Names: Group 2 Hu yiteng Yu zhenyue Wangqi

Date: 8/7

Course: ST8601

Score:

Project name: **Facenet : a cloud based AR social application framework**

Facenet : a cloud based AR social application framework

YiTeng Hu

Hangzhou Dianzi University, School of Honor, Hangzhou 310000, China

Abstract

With the development of social media, the social relationship between people becomes more and more complex, and the social network becomes bloated. We build a social network model and use the cloud database to store it. This social network clearly represents the social relationship between people. We have established a unique social ecosystem called **facenet**, which is different from Facebook or WeChat. It has the cutting-edge AR functionality and users can easily find the shortest **social connection** between them and any stranger in the world through this application. We also designed an application framework, which clearly shows the software architecture of facenet and the relationship between various modules. People can create their own cloud-based AR applications through this framework.

Keywords：social network, augmented reality, breadth first search, facenet

1. Introduction

1.1 Background

According to the **theory of Six Degrees of Separation**, there will be no more than five people between you and any stranger, that is, you can know any stranger through up to five people. We hope that through our project, people can see a person on the street, know the social relationship between users and the observed people, and display it on AR glass through AR technology.

1.2 introduction to AR

Augmented reality (AR) enables people to experience a 3D real-world environment where they can interact with 3D objects that exist in the real world. Augmented reality takes advantage of the information in the form of audio, graphics, and texts.

1.3 introduction to social network

A social network is defined as several nodes that are tied by one or more types of relationships. Social networking applications, such as Facebook or Wechat, play an increasingly important role in accessing information and making contact with other individuals in our daily life. Our application aims to utilize social network services as our resources and collect profile photos from users.

2. Related Work

In this section, we briefly introduce a social network model called Facenet and review how to realize the AR function through CNN models.

2.1 Alexnet

[6]Alexnet was proposed by Alex Krizhevsky in 2012 and won the champion of ILSVRC competition in 2012. The prediction error rate of top5 was 16.4%, far surpassing the first place. Alexnet uses 8-layer neural network, 5 convolution layers and 3 full connection layers (the maximum pooling layer is added after the 3 convolution layers), containing 630 million links, 60 million parameters and 650,000 neurons.

3. Solution Design

3.1 Overview

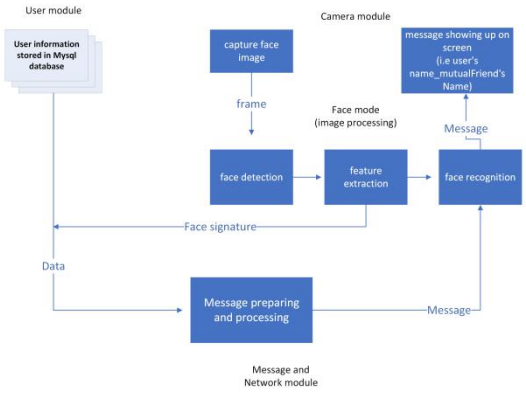


Figure 1: facenet framework

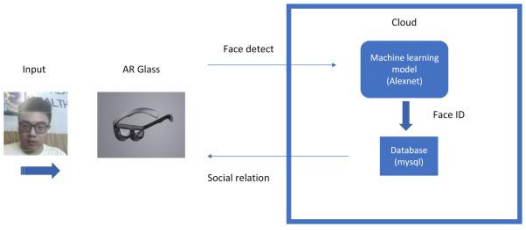


Figure 2 : facenet overview

3.2 construction of Social Network

3.2.1 rules

The social relationship network can be abstracted into an undirected graph without weight. Each person can be regarded as a node in the graph. If two people know each other directly, there is an edge between the two nodes. Therefore, If two people know each other through several mutual friends, there is an indirect path between the two corresponding nodes.

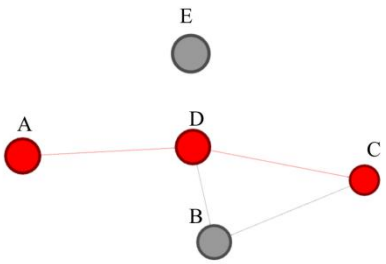


Figure 3: social network diagram

3.2.2 finding the shortest path between two people using BFS

With the increase of users, the size of social network will inevitably increase exponentially, so how to efficiently find the shortest social distance between two nodes has become an important problem. We decide to use breadth first search in our project. The pseudo code of BFS are shown in the figure2.

We draw a schematic diagram of a 4-size social network (Figure 1). Considering that there are two nodes A and C, we can see that the paths from A to C include A-D-C and A-D-B-C. although A and C can get to know each other through additional B, the path of A-D-C is the shortest, so we choose to display this relationship path in our application.

BFS Algorithm will establishing 2 empty lists and 2 queues which are user's and target's friend's list, then start looping. During looping the system will do a determination if a name in the user's list is a name in target's list, if true the whole process will return the combination of 2 already established list from the user and target, it is false then keep searching. The two for loop will add the name to the list from the queue if the if statement returns false.

3.3.3 storage of the social network in mysql database on the cloud

MySQL Database Service is a fully managed database service to deploy cloud-native applications.[5] We choose this databse because it has the advantages of small volume, high speed and low overall cost of ownership.Moreover, MySQL can run on multiple systems and supports multiple languages.

There are mainly two tables in our MySQL database. One is called network, which stores the structure of social network in the form of **adjacency matrix**, and the other is userinfo, which stores some users' personal information. It is worth noting that we use **blob** instead of path to store users' personal photos, The reason is that once the physical way of storing pictures changes, it is troublesome to modify the face picture field in each piece of data.

```
mysql> desc network
    -> ;
+------------+--------------+------+-----+---------+----------------+
| Field      | Type         | Null | Key | Default | Extra          |
+------------+--------------+------+-----+---------+----------------+
| edge_id    | int unsigned | NO   | PRI | NULL    | auto_increment |
| node_id_1  | int          | NO   |     | NULL    |                |
| node_id_2  | int          | NO   |     | NULL    |                |
+------------+--------------+------+-----+---------+----------------+
3 rows in set (0.01 sec)

mysql> desc userinfo;
+-------------+--------------+------+-----+---------+----------------+
| Field       | Type         | Null | Key | Default | Extra          |
+-------------+--------------+------+-----+---------+----------------+
| uid         | int unsigned | NO   | PRI | NULL    | auto_increment |
| name        | varchar(20)  | NO   |     | NULL    |                |
| sex         | varchar(6)   | YES  |     | NULL    |                |
| prefer      | varchar(60)  | YES  |     | NULL    |                |
| password    | varchar(20)  | YES  |     | NULL    |                |
| facepicture | blob         | YES  |     | NULL    |                |
+-------------+--------------+------+-----+---------+----------------+
6 rows in set (0.01 sec)
```

Figure 4: structure of our database

```
Q, Insert(x), then mark x as read
P, Inser(y), then mark y as read
while(Q not empty & P not empty)
    if(Q not empty)
        i = Q.getFirst()
        if(i = y || i belongs to P)
            return true
        for(u belongs to U(i))
            i' = f(i,u)
            if(i' not read)
                mark i' as read
                Q.append(i')
            else
                keep only 1 i'
    if(P not empty)
        i' = P.getFirst()
        if(i' = x || i' belongs to Q)
            return true
        for(v belongs to U^-1(i'))
            i = f^-1(i', u^-1)
            if(i not read)
                make i as read
                P.append(i)
            else
                keep only 1 i
return false
```

Figure 5:Breadth first search algorithm

3.3 Face Recognition module

As shown in the graph above, basically our face module has three phases of processing images that are captured by a camera. Detection is the process of finding an image: focusing on finding a human face based on the frame. Feature extraction is the second step that maps faces by determining the distinguishing elements of a unique face, including eyes distances, nose shapes, or lips. The final step is confirming the identity of a person and searching for the information of that specific person in our database.

As deep learning methods can lead us accurate results, the team decide to apply Convolutional Neural Network (CNN) into our face recognition systems. [2] A convolutional neural network is made up of many layers and is mainly used for image processing, classification, and segmentation. The structure of CNN contains Convolutional, pooling, and fully connected layers. The figure below is shown how CNN works.

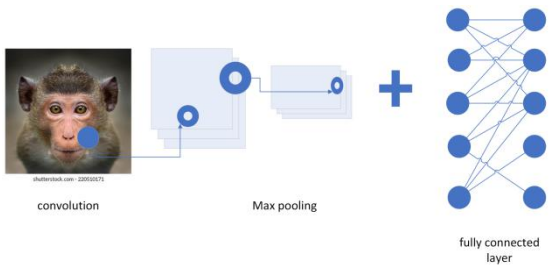


Figure : typical CNN structure

3.3.1 face detection

To recognize the person successfully in AR, the first step is to detect the face in the camera. In this step, we use the pre-rained face detection model provided by Opencv for unity [4].

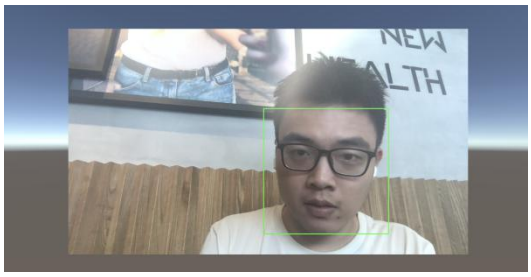


Figure 6: face detection demo

The above figure is the running result of our program in unity. It can be seen that the camera captures the face well and displays it in a green frame.

3.3.2 face search and recognition

When the face is detected in the AR device, the program will automatically save the screenshot to the local, extract the face features, encode them with Base64 coding, and upload them to Baidu AI cloud through socket connection. Baidu cloud stores our pre-trained model, and the user's face ID can be recognized after uploading the data.

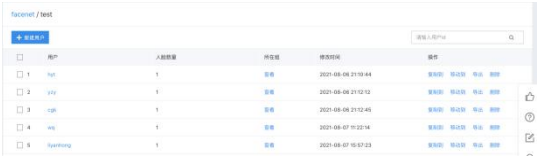
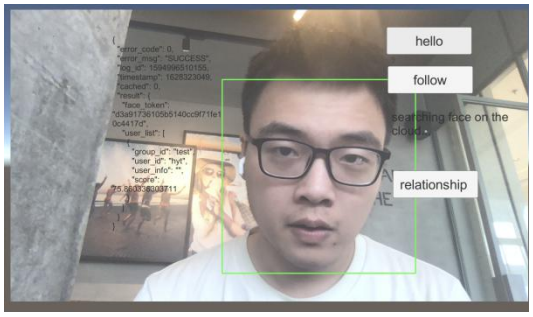


Figure 7: user-uploaded picture profile

3.4 message sending module

When another facenet user is detected by you, you can send it a request to establish a connection. When he passes, you can establish a socket connection through the facenet client to carry out a series of information exchange operations, such as sending information, voice call, etc.

3.5 software implementation



The above figure is a screenshot of the facenet software, in which the green rectangle shows the detected face, and the black font on the left is the face recognition result from the model deployed in the cloud. The buttons on the right show different functions. After clicking the relationship button, the social relationship between the user and the tested person will be found in the cloud user database, and the shortest of many social relationships will be returned and displayed through ar glass. Of course, you can also click the follow button to follow him on facenet.

3.4 augmented reality module

The augmented reality module in our application would focus on several points to enhance the feeling:

3.4.1 Remote Registering

To perform perfect padding around the reality and digital information, it requires a proper registering for both reality and network connection, and position tracking base on the context. By receiving the signal from the headset, our application can generate accurate 3D coordinates.

3.4.2 Vision Displaying

As AR stands for a vivid digital reality, seeing is the primary focus of our application, to bring up a close reality, a colorful and sensitive screen is needed. The headset would have camera-based equipment and non-camera-based equipment, to give as close as the visions customers expect to see by combining the digital information and the sight capture by the camera.

3.4.3 Digital Generating

Because AR's goal is to create a visual reality with not only seeing, but also touching and smelling, our application also implemented another feeling: touch. Using the information from remote registering and headset, our application will begin processing and generating a proper 3D model simulating the touching. To increase the efficiency of

processing, our application will use the visions from the camera and remote registering information to extract vital data, refine the quality of content, and brings up a feeling as close as you expect.

3.4.4 Data Transferring

In 21 century, non of the electronic devices won't connect to other applications such as Google, cloud, computer, etc. There are 3 different transferring for our application to use:

(2) Transferring base on postures: As a regular AR, our application will have "hotkeys" base on your postures, like right-hand slide right as "Enter", left-hand slide left as "Delete", to bring up a convenient operation and fasten your actions.

(3) Transferring base on specified application: AR needs connection, so why not use the connection to do other things like connecting to GPS to find out your coordinates? That is also included in our AR's functionality.

4. lessons learned and beyond

- **Software development technique**

We get familiar with android studio,unity and other powerful developing tools. Those tools are super useful in our later career.

- **Cloud platform**

Understand the use and principle of cloud platform, and successfully apply it to our project.. From the cloud, we can discover some brand new ways of combining and utilizing cloud techniques and assets with other new ideas, like multiple connections base on the dynamic cloud environment, using the dynamic cloud to save up the memory and volume in order to put more effort and focus on displaying high performance running for the more fantastic experience.

5. Conclusion

What we did in this paper is providing an application framework which can help people develop similar cloud based social app with ease.

- The results show that our face recognition model has an average classification accuracy of 97.81% when the number of samples is about 40.
- The introduction of augmented reality module can stimulate people's interest in using social software

6. Future work

6.1 optimization

Considering the limitation of mobile processor, there is a large optimization space in data structure storage and IO reading.

6.2 AR

The unique advantages of AR have not been well demonstrated in our project. We can also make more technological breakthroughs in AR，especially with the 5G technology.

6.3 Privacy

Privacy is bound to become an unavoidable topic. How to make users' privacy inviolable is extremely important for a software's popularity.

6.4 visualization

It would be cool to dynamically display the topology of social networks on the map

6.5 wider user range

Our test objects are relatively limited, so we need to extend facenet to a wider population

References

[1]Talk2Me: A Framework for Device–to–Device Augmented Reality Social Network

[2]Face Recognition Based on Convolutional Neural Network

[3]Six Degrees of Separation in Online Society .John Guare

[4] https://github.com/opencv/opencv

[5] https://www.mysql.com/

[6] Krizhevsky A, Sutskever I, Hinton G E. Imagenet classification with deep convolutional neural networks[C]//Advances in neural information processing systems. 2012: 1097-1105.

Appendix

1. link to our code

https://github.com/961853266hyt/ST8601-final-project