AR Marketing Application Based on iOS Platform

Zheng Zhang[1, a], Xin Wang[1, 2, b, *], Hongchen Jiang[1, c], Yuemeng Hu[1, d], Jiayu Du[1, e], and Qiang Zhang[1, f]

[1]College of Computer Science and Technology, Jilin University, Changchun, China

[2]Key Laboratory of Symbolic Computation and Knowledge Engineer of Ministry of Education, Jilin University, 130012 Changchun, China

[a]zhangzheng2116@mails.jlu.edu.cn, [b]w_x@mails.jlu.edu.cn, [c]jianghc2116@mails.jlu.edu.cn, [d]huym2116@mails.jlu.edu.cn, [e]dujy2116@mails.jlu.edu.cn, [f]zhangqiang2116@mails.jlu.edu.cn

*Corresponding author

Keywords: Augmented Reality, Entity marketing, Image Identification, Online and offline technique

Abstract: Due to the development of computer vision, in recent years, augmented reality grows rapidly. A lot of augmented technologies have been applied in actual problems. At present, the industry of traditional marketing requires a lot of manpower, and the application of AR technology in marketing industry can greatly reduce industrial costs. In this paper, an Augmented Reality Marketing Application based on iOS platform is designed. Users can use this APP to scan physical stores signboards, which are saved previously, to trigger a 3D stereo model combined with the real world. At the same time, through the interaction with the model, users can get the business information and its sales activities, thus enhancing the consumer interest in the physical stores.The experiment results show this application work feasible and reliable. It is an attempt to integrate online and offline economy.

1. Introduction

At present, small physical stores can't escape the crisis of bankruptcy because of online business. The online shops are changing the shopping behavior and habits of consumers. Physical stores need to seek new development models [1].

As one of the most activity part in CV, augmented reality is a hit point for the researcher. In recent years, augmented reality technology has been used more and more in our lives. For example, Matsutomo had made an augmented reality evaluation system of education [2], and W.Friedric applied augmented reality technology to the monitoring system [3]. In the economic field, JH Arrasvuori proposed the concept of augmented reality advertising [4].

Today, smartphone-based AR SDKs, like Vuforia, ARkit, EasyAR and FAR-Play, have matured. Combining AR with physical business marketing, and designing and developing an AR marketing application is feasible.

Compared with traditional business marketing, the AR marketing application designed in this paper provides a novelty AR experience for the users when they want to know something about a store --combining 3D models with scenes around stores.

2. Related Work

2.1 Target picture feature extraction

SIFT [5] is one of the most popular and classic feature point detection methods. It is considered to be the best method for image matching, and has good stability in the scale change of the object, rigid body transformation, illumination intensity and occlusion. In this paper, it is used to extract the feature points of each frame of the picture captured by the mobile phone.

2.2 Picture recognition

Image recognition consists of two parts: feature point matching and picture similarity measuring. The key issue is the matching of feature points. If we use the simplest linear matching, the speed is far from the performance requirement. In this paper, we use the data structure algorithm K-means clustering algorithm [6] to construct high-dimensional feature space in order to reduce the dimension of SIFT feature vector and speed up the matching.

2.3 Generation of AR

The effect of the AR directly leads to a good user experience. The construction of AR involves three parts: the understanding and tracking of the real world, the rendering of the 3D AR world, the interaction between the user and the 3D world [7]. Fortunately, the AR SDK based on the smartphone platform is mature enough to fulfill the task of generating the AR world in this application. The application designed in this paper uses ARkit and Scenekit.

3. Approach

In this section, we detail the working methods of the AR marketing application in this article.

The whole process is shown in Figure 1.We refer the real-time scene image captured by the mobile phone as the target image, and the merchant signature image saved in the background database as the template image. After the feature image is extracted by the SIFT algorithm, the target image is matched with the template image that has been subjected to feature extraction in advance, and the matching result directly affects whether the renderer is started to perform model rendering and model tracking.

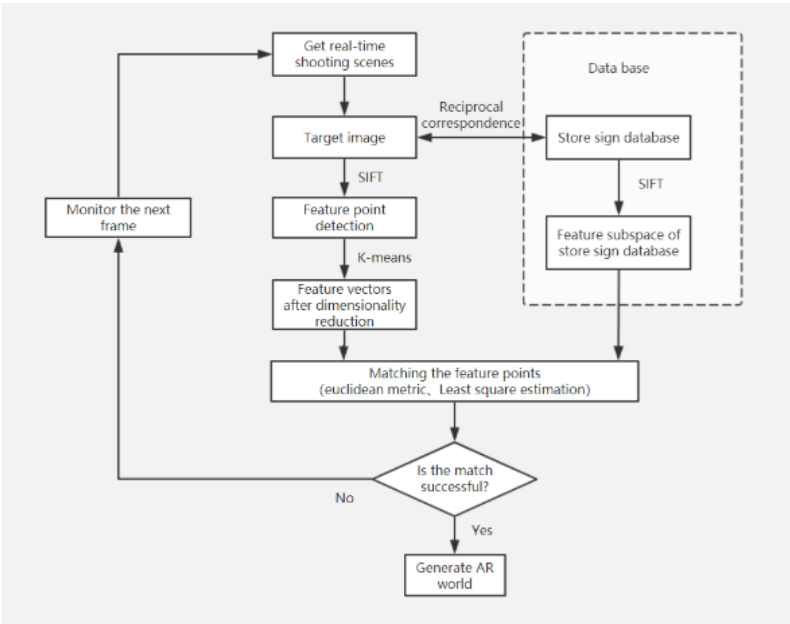

Figure 1. The whole process

3.1 Feature point detection

In our AR marketing application, the most important job is image recognition. The image feature point detection in this part is the basis of image recognition. The accuracy of the feature point's detection directly determines the accuracy of image recognition.

In the real scene, when the user open the camera to scan the store signboards, some external factors change, such as lighting changes, may bring difficulties to image feature point extraction. Considering the above problems, we decided to use the SIFT algorithm, which does not change along with the factors change,such as angle of scanning, brightness change, shooting angle and noise, can solve the above problem to some extents.

Because SIFT algorithm is very common, we don not describe the process in detail.

SIFT algorithm can be divided into four stages: scale space construction, feature point determination, direction determination and feature point descriptor generation. SIFT algorithm convolutes the original image with Gauss function to get a certain scale space of the image,then calculates the value of feature points using the magnitude and direction of the pixel gradient in the neighborhood of the Gauss pyramid. In the 4X4 neighborhood of feature points, gradient information of 8 directions in each neighborhood of 16 neighborhoods is recorded, and 128-dimensional feature description vectors are formed. At last step is normalizing the 128-dimensional descriptor vector to form the final descriptor.

3.2 Image recognition

The image recognition process is divided into two steps: feature point match and picture similarity measure. Considering the purpose of the application designed in this paper, the efficiency of image recognition is the key to this application.

Because the SIFT key point feature vector has a high dimension, its time complexity is high. If the simplest linear matching is used, the speed is far from the performance requirement. K-means clustering algorithm is a kind of linear or non-linear clustering method [8]. Based on this, in this paper, we use the K-means to reduce the dimension of the SIFT feature vector.

In the whole K-means algorithm, the selection of the initial cluster center is very important. In this paper, the SIFT feature vector matrix with the smallest average difference is calculated as the initial cluster center. The average difference is defined as follows formula (1):

$$d_i = \frac{\sum_{j=1}^{N} d_{ij}}{N} \quad (1)$$

$$d_{ij} = \sqrt{(x_{i1} - x_{j1})^2 + \cdots + (x_{im} - x_{jm})^2} \quad (2)$$

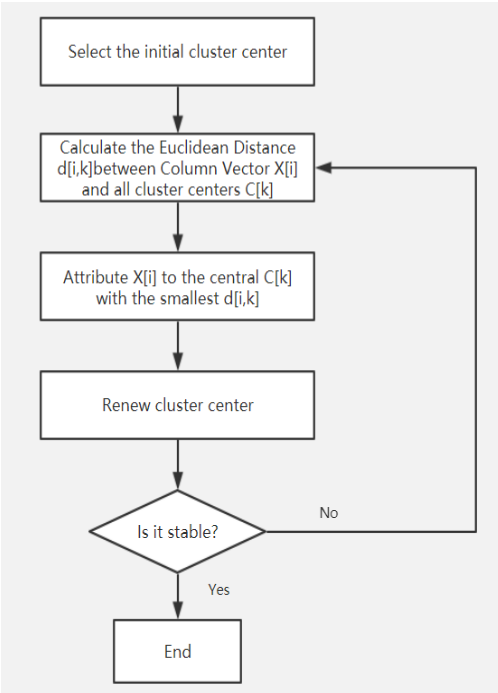

Figure 2. K-means process

m is the number of SIFT key points, and , are the two m-dimensional SIFT feature matrix column vectors; the distance is, and N is the total number of column vectors. The smaller the average difference, the higher the similarity between the sample and the whole. The K-means algorithm flow is shown in Figure 2.

The following is a step of matching the feature points extracted by the SIFT algorithm using the K-means algorithm:

1) The 128-dimensional SIFT feature vector matrix is clustered by K-means algorithm, and the column vector is divided into K classes.

2) Matching the t-dimensional feature vectors classified into the first class, the specific steps are as follows:

a) Calculate the Euclidean distance between the feature vectors between the feature vectors;

b) Traverse each key point in the template image to find the two key points closest to and closest to the Euclidean distance of this key point. If equation 3 is satisfied, it is considered to be a pair of copy-and-paste matching points.

$$\frac{D_{nearest}}{D_{hpyo-nearest}} < r \tag{3}$$

3) Through the setting of the initial test threshold, it is finally determined whether the target picture contains a template picture, that is, whether the user scans the signboard of the store in the database.

According to the result of image recognition, application decides whether to generate AR scene.

3.3 AR world generation

The construction of AR involves three parts: the understanding and tracking of the real world, the rendering of the 3D AR world, the interaction between the user and the 3D world [9].

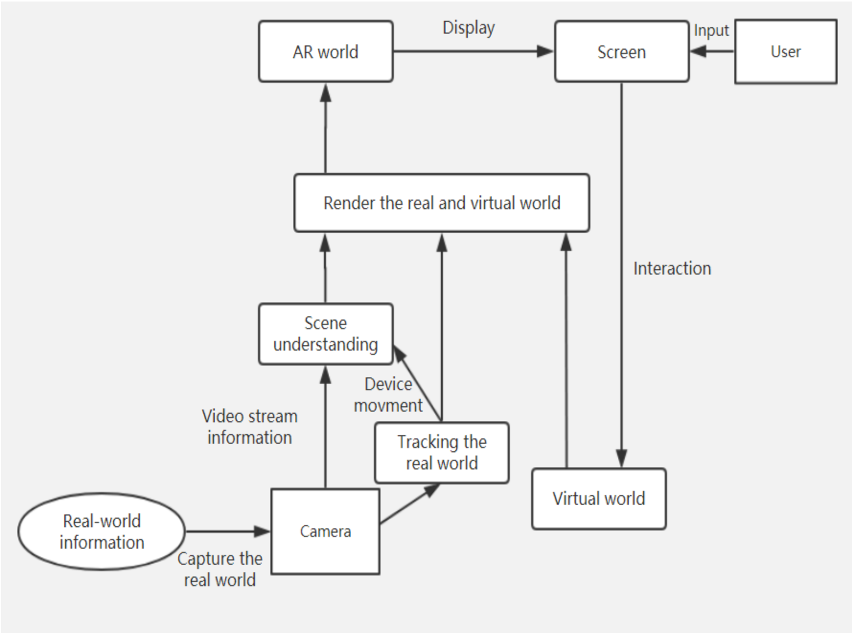

Figure 3. AR generation process

As shown in the figure3, real-world information is captured by the camera and converted into video stream information, which is used to understand the scene. At the same time, the virtual world is rendered to the real world designated position based on the understood scene, then application displays augmented reality world on the screen of the mobile phone.

The AR world generation module based on ARKit [10] and SceneKit is as follows:

(1) Create a world coordinate system. In SceneKit, SCNCamera is used as a shooting lens, and SCNScene is used as a real-time scene scanned. Everyone or object in the scene can be regarded as SCNNode. Set SCNCamera to Create RootNode as the origin of the world coordinate system. In this coordinate system, each SCNNode has a unique coordinate. Prepare to render the virtual world to the real world designated location.

(2) Capture the real world. By calling ARSession to manage and configure the main process of the entire AR experience, application reads the data from the device's motion sensor to control the built-in camera, captures the real-time image with AVCaptureSession, analyzes the captured image, and outputs it to the ARFrame instance. .

(3) Track the real world. The ARFrame receives the real-time image captured by the AVCaptureSession and generates an instance of the video image with position-tracking information. For the video frame of the device camera captured continuously bythe AR Session, ARKit will analyze the data from the device motion sensing hardware to estimate the actual position of the device, and pass the current frame position information and image parameters through the ARFrame.

(4) Scene Understanding. The purpose of scene understanding is to place virtual objects into the real world, which usually consists of Plane detection, Hit-testing, and Lighting estimation.

After the above steps, the virtual world, the captured real world, the world-tracked information, and the scene understanding information have all been acquired, and they are rendered together by calling the rendering engine to form an AR world.

3.4 User interaction

The user interaction part acts as a bridge to establish a connection between the user and the AR world. The user sends a command to the system to interact with the virtual world by clicking the screen. After the virtual world responds to the interaction, the rendering is performed again, and when the user moves the device, the captured real world changes, and the scene needs to be re-understood and re-rendered. Through interaction, the user can obtain the physical store information of interest, which improves the interest of the user in the physical store, and realizes online and offline integration.

4. Results

This section describes the results of the use of the AR application designed in this paper. The results based on the above algorithms, ARKit and SceneKit. The test computer used in this section is Samsung 500r5l, CPU frequency dual core 2.5GHZ, memory 4GB, hard disk capacity 500GB; operating system is Windows 10; test phone is iPhone8 Plus, memory 64G, mobile phone system is iOS 12.1.4.

4.1 Image recognition

The database of the experiment is a sample of 1000 advertisement images of the merchant's signboard. The test images are 100 merchant's signboard pictures captured by iPhone in different environments, different angles and different lighting conditions. As shown in Figure 4, three image recognition experiments are performed based on the SIFT algorithm and the K-means algorithm. The corresponding experimental data statistics are shown in Table 1.

Figure 4. Experimental results of image recognition

Table.1. Analysis of experimental

Experiment	Number of features	Accuracy	Time/s
One	437/386	94.00%	2.056
Two	1964/2124	84.50%	3.245
Three	1132/1235	91.65%	2.854

It can be seen from result that the correct recognition rate of the picture, the time and accuracy of the SIFT and the K-means algorithm used in this paper are significantly high. It can reach 91.65%, which is enough to meet the need of accuracy in this application.

4.2 AR scene

As shown in figure 5, the AR world has been displayed on the mobile phone after the user scanned the store signboard.The user can interact with the model in the AR world through the button, thereby obtaining the store information, which can raise the user's interest in consumption and connect reality and virtuality, online and offline.

Figure 5. AR effect of application display

5. Conclusions and expectations

In this paper, an AR marketing application based on iOS platform is implemented, including image feature point extraction, image recognition and matching, AR world construction and user interaction. The main contribution of this paper is presented. It integrates various algorithms in computer vision and apply those to actual problems. SIFT is used for image feature point extraction, K-means is identification of store signboard is improved. ARKit and SceneKit has been used for the construction of the AR world, from the experimental results, the experience for the user is good.

In future work, it is feasible to combine the characteristics of images of store signboards and find better algorithm to reach higher accuracy.

References

[1] T. Kot and P. Novák, “Utilization of the Oculus Rift HMD in Mobile Robot Teleoperation”, Applied Mechanics and Materials, 2014, pp. 199-208.

[2] Matsutomo S, Miyauchi T, Noguchi S, “Real-Time Visualization System of Magnetic Field Utilizing Augmented Reality Technology for Education”, IEEE Transactions on Magnetics, 2012, 48 (2): 531-534.

[3] Friedrich W, Wohlgemuth W, “Operating and monitoring system utilizing augmented reality technology”: WO, WO/2000/052536 [P]. 2000.

[4] Olsson T, “Concepts and Subjective Measures for Evaluating User Experience of Mobile Augmented Reality Services”, Human Factors in Augmented Reality Environments. Springer New York, 2013, pp.135-142

[5] Lowe DG, “Distinctive Image Features from Scale-Invariant Keypoints”, In: International Journal of Computer Vision. 2004, pp 91-110.

[6] Nguyen B, Baets B D, “Kernel-Based Distance Metric Learning for Supervised k-Means Clustering”, IEEE Transactions on Neural Networks and Learning Systems, 2019, pp (99): 1-12.

[7] Fenu C, Pittarello F, “Svevo Tour: The Design and the Experimentation of an Augmented Reality Application for Engaging Visitors of a Literary Museum”, International Journal of Human-Computer Studies, 2018:S1071581918300533.

[8] Georgios D, Fernando B, Felix L, “Improving Imbalanced Learning Through a Heuristic Oversampling Method Based on K-Means and SMOTE”, Information Sciences, 2018: S0020025518304997.

[9] Kinateder M, Gualtieri J, Dunn M J, “Using an Augmented Reality Device as a Distance-based Vision Aid-Promise and Limitations”, Optometry & Vision Science Official Publication of the American Academy of Optometry, 2018.

[10] UM Dilek, “Detecting Position Using ARKit”, Physics Education, 2018, 53 (2): 025011.