STUDY ON EMOTIONAL STATE CHANGE BASED ON DYNAMIC EXPRESSION SIMILARITY

Yan Zhang, Xiangyang Feng and Ming Zhu

College of Computer Science and Technology,
Donghua University, Shanghai, China

ABSTRACT

Facial expressions can express different emotions. Similar facial expressions usually correspond to the same emotions, and the changing process of emotional states is reflected in the dynamic changes of facial expressions. However, existing studies mainly focus on instantaneous emotional states, which cannot reflect the intensity of emotions. This paper proposes a method to study the process of emotional change based on dynamic expression similarity, which can assess not only the change of emotional state but also the change of emotional intensity. First, the features of dynamic expressions are extracted based on the VGG16 network model. Then, the cosine similarity of the expression features is calculated to match the corresponding emotions. At the same time, the expression intensity of each frame is calculated to evaluate the change in emotional intensity. The experimental results show that the similarity calculated in this paper is increased by 9.7% on average, which can be used for the study of emotional states.

KEYWORDS

Dynamic Expression Similarity, The Emotional State Change, Emotional Intensity.

1. INTRODUCTION

The external manifestations of emotions are often called expressions. It is a quantified form of movement of body parts in response to emotional changes. Because the facial expression is a pattern made up of all facial muscle changes, it can express a wide range of emotions and is the primary indicator for identifying human emotions. Emotion research based on facial expressions has been an active research topic in the past ten years. Ekman and Friesen [1] defined six basic expression categories: happy, angry, surprised, fearful, disgusted, and sad, corresponding to six basic human emotions. Facial expressions are recognized and classified according to the expression categories defined by Ekman. On the one hand, low-latitude facial expression features are extracted, such as Gabor wavelet transform, LBP local binary mode [2], HOG direction gradient histogram [3], etc. On the other hand, people classify expressions into pre-defined categories by seeking classifiers with better stability, such as BP neural networks and clustering [4].

However, most current facial expression studies aim to distinguish between different expressions and use static expressions to study instantaneous emotional states. However, in many scenarios, such as performance imitation, lie detection, driver fatigue detection, etc., people often do not pay attention to the category of expressions but pay more attention to the emotional information transmitted in the process of expression changes, so as to judge the current emotional state. Since similar expressions often express the same emotional information, it can be judged whether they

David C. Wyld et al. (Eds): CMLA, CSEIT, NETCoM, NLPD, GRAPH-HOC, WiMoNE, CIoT, NCS - 2022
pp. 33-45, 2022. CS & IT - CSCP 2022 DOI: 10.5121/csit.2022.121103

are expressing the same emotion by calculating the similarity degree of the expressions through the similar information between the expressions. At the same time, the dynamic process of facial expressions contains more abundant and accurate emotional information, and the similarity of dynamic expressions is closer to the state of human emotional changes. On the other hand, in addition to the type of expression, the degree of expression is also an indicator of human emotions. Expressions of the same category, such as smiling and laughing, displeasure and anger, have different levels of expression. The degree of expression corresponds to the intensity of emotion, and the change in emotion reflects the dynamic change process of emotional state. However, the current research on expression similarity is limited to the similarity between static expressions, while static expressions are only limited to instantaneous emotions and cannot reflect changes in emotional states. As a result, the ability to describe the changing state of emotions is poor, and the expression intensity has not been used in the evaluation of emotional intensity.

In response to the above problems, this paper proposes a dynamic expression similarity evaluation method to study the process of emotional change. First, the VGG16 network is used to extract the expression features, and then the KPCA principal component analysis method is used to perform dimensionality reduction analysis on the high-dimensional expression features, and finally the cosine similarity is used to reduce the dimension. The matching method compares the expression features of each frame, calculates the similarity, and calculates the intensity estimate for the expression of each frame, and combines feature similarity and intensity similarity to study the change of emotion.

1.1. Our Contributions

1. In this paper, we study the changes in emotional states from dynamic expressions and propose the use of the VGG16 network to extract expression features, which solves the problem of inaccurate feature extraction for face feature point localization. Based on this, the KPCA-based feature dimensionality reduction method and cosine similarity matching method are used to improve the similarity results.
2. We proposed to use expression intensity for the study of emotional states and to study the changes in emotional states by fusing intensity curves based on the results of dynamic expression similarity.

2. RELATED WORK

Facial expressions can intuitively reflect the emotional state of human beings and are the most direct way to express emotions. From the state of emotions, it can be divided into instantaneous emotions and dynamic emotional changes. Expression is the carrier of emotions. Instantaneous emotions are mainly expressed through static expressions [5], and dynamic changes in emotions can be reflected through dynamic expression changes [6]. Nasuha et al [7]. proposed a CNN emotion classification model that reduces parameters by separating convolutional layers and studied instantaneous emotions through the classification accuracy of seven basic emotions but did not consider the fusion of expression intensity and emotion recognition. Sui et al [8]. proposed the PSO-ELM model to recognize dynamic expressions; Fisher [9] studied emotion recognition through dynamic emotional expressions, which measure the understanding of others' emotions in everyday life but did not consider different levels of emotion.

Many scholars study expressions from the perspective of expression similarity. Vemulapalli [10] proposed a parsimonious space that is closer to human visual preferences to describe facial expressions and created a large-scale facial expression comparison dataset to obtain this simple space. This article argues that if the other two expressions are visually like the third, then the distance between these two expressions in the parsimony space will be much smaller than the distance between them and the third expression. However, only the similarity between static expressions is described. For example, if a static smile is displayed, whether it is a smile, a big laugh, a wry smile, a fake smile, or a sneer, it will be judged as similar. That is, a smile and a big laugh will also be judged to be similar in emotion, but in fact, the intensity of the two emotions is different. Unfortunately, little research on emotions takes emotional intensity into account. Schroff et al. [11] proposed a new method, FaceNet, which can directly learn the mapping from images to points in Euclidean space. The distance between the points in the Euclidean space of the features corresponding to the two images directly corresponds to whether the two images are similar. Similarly, Schroff et al. [11] only studied the similarity of static emotions, that is, transient emotions. And whether it is from the perspective of emotion recognition or similarity, it ignores the dynamics of facial expressions and the importance of expression intensity. Expression intensity can reflect the intensity of emotions and deeply understand the psychology of characters. This paper proposes an evaluation method based on dynamic expression similarity and emotional intensity to study the change process of emotion.

3. DYNAMIC FACIAL EXPRESSION SIMILARITY MATCHING

The same emotional state often has similar facial expressions, and the dynamic changes in facial expressions directly reflect a person's emotional changes. The dynamic face similarity matching algorithm proposed in this paper mainly includes expression feature extraction, dimensionality reduction of high-dimensional expression features based on KPCA, emotional intensity estimation, and expression similarity matching based on cosine similarity and determine emotional states using similarity results. Its experimental flow chart is shown in Figure 1.

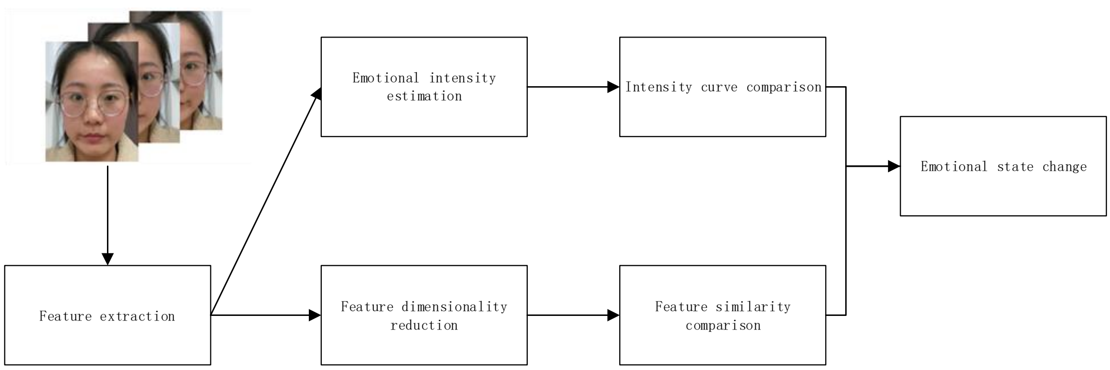

Figure 1. Emotional state change matching flow chart.

3.1. Feature Extraction Network

Facial expression feature extraction is mainly to locate and extract facial organ features, texture regions, and predefined feature points. General facial expression feature extraction methods are divided into two categories: traditional feature extraction and feature extraction based on deep learning. Traditional feature extraction can be based on geometric features, gray features, motion features, frequency features, etc., using the LBP local binary mode [2] and the HOG direction gradient histogram feature [3] [21]. The method based on deep learning [12] mainly uses convolutional neural networks to extract high-dimensional features of facial expressions in combination with the current popular network, which solves problems such as ignoring local feature information, feature loss, and inaccurate feature extraction in traditional manual extraction of facial features.

As one of the most representative networks in deep learning, the convolutional neural network has achieved great success in the field of image processing, relying on several convolutional layers for feature extraction, and many successful tasks such as image classification and recognition are based on CNN [13]. Compared with traditional image processing algorithms, CNN has the advantage of being able to directly input the original image to extract image features, avoiding the complex image preprocessing process. When using neural networks to extract expression features, the output of the last layer of the network is not directly used as expression features, because each value in the feature vector output by the last layer represents the confidence that the expression may be a certain expression. If the expression does not appear in the classified expression, its value will be suppressed to a very low level, which is not conducive to the subsequent similarity calculation.

In this paper, VGG16 is selected as the feature extraction network of facial expressions. VGG16 won the runner-up to ImageNet in 2014, and its basic mechanism follows the traditional CNN architecture. VGG16's network has a total of 16 layers, including 13 convolutional layers and 3 full connection layers [14]. In this experiment, the output value of the last fully connected layer of VGG16 was selected as the feature vector of the expression. According to the structure diagram of VGG16, the dimension of the feature vector obtained from this layer is 4096.

3.2. KPCA Feature Dimensionality Reduction

Dimensionality reduction of facial expression features is to select the most representative features from the original facial expression features and reduce the time complexity of the algorithm while preserving the facial expression features. Existing dimensionality reduction methods are mainly divided into linear and nonlinear dimensionality reduction [15]. Nonlinear dimension reduction [16] mainly includes local linear embedding (LLE) that preserves local features and isometric feature mapping (Isomap) that preserves global features. Linear dimension reduction mainly includes Principal Component Analysis (PCA) [12], Linear Discriminant Analysis (LDA), etc. PCA dimension reduction mainly maps high-dimensional data features to new low-dimensional feature spaces and transforms them into several new comprehensive features. The new low-dimensional comprehensive features are linear combinations of the original high-dimensional features. However, PCA dimension reduction only uses the global information of the face image, and the effect is not very good under different facial expressions and postures. Scholkopf [17] proposed kernel principal component analysis (KPCA) as a nonlinear extension of PCA, which performs linear principal component dimension reduction in the high-dimensional feature space by mapping the data features of the input space to the high-dimensional feature space. Therefore, this paper adopts the KPCA feature dimension reduction method to reduce the dimension of the extracted features.

The facial expression feature vector extracted based on VGG16 has 4096 dimensions, which is a deeper facial expression feature, and the computational complexity of facial expression sequence similarity is too high. Feature dimension reduction can reduce the complexity of facial expressions while preserving their original features. The KPCA feature dimensionality reduction method used in this article is a nonlinear principal component analysis method that introduces a kernel function based on PCA dimensionality reduction and uses a nonlinear mapping to map the feature to a high-dimensional or even infinite dimension. The dimensional space is converted to linear, and principal component analysis is performed in the mapped kernel function space [18,19], to achieve the purpose of feature dimensionality reduction. Its commonly used kernel functions generally have linear kernel function, Q-order polynomial kernel function, Sigmoid kernel function, Gaussian radial basis kernel function, and Laplace kernel function. The Gaussian radial basis kernel function is shown in Equation 3-1:

$$K(X_i, X_j) = \exp(-\frac{||X_i - X_j||}{\sigma^2}) \quad (3\text{-}1)$$

Where X_i, X_j can be represented as an eigenvector of input space, $||X_i - X_j||$ can be viewed as the square Euclidean distance between $X_i$and X_j,σ is a free parameter. Since the Gaussian radial basis kernel function can map the feature data to the feature space, the value of each element is between (0,1] and there is only one kernel parameter, which can alleviate the nonlinear component between the features, so this paper chooses Gaussian the radial basis kernel function is used as the kernel function of KPCA. The process of reducing the dimensionality of the extracted high-dimensional features using KPCA is as follows:

(1) Based on the high-dimensional features extracted from the VGG network, the feature values are centralized, and then the features are mapped to the high-dimensional feature space using the Gaussian radial basis kernel function.
(2) According to the eigen values mapped to the higher-dimensional space, the mean value of the features is calculated, and the covariance matrix of facial features is constructed.
(3) The eigen values and eigenvectors were calculated according to the constructed covariance matrix of facial expression features. The feature vector corresponding to the maximum eigen value is selected as the low-latitude feature after dimensionality reduction of the high-dimensional feature, namely the principal component feature.
(4) The eigenvectors corresponding to the eigen values constitute the subspace after dimensionality reduction, which is the expression feature after dimensionality reduction.

3.3. Matching Facial Expression Similarity

The common similarity match mainly includes distance measures and correlation measures. A distance measure is a metric that compares the definitions of two images using a distance function. The more similar the images, the smaller the distance. Distance measures include Euclidean distance, Manhattan distance, and Chebyshev distance. Correlation measures include cosine similarity, Pearson correlation coefficient, and so on.

The Euclidean distance measures the absolute distance between each point in multi-dimensional space [20]. Two n-dimensional eigenvectors $X = (X_1, X_2, X_3, \ldots\ldots, X_n)$ and $Y = (Y_1, Y_2, Y_3, \ldots\ldots, Y_n)$.Then the Euclidean distance formula of the two feature vectors is as follows:

$$D(X, Y) = \sqrt{\sum_{i=1}^{n}(X_i - Y_i)^2} \quad (3\text{-}2)$$

To reduce the amount of calculation in the experiment, the following Euclidean distance calculation formula is generally adopted, without the square root of the Euclidean distance:

$$D(X, Y) = \sum_{i=1}^{n}(X_i - Y_i)^2 \quad (3\text{-}3)$$

Cosine similarity evaluates the similarity of two vectors by calculating the cosine of the angle between them. Two N-dimensional feature vectors X, Y, the similarity formula of the two feature vectors is as follows:

$$\cos(\theta) = \frac{\sum_{i-1}^{N} X_i Y_i}{\sqrt{\sum_{i-1}^{N} X_i^2}\sqrt{\sum_{i-1}^{N} Y_i^2}} \quad (3\text{-}4)$$

The cosine value is in the range [-1, 1]. The cosine value of two coincident vectors is 0, the cosine value of two vectors in the same direction is 1, and the cosine value of two vectors in the opposite direction is -1. Compared with Euclidean distance, cosine distance pays more attention to the difference in direction between two vectors.

By comparing several groups of experiments, this paper finally decided to use the cosine similarity distance function as the similarity judgment function, because the cosine distance is used to measure the consistency of vector dimension direction, and the cosine value of the vector, including angle, is used to reflect similarity. In this paper, the low-latitude facial features obtained by KPCA dimensionality reduction are treated as vector features, and the degree of similarity is reflected by calculating the cosine value. The Euclidean distance function, on the other hand, does not pay attention to the difference in dimension. The greater the absolute distance value, the greater the similarity; the greater the absolute distance value, the less similar it is. And the algorithm complexity is higher than the cosine function.

According to the similarity results, it is judged that it is the same emotion, and then the relationship between its intensity changes is studied according to the intensity of emotion.

3.4. Emotion Intensity Estimation

For the task of expression recognition and classification, it only needs to identify what kind of expression it is, but for emotions, emotions are a dynamic process that requires not only similar emotion categories but also similar intensity of emotions. Expression intensity is a measure of emotional intensity that can reflect the degree of emotional change and is an indispensable part of emotional research based on dynamic expression similarity. However, there are relatively few studies on expression intensity at present. On the one hand, it is because the recognition and classification of static expressions and the research on similarity do not need to involve the knowledge of intensity, and on the other hand, there is no uniform definition of expression intensity. Prkachin et al. [22] used facial motion units to define expression intensity, but it took a lot of time and manpower; Hess et al. [23] defined expression intensity through the difference between expression images in different frames in the video, but the above the studies on expression intensity were all separate studies, and none of them combined studies on expression intensity and emotion.

Different levels of emotions reflect the psychology of different people. For dynamic emotions, expression intensity can dynamically reflect the process of emotional intensity changes and then reflect the state of emotions. Expression intensity estimation is the process of dividing each type of facial expression into degrees so that it better reflects the changes in facial emotion. For example, smiling and laughing represent two different degrees of happy emotions. In this paper, the expression intensity values are estimated based on the VGG16 network and Softmax, and the threshold value after the output of Softmax is defined as the intensity value of the expression. From the output of the Softmax function, the intensity value ranges from 0 to 1, which exactly corresponds to the process of intensity from none to climax. In this paper, the intensity curve of each group of emotions is obtained based on the intensity estimation value, and the state of the emotion is judged based on the matching degree of the intensity curve based on feature similarity.

4. EVALUATION EXPERIMENT

This paper conducts experiments based on dynamic expression similarity and emotional intensity to verify the change in an emotional state. First, use VGG16 to train an expression feature extraction model and obtain the emotional intensity value based on the model; secondly, use the cosine function to calculate the feature similarity in the similarity calculation. It is judged whether it is the same emotional state according to the feature similarity and emotional intensity of dynamic expressions.

4.1. Dataset

The data set for this experiment includes the Extended Cohn-Kanade Dataset (CK+), which is used for the training of the VGG16 expression feature extraction model, and the collected pictures of dynamic expression changes to evaluate the results of emotional changes. The CK+ dataset includes 123 subjects and 593 image sequences, of which 327 sequences have expression labels. And the labeled expression sequence contains the change of the subject's expression from calm to peak, which can be used for the study of emotional changes. The change sequence of surprise emotions of a subject in the data set is shown in Figure 2, and its surprise emotion gradually reaches its peak emotion.

Figure 2. A sequence of images in CK+ dataset.

4.2.Model Training

Model training uses TensorFlow framework, GPU model is NVIDIA GTX1650, 4G memory. The weights are initialized through TensorFlow's variable_initializer, the initial learning rate is set to 0.1 using the Adam optimizer, and the learning rate is optimized using cross-entropy with a step size of 0.0001. When the loss of the training set for three epochs is no longer reduced, the learning rate is reduced by 10 times, and the batch_size is set to 50 to prevent the model from overfitting. The training process of the model is shown in Figure 3, and its accuracy and loss values are saved every 100 iterations. On the left is the accuracy curve for the training set. The X-axis represents the iteration epoch, and the Y-axis is the change in accuracy. The figure on the right is the loss curve of the training set. The X-axis represents the iteration epoch, and the Y-axis is the loss value for the training set. From the 150th epoch, the accuracy has been infinitely close to 1. The model gradually converges.

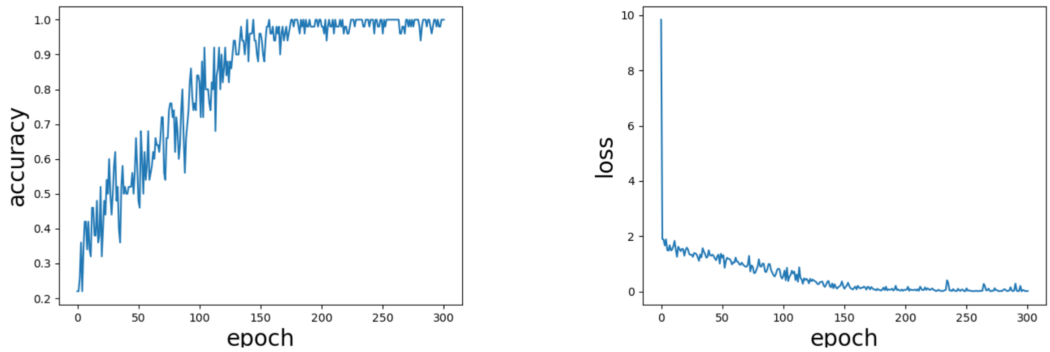

Figure 3. Loss and accuracy changes.

4.3. Expression Similarity Calculation

In this paper, the expression similarity calculation is proposed. Firstly, the expression features are extracted based on the full connection layer of the VGG16 model, and the extracted features have 4096 dimensions, which are used as the feature vector of the expression. The similarity was calculated based on two facial expression sequences, one of which was used as the standard sequence and the other as the test sequence. The feature vectors of each frame of the facial expression sequence were combined to form a high-dimensional feature matrix, and the kernel principal component dimension reduction algorithm was used to reduce the dimension of the feature matrix.

The dimensionality reduction calculation of an extracted single frame of facial expression was used to effectively reduce the latitude of facial expression while keeping the original features unchanged to the maximum extent. The cosine function was used to calculate the comparison of similarity between a single frame of facial expression after dimensionality reduction and before dimensionality reduction. The data for dimension accuracy before and after dimensionality reduction is shown in Table 1.

Table 1. Comparison of similarity of different dimensions

Dimension	4096	1024	256	64	32	16	8
Precision	0.9036	0.9037	0.9038	0.9111	0.9025	0.8256	0.7898

According to the results in Table 1, when the dimensionality of facial features is reduced by nuclear principal component analysis, its accuracy goes through a process of first increasing and then decreasing. When it is reduced to 64 dimensions, it has the highest similarity to the original feature and can retain the facial feature information to the maximum extent. As can be seen from the accuracy after dimensionality reduction in Table 1, there is only a 0.0086 difference between the accuracy of 32 dimensions and 64 dimensions. However, considering the complexity of dimensionality reduction, the complexity of facial features reduced to 32 dimensions is higher than that of 64 dimensions. Therefore, when We were reducing the dimensions of facial features, we chose to reduce the high-dimensional facial features to 64 dimensions.

4.4. Experimental Results

Table 2 shows the comparison of the calculation results of the similarity of expressions, where VGG16 indicates that the expression feature extraction method uses a deep learning model, and no indicates that the expression features are extracted according to the facial geometric features. PCA, KPCA, and Isomap represent three different dimensionality reduction methods, and the cosine function and Euclidean distance represent two different similarity matching algorithms. Figures 4 and 5 show the sequence of two groups of emotions, respectively. Figures 7 and 8 show the sequence of two groups of emotions, respectively. Emotional intensity is the process of having no emotion to experiencing the peak of emotion. Figures 6 and 8 show a comparison of the two sets of intensity curves.

Table 2. Comparison of similarity results of different dimensionality reduction methods

Method	VGG16	Cosine function（%）	Euclidean distance	Increase（%）
KPCA	√	**92.5**	**0.29**	**9.7**
	×	86.5	0.65	3.7
PCA	√	91.8	0.33	9
	×	85.7	0.87	2.9
Isomap	√	90.6	1.07	7.8
	×（standard）	82.8	2.17	0

According to the experimental results in Table 2, after extracting facial features based on geometric features, the benchmark method adopted isomap feature dimension reduction, and the cosine function was used to calculate the similarity of 0.828 and the European distance of 2.17. When VGG16 is used to extract facial features, the cosine similarity is 0.906 after isomap feature dimensionality reduction. Compared with facial features extracted without VGG16, the cosine function similarity improved by 7.8%, and the Euclidean distance is 1.07, which is also smaller than 2.17, and its similarity is higher. It shows that the facial features extracted by the deep learning model are more comprehensive and accurate than the facial geometric features. Based on the facial features extracted from VGG16, the feature dimension reduction method was changed. After PCA dimension reduction, cosine similarity was 0.918, and KPCA dimension reduction cosine similarity was 0.925, which were 1.2% and 1.9% higher than isomap similarity, respectively. The results show that KPCA can retain the high-dimensional facial features to a greater extent after dimensionality reduction. VGG16 was used to extract facial expression features, and the similarity of KPCA features after dimensionality reduction was 0.925, which was about 9.7% higher than that of the benchmark method.

Figure 4 shows a set of standard happy emotion sequences that contain 9 frames of emotion change. From the first frame in the upper left corner to the ninth frame in the lower right corner, the happy emotion frame has experienced expressionless to the highest climax of happy emotion. Its happy mood intensity goes from nothing to a climax. Figure 5 is a set of emotional frames imitating the happy emotion sequence in Figure 4, and the feature similarity between the two sets of emotions is 92.5%. Figure 6 shows the curve of the intensity change between the two sets of happy emotion sequences. From the template emotion sequence and the imitator's emotion sequence, the expressions in the two sequences have experienced the process from having no emotion on their faces to reaching a climax emotion. The emotional intensity curve is consistent with the change of the emotional sequence. The emotional intensity increases gradually and then peaks. If the comparison between the first frame in the emotional sequence 1 and the first frame in the emotional sequence 2 is considered alone, and the intensity curve is not considered, the interpreted emotion is not happy and cannot reflect the emotional change process. According to the similarity results combined with the emotion intensity curve to judge the emotion change process, it can be concluded from the trend of the intensity curve that the dynamic characteristics of the two emotion sequences are similar, and that the emotion change process is also similar.

Figure 4. Happy emotion sequence 1

Figure 5. Happy emotion sequence 2

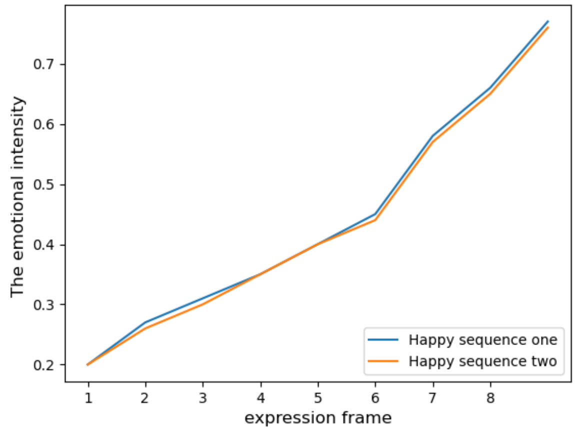

Figure 6. Comparison of the intensity curves of happy emotions

Figure 7 shows a sequence of dynamic changes of a group of template surprise emotions from the first frame in the upper left corner to the twelfth frame in the lower right corner. Figure 8 is a sequence of surprise emotions imitating Figure 7, and the feature similarity between the two sets of emotions is 92.48%. Figure 9 depicts a comparison of the intensity of the two surprised emotion sequences depicted in Figures 7 and 8. The two surprised emotion sequences are from no expression to a surprised climax expression, and their intensity has also experienced a change process from 0 to close to 1, which is the same as the change trend of expression. It can be concluded that the method of dynamic expression similarity combined with emotion intensity estimation proposed in this paper can be used to effectively evaluate the process of emotion change.

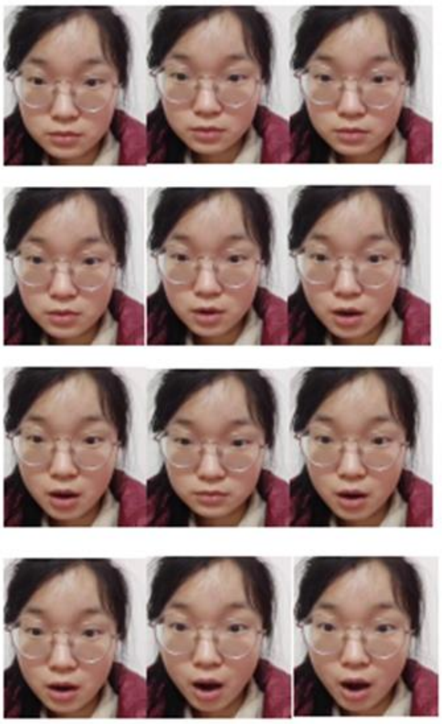

Figure 7. Surprise emotion sequence 1

Figure 8. Surprise emotion sequence 2

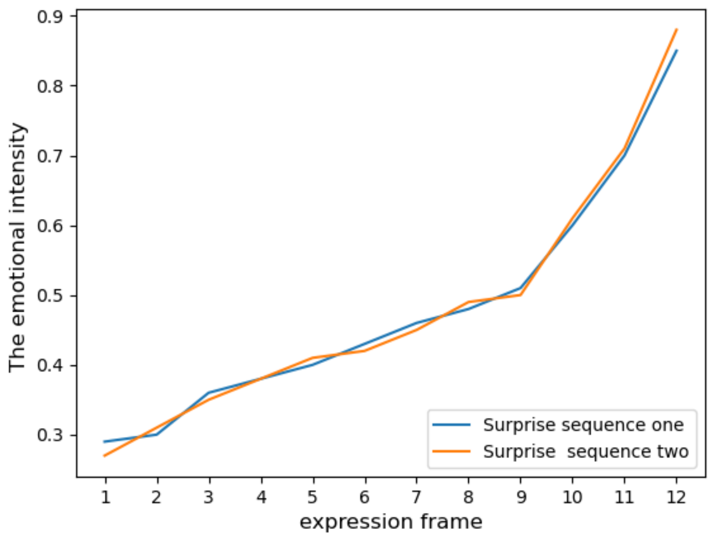

Figure 9. Comparison of surprise intensity curves

5. CONCLUSION

This paper proposes a method to study the emotion process based on dynamic facial expression similarity. Firstly, facial expression features are extracted through the VGG16 network to improve the inaccuracy of facial expression feature extraction based on facial geometric features. Secondly, facial expression feature similarity and facial expression intensity similarity are combined to make the results more accurate. Compared with facial feature extraction based on geometric features, the proposed algorithm can effectively extract facial features and then calculate the similarity of facial expression sequences. On average, compared with facial feature extraction based on geometric features, the similarity of facial expression sequences is improved by 9.7% on average, which can better evaluate the process of emotional change. However, the expression similarity measurement method proposed in this paper is mainly based on positive or near-positive expressions without considering the influence of side and facial occlusion on expression. Therefore, the future research direction is mainly to study the influence of head posture on the emotion change process.

REFERENCES

[1] P. Ekman and W. V. Friesen, "Facial action coding system (facs): a technique for the measurement of facial actions," *Rivista Di Psichiatria*, vol. 47, no. 2, pp126-38, 1978.
[2] J. ma, "Facial expression recognition based on feature fusion," *Industrial control computer*, vol. 33, no. 11, p4, 2020.
[3] E. Meyers and L. Wolf, "Using biologically inspired features for face processing," *International Journal of Computer Vision*, vol. 76, no. 1, pp93-104, 2008.
[4] T. Zavaschi, A. S. Britto, L. Oliveira, and A. L. Koerich, "Fusion of feature sets and classifiers for facial expression recognition," *Expert Systems with Applications*, vol. 40, no. 2, pp 646-655, 2013.
[5] Y. Tang, X. Zhang, X. Hu, S. Wang, and H. Wang, "Facial expression recognition using frequency neural network," *IEEE Transactions on Image Processing*, vol. 30, pp 444-457, 2021.
[6] N. Perveen, D. Roy, and K. M. Chalavadi, "Facial expression recognition in videos using dynamic kernels," *IEEE Transactions on Image Processing*, vol. 29, pp 8316-8325, 2020.
[7] A. Nasuha, F. Arifin, A. S. Priambodo, N. Setiawan, and N. Ahwan, "Real time emotion classification based on convolution neural network and facial feature," *Journal of Physics Conference Series*, vol. 1737, no. 1, pp 012008, 2021.
[8] X. Sui, L. Xue, and D. Li, "Dynamic Expression Recognition Based on Hybrid Features and Optimized Extreme Learning Machine Model," *IOP Conference Series*: Materials Science and Engineering, vol. 719, no. 1, p. 012080(6pp), 2020.
[9] H. Fisher, "Emotion recognition from realistic dynamic emotional expressions cohere with established emotion recognition tests: A proof-of-concept validation of the emotional accuracy test." *Journal of Intelligence*. vol. 9, 2021.
[10] R. Vemulapalli and A. Agarwala, "A compact embedding for facial expression similarity," in 2019 IEEE/CVF *Conference on Computer Vision and Pattern Recognition (CVPR)*, 2019, pp 5676-5685.
[11] F. Schroff, D. Kalenichenko, and J. Philbin, "Facenet: A unified embedding for face recognition and clustering," in *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2015, pp815-823.
[12] B. Yang and Y. Yang, "Pca is used to reduce dimension of deep learning image feature extraction," *Computer system application*, vol. 28, no. 01, pp 281-285, 2019.
[13] L. Rampasek and A. Goldenberg, "Tensorflow: Biology's gateway to deep learning?" *Cell Systems*, vol. 2, no. 1, pp12-14, 2016.
[14] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," *Computer Science*, 2014.
[15] J. Liu and F. Zhao, "High dimensional data reduction technology and research progress," *Electronic technology*, vol. 031, no. 003, pp 36-38, 43, 2018.
[16] J. Du, X. Wang, and L. Hu, "Nonlinear dimensionality reduction technology and visualization application," *Journal of Donghua University (Natural Science edition)*, vol. 46, no. 4, pp 6, 2020.

[17] B. Schölkopf, A. Smola, and K. Müller, "Nonlinear component analysis as a kernel eigenvalue problem," *Neural Computation*, vol. 10, no. 5, pp1299-1319, 1998.
[18] M. R. Galiaskarov, V. V. Kurkina, and L. A. Rusinov, "Online diagnostics of time-varying nonlinear chemical processes using moving window kernel principal component analysis and fisher discriminant analysis," *Journal of Chemometrics*, pe2866, 2017.
[19] A. A. Joseph, T. Tokumoto, and S. Ozawa, "Online feature extraction based on accelerated kernel principal component analysis for data stream," *Evolving Systems*, vol. 7, no. 1, pp15-27, 2016.
[20] L. Gao, "Research and application of text clustering algorithm," Ph.D. dissertation, *University of Electronic Science and Technology of China*, 2013.
[21] E. Meyers and L. Wolf, "Using biologically inspired features for face processing," *International Journal of Computer Vision*, vol. 76, no. 1, pp 93-104, 2008.
[22] K. M. Prkachin and P. E. Solomon, "The structure, reliability and validity of pain expression: Evidence from patients with shoulder pain," *Pain*, vol. 139, no. 2, pp267-274, 2008.
[23] U. Hess, S. Blairy, and R. E. Kleck, "The intensity of emotional facial expressions and decoding accuracy," *Journal of Nonverbal Behavior*, vol. 21, no. 4, pp 241-257, 1997.