Intelligent Emotion Evaluation Method of Classroom Teaching Based on Expression Recognition

Yanqiu Liang
Nanyang Institute of Technology Normal College, Nanyang, China
`rmgyrrezptpv7@163.com`

Abstract—To solve the problem of emotional loss in teaching and improve the teaching effect, an intelligent teaching method based on facial expression recognition was studied. The traditional active shape model (ASM) was improved to extract facial feature points. Facial expression was identified by using the geometric features of facial features and support vector machine (SVM). In the expression recognition process, facial geometry and SVM methods were used to generate expression classifiers. Results showed that the SVM method based on the geometric characteristics of facial feature points effectively realized the automatic recognition of facial expressions. Therefore, the automatic classification of facial expressions is realized, and the problem of emotional deficiency in intelligent teaching is effectively solved.

Keywords—Intelligent teaching, emotion recognition, support vector machine (SVM)

1 Introduction

With the development of information intelligence technology, artificial intelligence education faces challenges and opportunities. Multimedia computers are widely used in the field of education, which has a great impact on the traditional teaching process. A variety of emerging advanced teaching equipment have entered the classroom. The teaching form is developing in a diversified direction. The intelligent computer-aided teaching system is integrated with network, artificial intelligence and multimedia technology. It differs from traditional computer-aided teaching systems. The distinguishing feature is its intelligent and personalized teaching function, which has the advantages of interactivity, sharing, autonomy and efficiency. In the process of learning, human-computer interaction is realized. Teachers and students can realize the interaction between teaching and learning through the network. Based on the geometric features of the image, commonly used linear kernel functions, polynomial kernel functions, and radial kernel functions are applied. The issue of kernel function selection and parameter optimization in SVM (support vector machine) is further explored. Combined with the geometric features based on facial features, the SVM method is used to identify facial expressions. Finally, automatic classification of facial expressions is achieved. The experimental results were analyzed and compared.

2   State of the Art

With the continuous deepening of education informatization, the demand for intelligent teaching products in the education industry has grown rapidly. Intelligent computer-aided instruction was used in the field of medical education. As a new type of teaching media, virtual classrooms were the product of computer, multimedia, and network communication and other information technology, multi-disciplinary and multi-field integration. Through computer networks, multimedia communication technology was used to construct the learning environment, which can realize virtual classroom teaching monitoring, multimedia individualized interactive network learning and other functions. Teachers and students in different locations can carry out most teaching activities. Neoh et al. [1] studied the intelligent recognition of facial expressions based on hierarchical coding cascade optimization model. Liu et al. [2] discussed several designs for face-to-face digital classroom environments. Elżbieta et al. [3] studied the interaction between teachers and students in the digital interactive classroom to expand the teaching space of teachers. Its research and development were also combined with the publication of digital books, and teaching books were digitized. The system was being promoted on campus across the country. Lee and Shin [4] studied the virtual teacher software system. It creates a more user-friendly, flexible and smarter online learning system for users. In the Global Chinese Computer Education Application Conference, Wei and Zhao [5] proposed to introduce the role of virtual teacher or virtual instructor in the learning system. Based on the concept of mobile learning and hybrid learning, Ko [6] used a personal digital assistant (PDA) for teaching. Test and note functions were added when the student's handheld device was connected to the teacher's PC. At the same time, the application supports a single user to use the PDA as an external input and output device. PAD was used as a front-end device for personalization and intelligence. Other applications or device interfaces were the future development direction, which was also a research hotspot of intelligent teaching. Shah et al. [7] established the candle project. Based on the digital community of the network learning integrated environment, mobile technology was used to provide content and communication. This brings hope to information education in remote areas. Shojaeilangari et al. [8] studied dynamic facial expression analysis based on gradient direction extended space-time histogram. Based on the Japan Female Facial Expression (JAFFE) Database and grayscale features, the relationship between kernel function and feature quantity and expression recognition accuracy and time consumption was discussed. Combined with the experimental results, the relevant guiding principles were proposed for the selection of the kernel function and the selection of the number of features.

In summary, traditional teaching methods cannot meet current needs. Information technology such as computers, multimedia, and network communications were introduced into classroom teaching activities. However, previous studies have not applied expression recognition technology to the classroom. The student's emotions were ignored. Therefore, to solve the problem of emotional loss in teaching and improve the teaching effect, an intelligent teaching method based on facial expression recognition was studied. First, the research background and meaning were introduced. Sec-

ond, previous research findings and theories in this field were reviewed. Third, the facial expression data set was introduced. The issue of kernel function selection and parameter optimization in SVM was discussed. Finally, the proposed method was verified. Results showed that the SVM method based on the geometric characteristics of facial feature points effectively realized the automatic recognition of facial expressions.

3 Methodology

3.1 Emotional calculation

In 1997, Professor R. Picard of the Massachusetts Institute of Technology Media Lab defined the concept of emotional computing in her monograph "Affective Computing." Emotional computing is a calculation that derives from emotions. The purpose of emotional computing is to establish a harmonious human-machine environment. By giving computers the ability to recognize, understand, express, and adapt to human emotions, computers have higher and more comprehensive intelligence. Emotional computing theory and technology have been widely concerned by academic circles at home and abroad. It is one of the most challenging scientific issues in the field of harmonious human-computer interaction research. The emotion calculation includes several parts as shown in Figure 1.
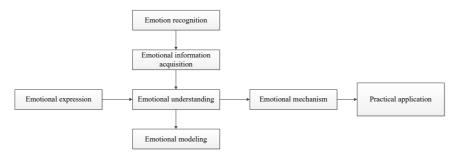

Fig. 1. Research content of emotional computing

3.2 Key technology of emotional computing

In the field of emotional computing, there are several key technologies, including the study of emotional signal sensors, biometrics, analysis of human emotion states based on expression features and physiological signals, emotional modeling and recognition, the effective expression of identified emotional outcomes, the fusion, integration, and knowledge reasoning systems of various perceptual data.

Psychology theory believes that emotion as a psychological process has a unique external manifestation. Expressions include facial expressions, gesture expressions, and tone expressions. In the process of emotional expression, psychologists have shown that facial expressions can best express one's emotions, which accounts for

55% of emotional expression. Sound accounts for 38% of emotional expression, while language accounts for only 7% of emotional expression. Figure 2 shows the specific distribution. It can be seen that in these three expressions, facial expressions best reflect one's emotions, followed by sound, and finally language.
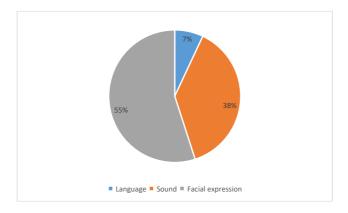
Fig. 2. Proportion of expression

In the process of using emotional computing and computer technology to evaluate the quality of classroom teaching, the facial and posture expression images of students are first obtained. Emotional calculations are then used to analyze the characteristics of the learner's facial and gesture expressions to identify student emotions. Finally, through the obtained student's emotional results, according to the learner's emotional model, the students' learning state and effect are analyzed, and the teacher's classroom teaching quality is judged.

Facial expression recognition technology covers three aspects: detection and recognition of facial expressions, feature extraction and classification. Figure 3 shows the flow of facial expression recognition.
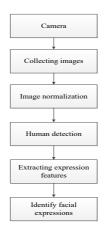
Fig. 3. Facial expression recognition process

In facial expression recognition, feature face is a common method, but this method is also insufficient. First, the amount of calculation is relatively large, and second, the requirements for pictures are relatively strict. The information for extracting facial expression images is based on the feature extraction points of the still images. The advantage of this method is that it is simple and fast. However, the disadvantage of this method is that the requirements for the extracted face image are relatively high. The expression of the face in the extracted facial expression picture must be particularly exaggerated. Therefore, the robustness of the method is bad and the recognition rate is low. Active shape model (ASM) is a popular face feature point localization algorithm. The global shape model is also known as the point distribution model (PDM). It is a shape model based on the statistical properties of training samples. The purpose of establishing a local texture model is to determine the best selected position for each feature point. By collecting the brightness of the face image, the local texture is normalized to obtain a specific vector. The facial motion coding system analyzes the characteristics and related expressions of such motion units. It consists of approximately 46 separate and connected motion units (AUs). This method is relatively straightforward and easy to understand. However, in terms of applications, the accuracy, speed and efficiency of the system need to be improved.

There are many characteristic parameters in a person's spoken voice that can explain a person's emotional state. People's emotions are different, and the corresponding emotional characteristics are also changed differently. Therefore, it is very important to study the characteristic parameters of speech that can express emotions such as speech rate, intonation, time structure, amplitude structure, fundamental frequency structure, and formant. If a person is excited now, the speed of speech will be faster than usual. Therefore, the speech rate characteristic parameter in the speech signal can be used to judge the degree of excitement of a person. If a person is in a sad state, the amplitude parameter in the speech parameter feature will be lower. If a person is in an emotional state of joy, anger, and surprise, the amplitude parameter in the speech parameter feature will be higher, the span of the amplitude parameter value will be larger, and the magnitude of the emotional change will be greater.

3.3 Facial expression data set

Facial expression recognition is affected by many factors such as illumination, background, ornaments, time, etc. These factors directly affect the robustness of facial expression recognition algorithms, which greatly hinders the application of facial expression recognition. Therefore, the current research on facial expression recognition still has great challenges, such as face recognition and expression recognition. The research, development and testing of algorithms require a lot of relevant facial expression images. Moreover, the relevant facial expression images are rich, and the negative influence of the above factors on the robustness of the facial expression recognition algorithm is easily overcome. This helps to further improve the recognition rate of the expression. At present, most algorithms are suitable for situations where the background is simple and the face pose is fixed, and the effect is good. However, for complex backgrounds and unknown face gestures, facial expression

recognition is still a difficult problem at this stage. Moreover, the detection speed and accuracy are difficult to achieve good results at the same time. Therefore, in addition to considering the situation of the positive face, the face expression with an unfixed posture should also be considered so that it can be better applied to intelligent teaching.

To increase the diversity of the data set, frontal and many facial expression images with deflection angles are included for expression recognition at different angles. To detect the recognition rate of the classifier, the plurality of expression databases is used as the test basis for the research method. In the experiment, a common facial expression database was used, such as JAFFE, Bio ID-Face Database and Yale Face of Kyushu University in Japan. The expression data set consists of 1603 384*286 images. Four expressions were shown, including happy, interest, confused, and tired. The data set contains the frontal face image and the picture with a deflected angle. The picture with the angle of deflection selects the frontal face with a tilt angle between (-30, +30). Images with different angles and different lighting conditions are included. The number of pictures for each expression is as follows: happy-442, interest-934, confused-459, and tired-145. Due to the limited number of facial expression images with the deflection angle, the number of various types of expression images in the expression database is not balanced.

Due to experimental conditions and personnel limitations, the emoticons in this experiment come from multiple libraries. Most of them come from the currently popular expression library. Since these expression libraries are created by professional departments, they are ideal for image viewing angles and facial expressions, and are suitable for related research. The images in the expression database are mainly derived from the following expression libraries.

Among them, the Japan Female Facial Expression Database (JAFFE) includes 213 gray-scale expression images of ten Japanese women, and each person displays a total of seven expressions including neutral expressions. Since the image is captured without limiting the illumination and head pose, some faces have a small angle of deflection to some extent. In the JAFFE library, each expression image has been semantically described, and the library can be used as a reference for other library expression classification. This experiment refers to the expression classification in the library when classifying facial expressions.

Bio ID Face Database, which includes 1521 face images in 384*286 grayscale natural scenes, provided by 23 testers. At the same time, it also includes the position of the eyes of each face. This database is also commonly used for face recognition and human eye positioning. Since the commonly used facial expression database is a face image collected under constraints, these databases ignore the influence of time and age information on facial expressions. The expression data set includes facial expression images of various ages, such as adding a plurality of child pictures, which makes the image samples more diverse. The recognition rate of expressions is further improved. Figure 4 shows a partial emoticon image in the dataset.

Fig. 4. Facial expression image

After the expression library is created, the library is divided into two parts: one part is used to train the face feature shape model, which is called training expression library. The library manually extracts features from the image and provides training data sets for subsequent experiments. The other part is used for testing, which is called the test expression library. This part of the data is mainly used for the test of the face feature shape model and the test of the expression classifier. The number of pictures in the training expression library is 1538. The four emoticons of happy, interest, confused, and tired are 423, 934, 133, and 448, respectively. The number of images in the test emoticon is 113. The four types of emoticons are 33, 25, 25, and 30, respectively.

3.4 Data preprocessing

Data preparation needs to be done before creating a facial expression classifier. Data preprocessing is performed on feature point information. Most importantly, the correctness and validity of the data in the test sample and the training sample data set are guaranteed. By checking the completed feature extraction image, the mark of the feature point is checked to determine the vacancy value. The location of the mark should be specified to ensure the validity of the data. For the problems that occur in the above data processing, the image needs to be remarked.

The data is processed into the following svm specified format:

[label][index1];[value1] [index2]; [value2]

Label is the kind of classification, which is usually some integer. Four types of emotions are included, and the values of the labels are defined as 1, 2, 3, and 4, respectively. Index is an ordered index, which usually takes a real number. With 68 feature points, the number of eigenvalues is 136, and the index value is [1, 136]. Value is the data value to be trained, that is, the coordinate value corresponding to each feature point. Each data should be separated by a space. Finally, the classifier is used to classify the test samples and the results are generated as labels.

Before training the expression classifier, to avoid an imbalance in the training due to a feature being too large or too small, all data is normalized. The original sample is scaled. The normalization of feature points not only facilitates data processing, but also speeds up the convergence of the training network. Normally, the zoom range is between [0,1] or [-1,1]. It should be noted that the original training set and the original test set are treated as the same data set during the normalization process.

The coordinate values of the feature points are normalized as input data to obtain normalized data. The data range is distributed between [-1, 1]. In this way, all the data

points are in a high-dimensional space with the origin as the center and the radius of 1 in the sphere. The entire data set contains m images, so it has m rows. Each image has n feature points, and a matrix of m*n is constructed. Each column constitutes a feature column. For a certain feature column, the maximum value G_{max} and the minimum value G_{min} are obtained. The normalized data range will be distributed between [R_{min}, R_{max}]. All feature values X on this column use the following normalization formula:

$$X_i = R_{max} + \frac{X - G_{min}}{G_{max} - G_{min}} \times (R_{max} - R_{min})$$

(1)

Feature points are normalized to the original training set and the original test set using SVM. After normalization, a normalized file is generated in the directory. The data in the file is the normalized data. This file can be used to create a face classifier, which is called the training set.

3.5 Optimal kernel function selection and parameter optimization

Kernel functions are widely used in the field of pattern recognition. The theory about kernel functions appeared earlier. In 1964, in the study of the potential function method, the kernel function was introduced into the field of machine learning. However, until 1992, linear SVMS research was extended to nonlinear SVMS research. Through experimental methods, the kernel function is discussed and the kernel function suitable for this study is selected. The kernel function method was introduced into the support vector machine, which is inseparable from its characteristics. First, the dimension of the input space does not affect the kernel function matrix. The kernel function method can use high-dimensional input, which effectively avoids "dimensionality disaster" and reduces the amount of calculation. Second, the function does not have to care about the form and parameters of the nonlinear transformation function. Third, when the input data is mapped to the high-dimensional feature space through the nonlinear function, the selected kernel function type and parameters will affect the properties of the feature space, which will affect the performance of various kernel function methods. The kernel function method is flexible in its specific use. It can form various kernel function methods with other algorithms. In the process of use, appropriate algorithms and kernel functions are selected according to actual needs.

Commonly used kernel functions include radial basis functions, perceptron kernel functions, Gaussian kernel functions and polynomial kernel functions. Based on the geometric characteristics of face features, the following three kernel functions are discussed, namely linear kernel function, radial basis function (RBF) and polynomial kernel function. Through the relevant experiments, the relationship between the kernel function and the recognition accuracy is further explored. Through comparative analysis, suitable kernel functions were selected for emotion recognition in this study, which prepares for the next step of recognition. At the same time, relevant conclusions can be extended to other data sets and features.

First, the kernel function selected in the experiment is introduced:

Radial basis function: The radial basis function is widely used in SVM, which is radial symmetry scalar function. For any point x in space, the monotonic function of Euclidean distance between it and a certain center xc can be represented by k(‖x-xc‖). When the distance between x and xc is far, the value of the function is small. In practical applications, the commonly used radial basis function is a Gaussian kernel function. The definition is:

$$k(\| x - xc \|) = \exp\{- \| x - xc \|^{\wedge} 2/(2*\sigma)^{\wedge}2\}$$

(2)

In the formula, xc is the center of the kernel function. σ is the width parameter of the function, which controls the radial extent of the function. The nature of the Gaussian function determines that it can filter the image well in both the spatial and frequency domains, so it is widely used in image processing.

Linear kernel function: $k(x_i, x_j)=x_i, x_j$, the linear kernel function actually performs a dot product operation on two vectors, which realizes the nonlinear transformation of the kernel mapping.

Polynomial kernel function: k(x, y) = (1 + x. y) d. The polynomial kernel d is the order of the polynomial. The larger the order, the larger the nonlinearity. The core is prone to infinity when the sample size is large.

Cross-validation is mainly used in modeling, such as regression modeling of PCR (Principal Component Regression) and PLS (Partial Least Squares Method). The sample to be modeled is divided into two parts: one part is used to build the model, and the other part is used to test the built model. The number of samples used for testing is relatively small. The prediction errors of the test samples are calculated and their squared sums are recorded. This process is repeated. When all samples are predicted one time later, the prediction error of each sample is squared and summed. This process is called PRESS. Cross-validation is also known as cross-matching, which is mainly to avoid over-fitting phenomena to obtain a reliable and stable model. It is a commonly used accuracy test method and is usually an important indicator to measure the quality of a trainer. For example, a 5-fold cross validation divides the data set into five parts. One of the four training sessions was used for each test. The accuracy of the algorithm is estimated using the mean of the five results. Usually, multiple times of cross-validation is used to find the average.

In the process of creating a facial expression classifier, when using the kernel function, two parameters c and γ are considered, where c is the penalty coefficient, and γ is different for different kernel functions. Since there is no prior knowledge of the choice of parameters, a parametric search is done to get the best (c, γ). c and γ suitable for creating this classifier are used. In this way, the classifier can better predict the test set data and improve the recognition rate of the expression. The expression data sets were respectively subjected to 3, 5, and 10 times cross-comparison, and the corresponding parameters c and γ were recorded.

The algorithm is as follows:

First, the training set is disrupted;

Second, if the cross-validation coefficient is selected as n, the training set is divided into n parts;

Third, if i=O, when i≤n-1, the following loop is performed;

Fourth, when each training is carried out, the ith share is reserved for testing;

Fifth, according to the set parameters, the remaining training sets are trained to obtain the model;

Sixth, for all vectors in the ith, the model in the fifth step is used for prediction. The classification result of the test is saved. This process is complete.

For the normalized training set, the training set is first loaded and then trained. Tests were performed using three kernel functions, including test set data and data in the training set. v is the number of cross-comparisons. Table 1 shows the experimental results.

Table 1. Kernel function test

Kernel function	Vvalue	Cvalue	γvalue	Total recognition rate of training data (%)	Total recognition rate of test set data (%)
Radial basis function	3	127.0	0.124	88.2	66
	5	127.0	0.124	88.2	73
	10	127.0	0.124	88.2	71
	3	127.0	0.124	78.6	60
Linear kernel function	5	127.0	0.124	88.2	72.6
	10	127.0	0.124	88.2	73.1
	3	127.0	0.124	84.1	63
Polynomial kernel function	5	127.0	0.124	88.2	70.4
	10	127.0	0.124	88.2	70.5

As can be seen from Table 1, not the larger the v value, the higher the total recognition rate of the data in the training set and the total recognition rate of the test set data. Here, the recognition rate refers to the average recognition rate of the four expressions in the test set. The 5-fold cross-validation effect is the best, followed by a 10-fold cross-validation. Based overall experimental situation, the creation of the expression classifier will use the radial basis function. Among them, the values of v, c, and γ are 5, 127.0, and 0.124, respectively.

3.6 Automatic recognition of facial expressions

The recognition of facial expressions involves two processes. The first process is to create an expression classifier using the training data set. The support vector machine constructs an expression classifier by analyzing the attributes in the data set. The support vector machine (SVM) was introduced for training to generate an automatic facial expression classifier. The second process is to use the built-in classifier for classification, and automatic classification of facial expressions is implemented.

Different parts of the face contribute differently to the recognition. For example, the eyes and mouth are more important than the nose. Facial expression recognition has different degrees of information on different parts of the face, and the eye information plays a greater role. In the process of expression recognition, the information of the face part has a great influence on the recognition.

Facial expressions accurately reflect people's emotions. Therefore, different facial muscles play a leading role in facial expressions that express different emotions. For example, when people show surprises, the eyes and mouth are opened a lot. When people are sad, their eyebrows and mouth are drooping. When people feel happy, there will be changes in eyebrows, small eyes, upturned corners, etc. Among them, the shape of the mouth changes most obviously. In addition, the shape of the eyebrows, the spacing of the upper and lower eyelids of the eyes, the position of the pupil in the eye, the position and shape of the lips, etc., can all be used to reflect facial expressions. Through the observation and analysis of various expression features, facial expressions are complex and rich. However, because each facial expression is caused by a series of muscle activities, different muscle activities will show different characteristics. Therefore, various facial expressions have a certain distinction between certain parts. To better distinguish each expression, the geometric relationship between the feature points is used to generate corresponding feature points. Then, it is used together with 68 feature points as training set data. The geometric feature points are the features that best represent and distinguish the expressions, which can effectively improve the recognition accuracy.Figure 5 shows the calibration of feature points.

Fig. 5. Calibration of feature points

Table 2 shows the geometric feature selection in the experiment. When each expression changes, the facial organs will be distorted accordingly. Based on this expression, for example, when people are happy, the mouth is enlarged and the distance between the upper and lower lips is increased. However, for neutral expressions and angry expressions, the upper and lower lips are spaced less because the mouth is usually closed. The ratio of the width to the height of the mouth is calculated to distinguish between happy and neutral expressions. Therefore, the new eigenvalue D1 is added. It is the ratio of the distance between the upper and lower lips and the Euclidean distance between the left and right corners. In addition, when people are confused, the eyebrows will wrinkle together and the distance between the two brows will become smaller. However, for the happy expression, the eyebrows will bend down and

the brows will be relaxed, but the inner corners of the eyes will not change. Based on this feature, the Euclidean distance between the left and right brows and the two inner corners of the eye is calculated to better distinguish the confused and happy expressions, such as D9.

Table 2. Geometric feature selection

Dm	Euclidean distance ratio	Dm	Euclidean distance ratio
D1	d52,53/d49,55	D6	D24,25/d25,26
D2	D25,30/d28,49	D7	D29,31/d34,36
D3	D19,35/d33,55	D8	D34,36/d33,35
D4	D19,25/d25,55	D9	D19,25/d30,35
D5	D18,19/d19,20	D10	d52,55/d49,55

4   Result Analysis and Discussion

The expression data set used in the experiment consisted of 1603 384*286 images, which displayed four expressions, including four emotions: happy, interesting, confused and tired. The picture is selected as a frontal face with a tilt angle between (-45, +45), including pictures with different angles and different lighting conditions. Figure 6 shows the frontal face.

Fig. 6. Frontal face

Table 3 shows the classification results for each expression.

Table 3. Classification of facial expression recognition

Expression	Number of test pictures	Number of interests	Number of happy	Number of confused	Number of tired
Interest	33	30	2	1	0
Happy	25	3	21	1	0
Confused	25	4	2	19	0
Tired	30	3	3	4	20

As can be seen from Table 3, the classification effect of "interest" is good, which can reach 70%. For the case of misclassification, "interest" may be judged as "happy" or "confused". Through the observation of the wrong picture, it is found that the "interest" expression is similar to the "happy" or "confused" expression. The corner of the mouth rises slightly, but the lips and teeth are not separated. In the process of identification, it may be greatly affected by the pre-classification features D1 and D9. The reason why the "interest" expression is not regarded as "tired" expression is that from the spatial analysis based on the basic emotion, the two types of expressions are in the opposite angles of the cone model, and they are opposite each other. From the point of view of the expression picture, the two types of expressions are clearly different. For the "tired" expression, the eyebrows are depressed, causing the upper eyelid to be depressed and the upper and lower lips to be closed. In the process of identification, it may be greatly affected by geometric feature points D3, D4, D5 and D9.

For the classification of "happy" expressions, most of them are classified into "interest" expressions. The expressions in this type of wrong picture are more relaxed and the shape of the mouth changes less. The two expressions belong to the neighboring emotions, the nature is similar, and the difference is not significant. In the process of identification, it may be greatly affected by geometric feature points D2, D3, D9. "happy" is rarely judged as "confused". The main reason is that the brows in both types of expressions have the characteristics of being lifted, which may be affected by the geometric feature point D5. However, there are many differences between the two and other expressions, such as the shape of the eye, the position of the corner of the mouth, and the shape of the mouth. It may be affected by the classification features D1, D2, D3. The main reason why the "happy" expression is not considered as the "tired" expression is that it is similar to "interest". In addition, the "happy" expression is similar to the "interest" expression. They are opposite to the "tired" expression and are also antagonistic in nature.

Part of the "confused" expression was judged as "interest". These expressions have the characteristics of large eyes and a slight opening of the mouth. This may be affected by the classification features D1, D7 and D8. The "confused" expression is treated as a "happy" expression. These expressions have the characteristics of a large mouth and a downward bend of the eyebrows. This may be affected by geometric feature points D1, D5, D6. There is a significant difference between "confused" and "tired" expressions. For example, for the "tired" expression, the corner of the mouth is squatting and closed, the brow is depressed, and the upper eyelid is lowered. The shape of the eye changes more obviously. In comparison, the "confused" expression is opened and the brow is raised. The difference between the two can be distinguished by the D1, D5 and D6 geometric feature points.

The "tired" expression was judged as a "confused" expression. The reason is that it is affected by the D2 and D5 geometric feature points. Although the eyebrows of the two are reversed, the values of D2 and D3 are very close due to the pull-down of the mouth of the "tired" expression. Through the analysis of the influence of features on various expressions, this method can be used as a reference for the selection of classification features in expression recognition. In future experiments, the suitable combination of geometric feature classification features can be selected based on the recog-

nition of the expression. This provides ideas for future related research. For different expressions, the combination of the two methods is adopted for identification. Its characteristics and laws were found, and the ideal recognition of expressions was realized.

5 Conclusion

First, based on the geometric features of the image, the commonly used linear kernel functions, polynomial kernel functions and radial kernel functions are discussed. Through the experimental method, the problems of kernel function selection and parameter optimization in SVM are further discussed. Then, the proposed geometric characteristics and SVM classification methods are introduced in detail. The results of the identification were analyzed, which provided ideas for future related research. Finally, the SVM is used to construct the expression recognition classifier, and the face expression automatic recognition model is used. The automatic classification of facial expressions was realized, and the experimental results were analyzed and compared. The proposed emotion recognition method can be applied to intelligent teaching, and a good recognition rate is obtained for some facial expressions in the four expression data sets. The validity of the method was verified. The method of facial expression recognition in intelligent teaching is studied. Facial expression recognition is applied to intelligent teaching. Emotional problems in traditional intelligent teaching are studied. At present, emotional teaching is applied in situations where the background is not too complicated and the face posture is fixed (i.e., frontal face). In view of this, an improved active shape model method is proposed to extract facial feature points. The expression classifier is built using a support vector machine. In this process, the SVM method is used to effectively realize the automatic recognition of facial expressions based on the geometric characteristics of facial feature points.

6 References

[1] Neoh S C, Zhang L, Mistry K, et al. Intelligent facial emotion recognition using a layered encoding cascade optimization model, Applied Soft Computing, 2015, vol. 34(C), pp. 72-93 https://doi.org/10.1016/j.asoc.2015.05.006

[2] Liu Z, Wu M, Cao W, et al. A Facial Expression Emotion Recognition Based Human-robot Interaction System, IEEE/CAA Journal of Automatica Sinica, 2017, vol. 4(4), pp. 668-676

[3] Elżbieta Kukla, Paweł Nowak. Facial Emotion Recognition Based on Cascade of Neural Networks, 2015, vol. 314, pp. 67-78

[4] Lee J S, Shin D H. The Relationship between Human and Smart TVs Based on Emotion Recognition in HCI, 2014, vol. 8582(2), pp. 652-667

[5] Wei Y, Zhao J. Designing robot behavior in human robot interaction based on emotion expression. Industrial Robot, 2016, vol. 43(4), pp. 380-389 https://doi.org/10.1108/IR-08-2015-0164

[6] Ko B C. A Brief Review of Facial Emotion Recognition Based on Visual Information. Sensors, 2018, vol. 18(2), pp. 401 https://doi.org/10.3390/s18020401

[7] Shah M, Chakrabarti C, Spanias A. Within and cross-corpus speech emotion recognition using latent topic model-based features. Eurasip Journal on Audio Speech & Music Processing, 2015, vol. 1, pp. 4 https://doi.org/10.1186/s13636-014-0049-y

[8] Shojaeilangari S, Yau W Y, Li J, et al. Dynamic facial expression analysis based on extended spatio-temporal histogram of oriented gradients. International Journal of Biometrics, 2014, vol. 6(1), pp. 33-52 https://doi.org/10.1504/IJBM.2014.059640

7   Authors

Yanqiu Liang is a lecturer at Nanyang Institute of Technology Normal College. His research interests include intelligent teaching, emotion recognition, support vector machine (SVM).