International Core Journal of Engineering Volume 8 Issue 6, 2022
ISSN: 2414-1895 DOI: 10.6919/ICJE.202206_8(6).0019

Face Expression Recognition Algorithm based on Image

Yali Yu, Hua Huo*

Medical Big Data and Computational Intelligence Engineering Technology Research Center, Henan University of Science and Technology, Luoyang 471003, China

Abstract

This paper mainly expounds the basic knowledge of facial expression recognition. Firstly, it introduces the four main stages of facial expression recognition and the work tasks of each stage, then introduces the traditional feature extraction algorithm based on frequency domain and local features and the commonly used expression feature classification method, and analyzes the accuracy and effective performance of expression feature extraction to improve the effect of facial expression recognition. At the same time, two kinds of facial expression datasets and the difficulties of facial expression recognition are introduced.

Keywords

Expression Recognition; Feature Extraction; Feature Classification.

1. Introduction

With the development and progress of society and the spread of COVID-19 in recent years, the application of facial expression recognition is more and more extensive, and more and more scholars have done more in-depth research. During the epidemic, students are isolated at home, and teachers teach through remote classroom. Facial expression recognition technology can assist teachers to judge whether students pay attention in class and whether students understand what is taught in class [1]; In terms of human-computer interaction, especially in this society where left-behind children and empty-nesters are becoming more and more serious, human-computer interaction technology can accompany and care for the elderly and children to a certain extent, alleviate loneliness and make up for emotional loss in life; In terms of healthcare, COVID-19 has spread at home and abroad in recent years, seriously affecting people's life and work. In the treatment of patients, facial expression recognition can help medical staff analyze the changes in the condition, and timely find the patients' discomfort and medication effect [2]; In terms of traffic safety, image acquisition device can be installed in the driving position to detect the driver's state in real time, analyze whether the driver is emotionally stable, whether he is tired driving and whether he is physically unwell due to sudden diseases through expression recognition, and set up an alarm system to remind the driver to adjust his mood or stop driving for rest in time. This can not only reduce the occurrence of traffic accidents, but also improve the quality of life and promote social harmony and stability. Therefore, with the improvement of the practical application value of facial expression recognition in daily life and its important theoretical research value, it has become a hot research direction [3].

In 1971, American psychologists Ekman and Friesen made pioneering work on modern facial expression recognition [4], dividing facial expressions into six basic expressions, namely happy, sad, surprise, fear, anger and disgust. In facial expression recognition, the expression forms of different motion characteristics of facial muscles express different emotions, such as frowning tightly, vertical wrinkles between eyebrows, wide eyes, wide nostrils, and closed lips to express anger; The eyebrows are slightly bent down, there are wrinkles and bulges under the lower eyelids, but they are not nervous, the corners of the lips are pulled back and raised, the mouth is opened, and the teeth are exposed to

ISSN: 2414-1895 DOI: 10.6919/ICJE.202206_8(6).0019

express happiness. At the same time, the contribution of different feature parts of the face to expression recognition is also different. For example, the importance of the eyes is greater than the nose, and the importance of the mouth is greater than the eyebrows. It can be seen that the accuracy and effectiveness of expression feature extraction directly affect the accuracy of expression classification. Luo et al. used the combination of Principal Component Analysis (PCA) and Local Binary Patterns (LBP) to extract global gray features of the face [5], and improved facial expression classification through Support Vector Machines (SVM). This experimental method can effectively improve the accuracy of expression recognition. Zou et al. proposed a lightweight convolutional neural network (MFF-CNN) using the multi-feature fusion method [6]. The network model extracted the mid-level and high-level global features of each image. Experiments were conducted in CK+, JAFFE and Oulu CASIA datasets and comparison experiments with other models show that the network model improves the average recognition accuracy. Li et al. segmented the region of interest for expression recognition in face images [7], and used depthwise separable convolution to establish classifiers, and used a fine-tuning method to improve the expression recognition rate through experimental training. Zhang et al. extracted the shallow features of facial expressions through the convolution layer, and used the Inception network module as the basis and dilated convolution parallel to extract the multi-scale features of facial expressions [8]. At the same time, the attention mechanism was introduced. Compared with other traditional feature extraction algorithms, the method improves the recognition rate.

2. Basic Knowledge of Facial Expression Recognition

2.1 Facial Expression Recognition Process

Facial expression recognition mainly includes four stages. The first stage is face detection, which detects whether there is a face in the image and marks the face part with a rectangular frame; The second stage is image preprocessing, which mainly processes the image information, reduces the influence of illumination and noise, and removes the redundant information of the image, so as to prepare for the next step of extracting effective features; The third stage is feature extraction, which extracts the characteristic information in the image. This step is the key and has a great impact on the accuracy of expression recognition; The fourth stage is feature classification, which maps the recognized expression results to the corresponding labels, which affects the final accuracy of facial expression recognition. The basic flowchart of facial expression recognition is shown in Fig.1.

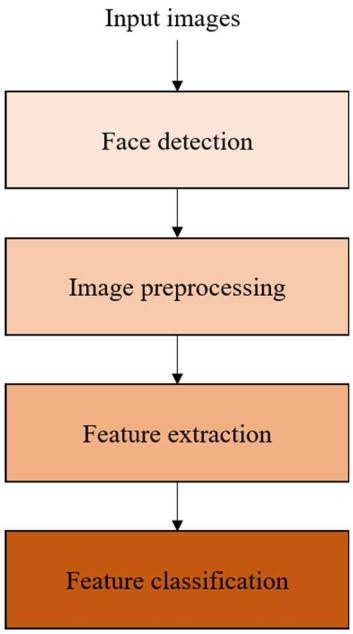

Fig. 1 Facial Expression Recognition Flowchart

ISSN: 2414-1895 DOI: 10.6919/ICJE.202206_8(6).0019

2.2 Expression Feature Extraction Algorithm

Image feature extraction refers to the processing and analysis of the information with obvious physical or statistical significance contained in the image. The extracted effective features usually have the characteristics of distinguishability, stability and independence.

2.2.1 Gabor Wavelet Transform

Wavelet transform is a local transform of time and frequency, which can extract relevant features in different scales and directions in the frequency domain, overcoming the defect that Fourier analysis can only describe signals with a single variable [9]. Gabor wavelet transform has the advantages of zoom, covering a wide frequency domain and weakening the correlation between different features. Performing Gabor wavelet transform on an image and then performing dimensionality reduction processing can improve the computational efficiency.

Wavelet transform can not only extract the features of image texture, but also reduce the interference of illumination and position on image recognition. The process of wavelet denoising is to decompose the noisy signal in multi-scale first, transform it from time domain to wavelet domain, then extract wavelet coefficients in each scale, and finally reconstruct the signal through inverse wavelet transform.

2.2.2 Scale-invariant Feature Transform(SIFT)

Scale-invariant feature transformation sift was first published at the International Conference on Computer Vision (ICCV). It is an algorithm in the field of image processing that detects local features in the image. It is invariant to rotation and scale scaling, and has certain stability in noise, illumination, angle change and affine transformation. The SIFT algorithm mainly finds key points in different scale spaces, and then determines its location, scale and gradient direction.

2.3 Expression Feature Extraction Algorithm

After the above feature extraction, it is necessary to further classify the facial expression features, and selecting an appropriate classifier according to different feature requirements can improve the effect of facial expression recognition.

2.3.1 Support Vector Machine (SVM)

Support vector machine has achieved remarkable results in text classification tasks, and quickly became the mainstream technology in machine learning. SVM is a classifier based on linear discriminant function, which is generally suitable for binary classification problems. The maximum margin hyperplane is used as the decision surface of classification. For multi- classification problems, it is necessary to combine multiple binary classifiers to construct appropriate multi-classifiers. For the linear inseparable problem, the low dimensional feature space can be mapped to the higher dimensional Hilbert space by nonlinear function, which can be transformed into a linear separable problem.

The construction of SVM multi-classification mainly includes direct method and indirect method. The direct method has high complexity and difficult implementation process, which is more suitable for small sample classification. The indirect method is to first construct a hyperplane to separate one class from other classes, and then construct a hyperplane to classify one class and other classes in the subclass, and cycle the above operations in turn until all classes are separated separately. SVM multi-classification can solve the machine learning problem of small sample data, and can be applied to data mining, pattern recognition, etc. It has the advantages of good learning ability and low generalization error rate.

2.3.2 Sparse Representation Classification (SRC)

In 1993, Mallat proposed a method to represent signals with a super-complete dictionary, and suitable super-complete dictionaries could be selected according to the characteristics of signals, laying a foundation for the appearance of sparse representation. Studies have found that after sparse representation of signals, the more sparse the signal is, the higher the accuracy of the signal [10].

ISSN: 2414-1895 DOI: 10.6919/ICJE.202206_8(6).0019

Signal sparse representation is mainly divided into two tasks: dictionary generation and signal sparse decomposition. Generally, the selected dictionaries include analysis dictionary and learning dictionary. Although the analysis dictionary is easy to implement sparse representation of signals, the disadvantage is that the adaptive ability is poor and the representation form is single. However, learning dictionary can well adapt to different types of image data and has strong adaptive ability. Sparse representation has low accuracy for posture changes and uneven in images, but it is still suitable for illumination changes, image occlusion, corrosion, etc.

3. Facial Expression Database

3.1 JAFFE Database

The JAFFE database (The Japanses Female Facial Expression Database) selects 10 Japanese women's facial expression pictures taken in a laboratory environment [11], and each person makes 7 expressions according to the instructions, and these 7 expressions are Angry, Disgust, Fear, Happy, Sad, Surprised, Neutral. The database has a total of 213 face expression images, the size of the images is 256×256 pixels, and they are stored in TIFF format without compression. The database collection environment is fixed and single, the expression is relatively standardized, and has achieved good results in some traditional algorithms. However, due to the small number of images, and the images have no occlusion, which is not close to the scene of real life, it is not suitable for the deep neural network model, which is not conducive to training the robustness of the model.

3.2 RaFD Database

The RaFD database (Radboud Faces Database), created by graduate students in Behavioral sciences at Nijmegen in the Netherlands [12], is a high-quality face database containing 67 objects. The database contains a total of 8040 facial expression images with emotional expression, including eight expressions: Anger, disgust, fear, Happy, Sad, Surprised, Contempt and Neutral. Each expression contains 3 different gaze directions and uses 5 cameras to capture the face from different angles.

4. Current Difficulties in Facial Expression Recognition

With the rapid development of facial expression recognition technology, researchers at home and abroad have made some achievements. However, due to the diversity and complexity of facial expressions, and involving physiology and psychology, the accuracy of expression recognition needs to be further improved. There are still some difficulties to be solved in the existing facial expression recognition technology. After analysis, the main research difficulties are summarized as follows:

1) In the process of image collection, due to the different collection equipment and surrounding environment, it is easy to lead to problems such as high image noise, low resolution and serious occlusion, which affects the extraction of effective features.

2) Due to the limited amount of data available at home and abroad, it is easy to cause over-fitting when training neural network models, and it is also easy to lead to the problems of poor robustness and weak generalization ability of the trained network models. At present, strengthening the scale of labeled databases is the primary problem to be solved, especially for the popular deep neural network model.

3) Due to the differences in the facial structure of people in different countries and regions, different ages, genders, skin colors, races, etc. will have an impact on expression recognition, so it is difficult to build a unified neural network model for facial expression recognition, which will affect the accuracy and generalization ability of expression recognition.

5. Summary

Facial expression recognition has great potential application value in psychology, health care, safe driving and other fields, and is a very popular research topic in the field of image processing. At present, researchers in different fields have proposed different facial expression recognition

ISSN: 2414-1895 DOI: 10.6919/ICJE.202206_8(6).0019

algorithms and built different neural network models, which has achieved good recognition results and promoted the development of facial expression recognition technology. However, due to the limited number of images in the existing public database and can not fully reach the actual scene of real life, improving the robustness and generalization of network model is still the focus of future research.

References

[1] Yongfan L, Licheng L. Application of affective computing in web-based distance education system: Functions, Current research and key problems [J]. Modern Distance Education Research. Vol. 3 (2013) No. 10, p. 32-35.

[2] Ma Lin,Chen Wenfeng,Fu Xiaolan,Wang Tongtong. Emotional expression and micro-expression recognition in depressive patients[J]. Chinese Science Bulletin. Vol. 63 (2018) No. 20, p. 2048-2056.

[3] Barros P, Churamani N, Sciutti A. The facechannel: a fast and furious deep neural network for facial expression recognition[J]. SN Computer Science. Vol. 1 (2020) No. 6, p. 1-10.

[4] Ekman P, Friesen W V. Constants across cultures in the face and emotion[J]. Journal of Personality and Social Psychology. Vol. 17 (1971) No. 2, p. 124.

[5] Luo Y , Wu C M , Zhang Y . Facial expression recognition based on fusion feature of PCA and LBP with SVM[J]. Optik - International Journal for Light and Electron Optics. Vol. 124 (2013) No. 17, p. 2767-2770.

[6] Zou W, Zhang D, Lee D J. A new multi-feature fusion based convolutional neural network for facial expression recognition[J]. Applied Intelligence. Vol. 52 (2022) No. 3, p. 2918-2929.

[7] Li Chunhong, Lu Yu. Facial expression recognition based on depth separable convolution [J] Computer Engineering and Design. Vol. 42 (2021) No. 5, p. 7.

[8] Zhang Peng, Kong Weiwei, Teng Jinbao. Facial expression recognition based on multi-scale feature attention mechanism[J] Computer Engineering and Application. Vol. 58 (2022) No. 1, p. 8.

[9] Zhang Jingyuan, Zhang Bing, Jiang Xingzhou. Analysis of feature extraction method based on wavelet transform [J] Signal Processing. Vol. 16 (2000) No. 2, p. 156-162, 155.

[10] Luan Sidao, Wang Weiwei, Xie Yuxiang, et al Nonlinear sparse representation theory and its application [J] Computer science. Vol. 41 (2014) No. 8, p. 13-18.

[11] Lyons M J, Kamachi M, Gyoba J. The Japanese female facial expression (JAFFE) database[C]. Proceedings of Third International Conference on Automatic Face and Gesture Recognition. 1998, p. 14-16.

[12] Langner O, Dotsch R, Bijlstra G, et al. Presentation and validation of the Radboud Faces Database[J]. Cognition and Emotion. Vol. 24 (2010) No. 8, p. 1377-1388.