures divided into the static five fingers open gesture and the dynamic five fingers grasp gesture. Then the gesture images are processed with OpenCV. Thanks to skin color detection, we translate the hand space to HSV space color segmentation. Finally, we can get the value of each gesture image in a particular area and outline extraction of images. After getting the gestures outline, we get the gesture information achieve gesture recognition.

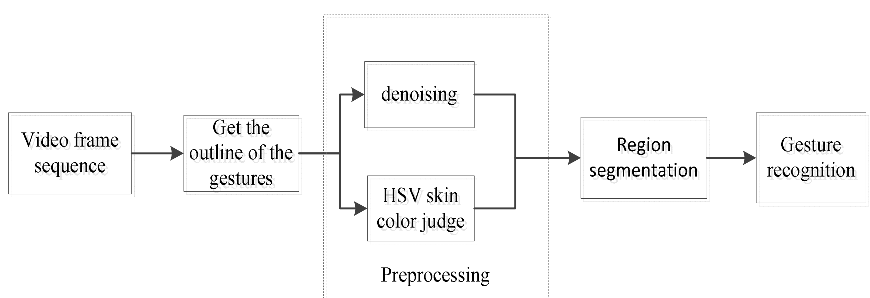


Figure 1. OpenCV gesture recognition processing.

Vuforia Tracking Algorithm.

The Vuforia recognition classes used are Image Targets in this article and single instance Camera. In Image Targets, we can detect the image's target and render the simple three-dimensional model, multiple target pictures and 3D model at the same time, and also activate the database in multiple devices (Fig. 2). In the Vuforia tracking algorithm, the Tracker tracker can be used to recognize the images on the camera side and store the target object in a way of State Object. It can simultaneously track multiple data sets. In this article, the hidden settings for the images are recognized. When the tracker tracks the identified object, reload the target object before destroying it. After the target object is lost, I use to another value instead of a layer of camera to render the off card three-dimensional model. Last the model position is recorded.

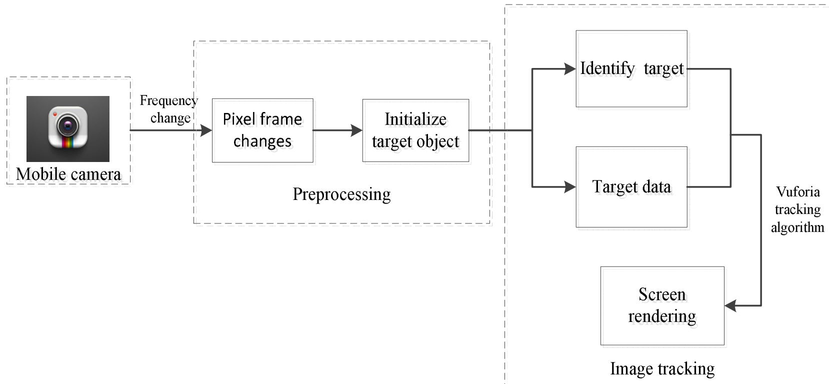


Figure 2. Vuforia SDK flow diagram.

The tracking algorithm is described as:

Step 1: Initialize the Tracker class and register the tracking event processing supervision.

Step 2: To determine whether the target object is to detect or track the state and start execution when it is in the tracking state.

Step 3: When the target object is tracked, execute the object traceable method to get the Render and Collider components of the child object. The Render iterated through component array and makes it available, then the Collid iterated through component array and makes it available.

Step4: Follow each of the above Step 1 and 2 to track each target and traverse the array.

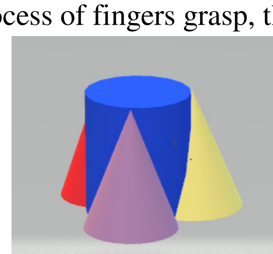
According to the gesture tracking in the Vuforia SDK, the three-dimensional model appeared after the target object been recognized(Fig. 3a). The center point of the model is determined as the center point of the three-dimensional coordinate system. When I join this hand, the model under the phone will move with the hand. Put your hand on the virtual model and determine the five fingertips of the camera coordinate system(Fig. 3b), and create a coordinate set $A=\{P_1, P_2, P_3, P_4, P_5\}$, among $P_i=(X_i, Y_i, Z_i)$, $i \in \{1,2,3,4,5\}$. Initial length,

$$H_i = \sqrt{X_i^2 + Y_i^2 + Z_i^2}\,. \tag{1}$$

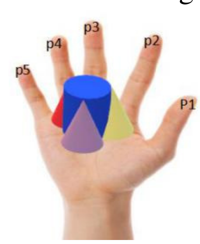
$$h_i = \sqrt{x_i^2 + y_i^2 + z_i^2}\,. \tag{2}$$

$$h_i < \min\{H_i\}\,. \tag{3}$$

When the distance between five fingertips and three-dimension mode is shorter than the initial length during the process of fingers grasp, the three-dimension mode will get smaller.



(3a)



(3b)

Figure 3. Three-dimension mode(3a) and integration of hand and three-dimension model(3b).

Algorithm Performance

The algorithm is transformed into a two-point distance calculation for the dynamic transformation of hand in this paper, and used each coordinate point to change the size of the model. The running time of the algorithm is reduced, and the real-time performance of the algorithm has been improved. In the case of changing the parameters, the performance of the algorithm does not change more. From the above, the method we have proposed increase the recognition rate of the model and greatly increases the robustness of the algorithm.

Experimental Results and Analysis

We should adjust the position of the AR Camera, so that the target object and the three-dimensional model can appear in the middle of the camera screen. Then the camera at the mobile end is aimed at the target object, and the three-dimensional geometry is presented when the target object is recognized. Finally, the hand slowly enters the AR scene and stays below the model, and the three-dimensional model is connected with hand, thus we can see that three-dimensional geometry gradually becomes smaller in the process of grasping five fingers. During this process, so as to maintain a smooth movement, it is very important that manual amplitude should not be too large to dynamic gesture.

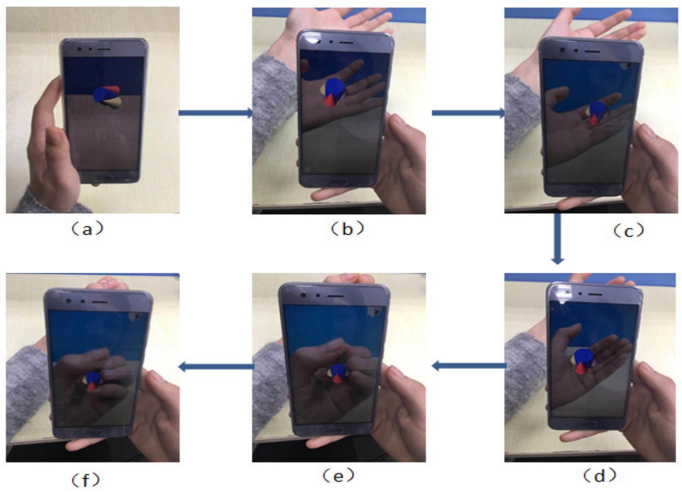


Figure 4. Experiment results.

Figure 4. With moving the mobile terminal model closer to the center of the screen (a), the hand slowly enters the scene until the camera captures the hand in the scene and keeps it stationary state, and try to make the three-dimensional model in the center of the hand (b). The hand begins to move slightly after a while, during the process the model moves in the same direction with the hand. After stopping the movement, the hand begins to gesture of five-finger grasp, and five fingers slowly move closer to a smaller model (c-d). At last, the contraction of the five fingers gradually increases, and when the fingers are almost all close together, the three-dimensional model gradually shrinks. (e-f).

The results show that the AR makes our better immerse in virtual space through gesture recognition. But there are many influencing factors in this experiment. Firstly, the distance between three-dimensional model and the fingers is not well controlled during the experiment. Secondly, when the interior is dark, the recognition performance of the mobile camera is reduced, so the focus ability of the camera will be weaker and it is easy to reduce the target recognition rate. Thirdly, it appears screen unstable and sluggish when we run in the Vuforia SDK. Vuforia SDK also has limitations.

Conclusion and Future Work

We can obtain a gesture-based AR system at mobile terminal. AR is realized by Vuforia, and the distance judgment model of the fingertip is determined according to the gesture. Ultimately, we have achieved virtual reality integration of hands and three-dimensional models through this experiment.

There are still some deficiencies in this paper. For example, we need to integrate gestures and 3D models more precisely, further improve the robustness of gesture recognition, and improve the map of experimental scenes. The further research work, on the one hand, is to get rid of the Vuforia SDK resource package, using the Android development environment and OpenCV algorithm to achieve gesture recognition in augmented reality. On the other hand, we are ready to add various gestures to manipulate 3D models and increase the speed of gesture recognition under the camera.

Acknowledgments

This paper is supported by the National Key R&D Program of China (No. 2016YFB1001403) and the National Natural Science Foundation of China (No. 61472163, 61603151).

References

[1] Aashni Haria, Archanasri Subramanian, Nivedhitha Asokkumar, Shristi Poddar, Jyothi S Nayak, Hand Gesture Recognition for Human Computer Interaction, J. Procedia Computer Science. 115(2017) 367-374.

[2] Feng Zhiquan, Jiang Yan, A Survey of Hand Gesture Recognition, J. Journal of University of Jinan (Sci. & Tech.). 27(2013) 336-341.

[3] Yelda Turkan, Rafael Radkowski, Aliye Karabulut-Ilgu, Amir H. Behzadan, An Chen, Mobile augmented reality for teaching structural analysis, J. Advanced Engineering Informatics. 34(2017) 90-100.

[4] João Paulo Lima, Rafael Roberto, Markerless tracking system for augmented reality in the automotive industry, J. Expert Systems with Application. 82(2018) 100-114.

[5] Chung-Hsien Tsai, Jiung-Yao Huang, Augmented reality display based on user behavior, J. Computer Standards &Interfaces, 55(2018) 171-181.

[6] Ting-Chia Hsu, Learning English with Augmented Reality: Do learning styles matter?, J. Computers & Education. 106(2017) 137-149.

[7] Moro Christian, Stromberga Zane, Raikos Athanasios, Stirling Allan, The Effectiveness of Virtual and Augmented Reality in Health Sciences and Medical Anatomy, J. Anatomical Sciences Education, 10(2017) 549-559.

[8] Xie Shuxiao, Chen Yuchao, Liu Guan, Exploration and practice for construction of smart classroom, J. Experimental Technology and Man. 33(2016) 221-224.

[9] Ram Pratap Sharma, Gyanendra K. Verma, Human Computer Interaction using Hand Gesture, J. Procedia Computer Science. 54(2015) 721-727.

[10] Zhou Zhong, Virtual reality augmented reality review, J. Science China Press. 45(2015) 157-180.

[11] José Miguel Mota, Iván Ruiz-Rube, Juan Manuel Dodero, Inmaculada Arnedillo-Sánchez, Augmented reality mobile app development for all, J. Computer and Electrical Engineering. 65(2018) 250-260.

[12] Zhimin He, Tao Chang, Siyu Lu, Hong Ai, Dong Wang, Qiang Zhou, Research on Human-Computer Interaction Technology of Wearable Devices Such as Augmented Reality Supporting Grid Work, J. Procedia Computer Science. 107(2017)170-175.

[13] Yelda Turkan, Rafael Radkowski, Aliye Karabulut-Ilgu, Amir H. Behzadan, An Chen, Mobile augmented reality for teaching structural analysis, J. Advanced Engineering Informatics. 34(2017) 90-100.

[14] Sun Chenxi, The Design and Implementation of Children's Education System based on augmented reality, Shandong University. 2017.

[15] Elakkiya R, Selvamani K, Kanimozhi S, Velumadhavva Rao, Kannan A, Intelligent System for Human Computer Interface Using Hand Gesture Recognition, J. Procedia Engineering. 38(2012) 3180-3191.