The Study and Application of Facial Recognition Models

Dihua Feng [1, †], Sixi Peng [2, †], Jiayong Wang [3, †, *]

[1] Big Data Application College, Zhuhai College of Science and Technology, Zhuhai, China

[2] Modern Information Industry College, Guangzhou College of Commerce, Guangzhou, China

[3] International Education College, Henan University, Zhengzhou, China

* Corresponding author email: WjyVanessa@henu.edu.cn

†These authors contributed equally.

Abstract. Facial recognition technology is a biometric application that uniquely identifies or authenticates a person by comparing and analyzing facial features based on the width of a person's face. Face recognition technology has gained popularity recently thanks to the development and invention of artificial intelligence and other technologies. The traditional methods of face recognition technology development—the method based on geometric features and the method based on 3D models—as well as the deep learning Convolutional Neural Network (CNN) model VGG-Face for face recognition—are analyzed, compared, and described in this paper. It also provides a detailed description of face recognition technology. In this work, accuracy and TAR values are employed for geometric feature matching by comparing the metric results of each method on a dataset for better comparison. Our results show that while both traditional geometric feature modeling methods and 3D modeling methods can achieve good accuracy, however, they do not perform as well as humans in terms of accuracy. The Vgg-Face based on the deep learning model completely exceeds the 3D model method in terms of accuracy and TAR value, and is even more accurate than humans.

Keywords: Geometric Features; 3D Model; VGG-Face; Face Recognition.

1. Introduction

Biometrics refers to the technology that uses human's physiological characteristics (e.g., fingerprint, face, iris) to recognize personal identity by the combination of computer technology and optics, biostatistics, etc. Face recognition technology has become one of the hottest study topics in biometrics due to the rapid growth of artificial intelligence and big data. Face recognition is based on the information of human facial features, and compared to other biological characteristics, facial features are easy to capture and understand [1]. Face recognition has been widely used in a variety of fields, including finance, justice, etc., due to its acceptance and ubiquity. Understanding the face recognition development process and identifying the accomplishments and shortcomings of each technical level are therefore essential.

One of the earliest approaches to facial recognition was based on geometric features. In 1965, Woody Bledsoe and Helen Chan Wolf published a technical report on automatic face recognition. They proposed that before computers recognized faces, facial feature coordinates must be first manually establish for recognition [2]. In 1971, the strategy to cut down on feature points was developed by Goldstein et al. They graded features and calculated the mean and variance of the data samples to choose the most reliable 22 features [3]. Several years later, some approaches to construct feature vector by selecting feature points automatically were developed. In 1977, Leon D. Harmon and Willard F. Hunt designed an algorithm that allowed to define a unique description vector from the input contour paths [4]. Additionally, the way to extract facial features in layers and areas was measured by Hiroshi Sako and Anthony V. W. Smith in 1996 [5]. To deal with strict requirements on face images, Yuille et al. raised a method of using a parameter-adjustable variable model to represent facial features [6].

The emergence of 3D model based on face recognition methods overcame the problems of insufficient accuracy of traditional 2D face recognition and pose, lighting and other problems. The key of this method is to reconstruct 3D face model through a certain algorithm. Among them, the popular 3D face reconstruction methods are CANDIDE-3 and 3D Deformable Model (3DMM). This article will mainly introduce 3DMM. In 1999, Blanz and Vetter proposed 3DMM algorithm, which is based on a 3D face database and generates new models from existing models in the database. In 2014, a facial expression database was published, which improved the robustness of 3DMM algorithm when facing different expressions. In 2017, Booth, James and others proposed LSFM. It is by far the largest 3DMM, and its database collects most of the global face information.

The most promising and popular method today is Convolutional Neural Network (CNN)-based face recognition. As early as 1989, there was a CNN model for image recognition, LeCun et al. proposed the LeNet model, but limited by the amount of data and hardware performance at that time, it did not achieve the desired performance. It was not until 2012 that thanks to the development of hardware, CNN broke through and the AlexNet model appeared. And in the next 10 years, more excellent models appeared e.g. VGGNet, GoogleNet, ResNet.. In 2014, Facebook proposed DeepFace that is based on AlexNet, the first model to apply CNN to face recognition in CVPR. In 2015, the Googlenet-based FaceNet model and VGGNet-based VGGFace model were proposed successively.

The purpose of this study is to examine the current progress and main technology of face recognition in last decades. The remaining sections of the paper are structured as follows. We will introduce the use of face recognition and its corresponding methods in section 2. Then, three main face recognition methods based on geometric features, 3D models, and the deep learning model will be researched and discussed in general in section 3. Also, several evaluation metrics will be presented in this part. Section 4 will offer several compared experiment data results ,and finally, section 5 will summarize the paper and draw conclusions from the discussion of developing processes of three methods.

2. Application

2.1 Application in Security Field and B-End Market

Policy decisions have an impact on face recognition technology. Face recognition technology is supported by the government due to the recent rapid growth of society and the economy, which has caused the technology to advance ever more quickly. The expression of human face is gradually developed by technology to the direction of commercialization [7]. For example, China's financial, government, education, medical and other enterprises and institutions. Among them, the most widely used fields of face recognition technology are finance and security. In addition, based on the continuous updating of technology, China's public transportation industry has gradually introduced face recognition technology. For example, in the Spring Festival in 2017, major transit stations have established passenger passages with face-scanning progress. This has prompted the widespread adoption of facial recognition technology in the transportation industry.

2.2 Application of Smartphone and C-end Market

The primary application areas of the C-end market are payment brushing and face unlocking. Among these, the use of face recognition technology in smartphones serves as the primary example of this technology's use in the lower end of the market. The 2017 release of the iphoneX serves as a precursor to the use of face recognition technology in smartphones. After that, the use of face recognition technology in Android phones will be entirely available under the application of FacelD. Needs for 3D facial recognition technology Due to the high requirements, 3D face recognition is not commonly employed in practical mobile phones. Sufficient hardware requirements and significant chip computation power are also needed. Additionally, while 2D face recognition technology has evolved to the point where it can be found in low-end mobile phones, 3D face recognition technology has encouraged the development of this technology as well.

2.3 Intelligent Office System

The smart workplace has drawn a lot of interest and is unquestionably going to be a trend in future societal development as the application possibilities of artificial intelligence are constantly expanded. Intelligent management is achieved through the use of face recognition software on a cloud platform, face equipment on mobile terminals, visitor management systems, access control attendance systems, office control systems, employee recognition systems, and other recognition systems. This allows for the professional creation of a more humanized and scientific environment, which boosts office productivity and quality. Face recognition technology has been extensively employed and played a significant influence in many other fields in addition to the ones mentioned above.

3. Methods

3.1 Geometrical Feature Matching

The geometric feature-based face recognition technique is realized by detecting the feature vector matching between the input image and the template picture. Especially, feature vectors are constructed by the shape and geometric relationship of facial organs. In 1965, Bledsoe et al. constructed a database concerning distance by manually establishing facial feature coordinates and calculating the distance of different features (e.g., pupil center, nose width). Thus, when they entered a new photo, the computer would use the previously computed database to compare the feature distances in each photo, returning possible-matched records of for identification [2]. The method of manually building feature databases was also used by Goldstein et al. in 1971. The key step in their experiment was to classify the feature points (e.g., setting level 1-5 based on lip width), and then analyze the assigned data samples [3].

Nevertheless, the methods of selecting feature points automatically gradually replaced the manual ones around 1980s. Leon et al. proposed an algorithm for automatic positioning of facial feature points by using the database constructed by the profile curve of the side face photo [4]. Subsequently, Sako et al. improved the feature extraction, and adopted the method of partition and stratification when extracting face features in real time. For example, they used Color Matching (CM) for feature extraction of face and mouth regions, where the value of the color pixel p was stored as $Co \times M(j)$ (Co is a constant, $M(j)$ is the probability that the pixel p belongs to a pre-determined normalized color histogram of the model face image). Additionally, Template Matching (TM) was used for the feature extraction of the eye region. For instance, the position of the eye was extracted by using the pixel density distribution of the eye shape in the image space. Moreover, Sako et al. added local image analysis in the eyebrow pupil region. The feature extraction rates of face, mouth and eye obtained by the above methods were 100 %, 96 % and 90 % respectively, which proved that the methods were effective [5].

However, the above methods have strict requirements on face images. For example, face images must be frontal face images and cannot be deformed or rotated. To address these issues, Yuille et al. offered a methodology to represent facial features by a parameter-adjustable variable model. The method firstly constructed the corresponding energy function through the feature parameter model with variable parameter values, and then found the minimum value of the energy function using the gradient descent method. The experiment verified that the parameter values in the energy function were the best matching values of the features. Therefore, the model parameters would be used as the geometric features of the organ [6].

3.2 3D Model

Face recognition based on 3D models requires 3D reconstruction of faces before recognition. There are many 3D reconstruction methods based on face images, including 3D Morphable models and CANDIDE-3. This article will focus on the 3D Morphable models proposed by Blanz et al. in 1999 [8].

The general 3D face model is a 3D deformable model. It can use multiple fixed points (x, y, z) to display the face. Its central idea is that the 3D face model can be fitted with the real face through these points (x, y, z), and can be linearly added through the weighted orthogonal basis of many other faces. Therefore, each 3D face model can be composed of all the base vector coefficients representing the face in the data set. The establishment of any 3D face model is actually a rearrangement and combination of the coefficients of each base vector.

In the research of 3D face, the basic features of face can be divided into shape features and texture features as shown in Fig. 1 and Fig. 2, so 3D face modeling can be represented by linear superposition of shape feature vector and texture feature vector.

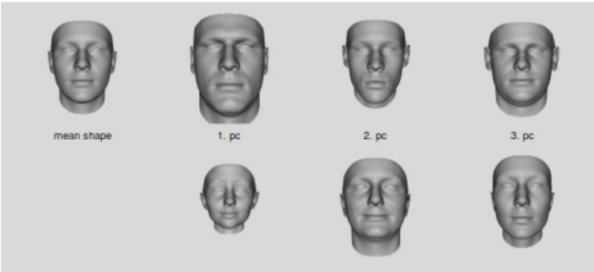

Fig 1. Shape vector based on the face [9], the Shape Vector S=(X1,Y1,Z1,X2,Y2,Z2,...,Yn,Zn).

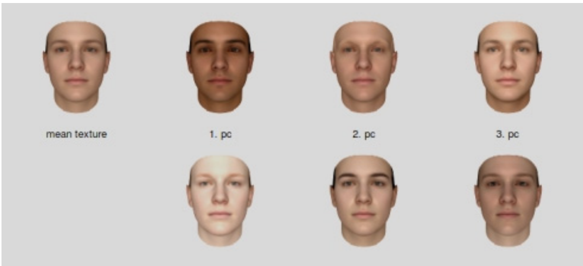

Fig 2. Texture Vector [9], the Texture Vector T=(R1,G1,B1,R2,G2,B2,...,Rn,Bn).

Therefore, any facial model can be fitted by numerical weighting and permutation combination method through the eigenvector value of m personal facial model in the data set, as shown below:

$$S_{mod} = \sum_{i=1}^{m} a_i S_i \ , \quad T_{mod} = \sum_{i=1}^{m} b_i T_i \ , \quad \sum_{i=1}^{m} a_i = \sum_{i=1}^{m} b_i = 1 \ . \tag{1}$$

S_i, T_i are the shape feature vector and texture feature vector of the ith personal face model in the dataset, respectively.

Since they are not orthogonally correlated, the "Si" and "Ti" cannot be used directly, as the basis vector when building the model. In this case, Principal Component Analysis (PCA) shown in Fig. 3 was considered for dimension reduction decomposition.

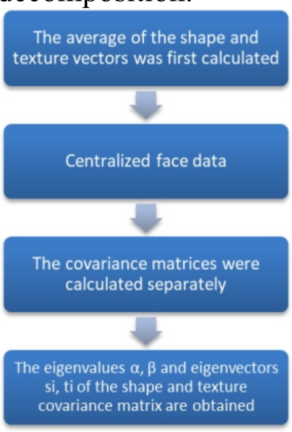

Fig 3. Principle Component Analysis (PCA).

Thus, the formula1 can be converted to formula (2).

$$S_{model} = \bar{S} + \sum_{i=1}^{M-1} \alpha_i s_i \,, \quad T_{model} = \bar{T} + \sum_{i=1}^{M-1} \beta_i s_i \tag{2}$$

When 3DMM is used for face modeling, a dataset with facial texture feature vectors and shape feature vectors is essential. Unfortunately, Blanz et al. only proposed the feature vector extraction method in 1999, but they did not open source their datasets. In 2009, Pascal Paysan and others used laser scanners to accurately capture the data of 200 faces, and opened the Basel Face Model (BFM) dataset they extracted, as shown in Fig. 4 [10].

Basic information of BFM database:

They collected with ABW-3D structured light system. The entire data set is composed of 200 3D face models, 100 of which are male, 100 of which are female, and most of them are white faces. The average age is 24.97 years old, ranging from 8 to 62 years old. Everyone was asked to collect 3 natural expressions and choose the best one.

In the process of point location fitting, the position of each point of the model should be accurately fitted. In the process, each model is fitted by 53490 points. And the expression coefficient was added to the BFM2017(Fig. 5) dataset in 2017[11]. In the same year, Booth et al. proposed the Large-Scale Face Model (LSFM) (Fig. 6) [12], which is a 3D Deformable Model (3DMM) automatically constructed by 9663 faces from different regions and ethnic groups in the world. According to them, LSFM is by far the largest deformable model, which basically contains face feature data of all races in the world.

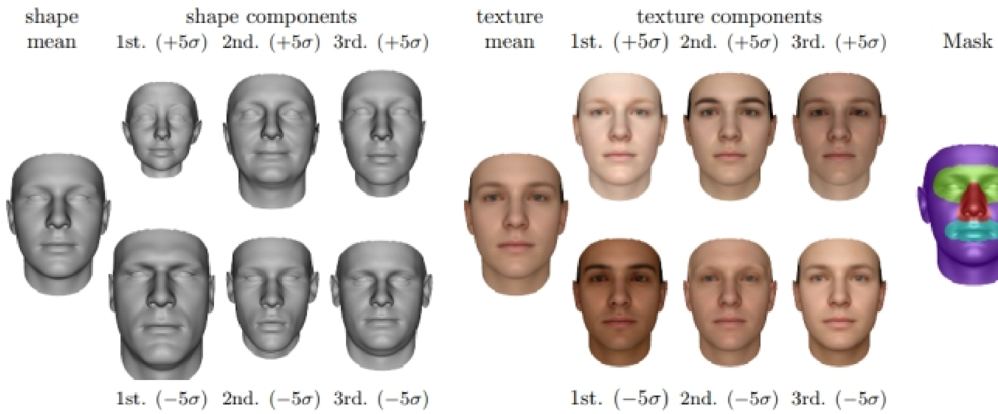

Fig 4. Average Face Shape and Average Face Texture in BFM Database [10].

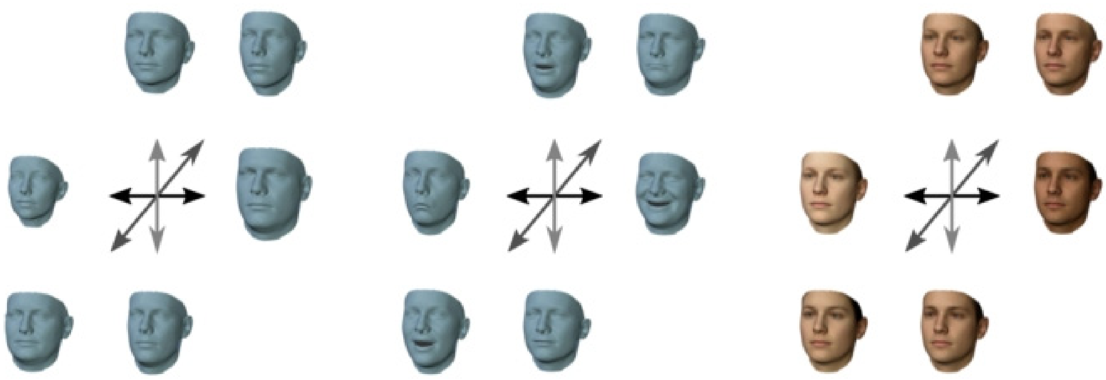

Fig 5. BFM 2017 [11].

Fig 6. LSFM [12].

3.3 The Deep Learning Model

Deep learning model is also proposed in the last decade to carry out the face recognition task. The VGGF model, one of the typical deep learning-based models — Oxford Visual Geometry Group created it [13]. The model was developed using a sizable dataset that included 2.6 million facial photos of over 2.6 thousand people. From the input layer to the output layer, there are 38 layers in the VGGF design. An average is typically calculated from the input image as part of the pre-processing step. The input should be a color image with a size of 224 by 224.

The VGGF generally consists of thirteen convolutional layers, each of which has a unique set of hybrid parameters. There are 15 Rectified Linear Units and 5 maxpooling layers in each group of convolutional layers (ReLUs). Three completely connected levels—the FC6, FC7, and FC8—follow these layers. The first two have 4096 channels each, while FC8 is used to categorize the 2622 identities and has 2622 channels. The classifier, a SoftMax layer that determines the particular face class that an image belongs to, is the final layer.

3.4 Evaluation Metrics

The performance of facial recognition methods is mainly measured by Precision (Eq.1), Recall (Eq. 2), and Accuracy (Eq. 3). True Accept Rate (TAR) is also used to evaluate the performance of recognition.

$$\text{Precision} = \frac{TP}{TP+FP} \tag{3}$$

$$\text{Recall} = \frac{TP}{TP+FN} \tag{4}$$

$$\text{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN} \tag{5}$$

Whereas TP stands for True Positive, TN stands for True Negative. Additionally, FP represents False Positive, and FN represents False Negative.

4. Results and Discussion

4.1 Results

We compare the metric results of each method on the according dataset in the Table 1 above. The precision and accuracy values have been adopted for Geometrical Feature Matching, while the accuracy and TAR values are used for better comparison. All results are from their original experiments.

Table 1. The performance comparison based on three Methods.

Method	Dataset	Precision	Recall	Accuracy	TAR-10%
Geometrical Feature Matching	ImageFrame	85.40%	-	70.78%	-
3D Model(3DMM)	MICC	-	-	75.25%	59.4%
Deep Learning Model (VGGface)	LFW	-	-	98.78%	99%

By comparing the experimental results of several models in the table, it can be found that the traditional geometric feature model method and the three-dimensional model method can achieve good accuracy, but they still did not perform as well as humans in terms of accuracy. However, the VGGface method based on deep learning models completely surpasses the 3D model method in accuracy and TAR value, which proves the superiority of neural networks.

4.2 Discussion

From the comparison of the above methods, it can be observed that the deep learning method is the main research direction of face recognition in the future. Although the face recognition method based on deep learning can win over any traditional face recognition method, it does not mean that the traditional method is useless. One of the characteristics of deep learning is its applicability. As long as there are enough network layers for deep learning, any function can be mapped theoretically, so the deep learning method can combine and optimize many excellent traditional algorithms. For example, most of today's face recognition methods based on deep learning are 2D fitting 2D, so the fitting process is inevitably affected by light and angle. Therefore, deep learning can be combined and optimized with 3D model method. For example, Prnet is a deep learning method that combine the 3D method [14]. They used deep learning to solve the difficulty of 3D method in collecting 3D face models. 3D face recognition with better physical performance solves the complicated process. Therefore, we think that deep learning algorithm and other traditional methods are indispensable in the future development of face recognition.

5. Conclusion

In this study, we reviewed the application, principal methods and corresponding performance based on the facial recognition task. Facial recognition methods can be broadly divided into three parts, namely geometrical feature matching, 3D model and deep learning-based models. Additionally, we merely present the possibilities of CNNs in this study. There are a variety of additional approaches that could be utilized in the future to create and evaluate facial recognition cues. General Adversarial Networks (GANs), which are becoming more and more popular as machine learning tools and which only need a small amount of data to train neural networks, are one such technique.

References

[1] Senior, A.W., Bolle, R.M. Face Recognition and its Application. In: Zhang, D. (eds) Biometric Solutions. The Springer International Series in Engineering and Computer Science, 2002, 697: 83-97.

[2] Woodrow Wilson Bledsoe. A Facial Recognition Project Report. Panaromic Research, Inc, 1963.

[3] Goldstein, A.J., Harmon, L.D., Lesk, A.B. Identification of human faces. Proceedings of the IEEE, 1971, 59(5): 748–760.

[4] Leon D. Harmon, Willard F. Hunt. Automatic recognition of human face profiles. Computer Graphics and Image Processing, 1977, 6(2): 135–156.

[5] Sako, H., Smith, A.V.W. Real-time facial expression recognition based on features' positions and dimensions. Proceedings of 13th International Conference on Pattern Recognition, 1996, 3: 643–648.

[6] A. L. Yuille et al. Feature extraction from faces using deformable templates. Proceedings CVPR '89: IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 1989: 104-109.

[7] Li Guoxin. Development and Application of Face Recognition Technology. Science and Technology Communication, 2018, (19): 102-103.

[8] Blanz V, Vetter T. A morphable model for the synthesis of 3D faces. Proceedings of the 26th Annual Conference on Computer Graphics and Interactive Techniques, 1999: 187-194.

[9] Morphface. Basel Face Model, 2009, https://faces.dmi.unibas.ch/bfm/index.php?nav=1-1-0&id=details.

[10] P. Paysan et al. A 3D Face Model for Pose and Illumination Invariant Face Recognition. 2009 Sixth IEEE International Conference on Advanced Video and Signal Based Surveillance, 2009: 296-301.

[11] T. Gerig et al. Morphable Face Models - An Open Framework. 2018 13th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2018), 2018: 75-82.

[12] Booth J, Roussos A, Ponniah A, et al. large scale 3d morphable models. International Journal of Computer Vision, 2018, 126(2): 233-254.

[13] Parkhi, O.M., Vedaldi, A. and Zisserman, A. Deep Face Recognition. Proceedings of the British Machine Vision Conference (BMVC), 2015, 1(3): 6.

[14] Feng Y, Wu F, Shao X, et al. Joint 3D Face Reconstruction and Dense Alignment with Position Map Regression Network. European Conference on Computer Vision, 2018.