

Research on Theory and Method for Facial Expression Recognition System Based on Dynamic Image Sequence

Yang Xinfeng^{1,*} and Jiang Shan²

¹*School of Computer & Information Engineering, Nanyang Institute of Technology, Henan, Nanyang, 473000, P.R. China*

²*Henan University of Traditional Chinese Medicine, Henan, Zhengzhou, 450000, P.R. China*

Abstract: We studied the tracking theory based on Candide3 facial model and according to the research we further developed the tracking process. Based on the research of facial tracking, we put forward a dynamic feature extraction method based on six parameters of facial model. We introduced the active appearance model algorithm to locate and track feature points of facial expression, then studied the tracking theory. Based on Candide3 facial model and according to the research we further developed the tracking process. We use the Dynamic Time Warping (DTW) technique to align the image sequence, and then extract the feature vector. The results show that the method based on the model parameters has a good classification results when extract dynamic characteristics.

Keywords: Facial expression recognition, face detection, extraction of dynamic feature, the Dynamic Time Warping technique.

1. INTRODUCTION

Expression feature extraction in face recognition is a vital part of the process, at present, domestic and foreign scholars on the facial expression made a lot of feature extraction algorithms, These algorithms are based on the 6 basic expressions of Ekman [1] and Frisen under the proposed framework. Facial expression recognition based on static images with high efficiency, high-speed characteristics, but there are a lot of limitations, due to the amount of information contained in the image is small, so easily affected by many factors such as the external environment and individual differences, such as different colors, differences in facial appearance, uneven illumination, final results are likely to interfere with expression recognition, making the system robust decrease. Facial expression is a continuous process of change, and dynamic image sequence contains a continuous movement or change of image, image sequences can be extracted from more richer facial expression information [2, 3], thereby reducing even eliminate the interference of the individual and the external environment, making facial expression recognition under various conditions to achieve better results, so many people began to study the dynamic image sequence based on facial expression recognition, then, how better to extract dynamic images sequence expression feature has become an important issue. In this article we discuss the two dynamic feature extraction methods, which are based on the feature point tracking algorithms and model parameters ASM Candide3 tracking algorithm, and the algorithm is how

to extract the motion feature image sequence expression were studied.

2. MOTION FEATURE EXTRACTION BASED ON ACTIVE APPEARANCE MODEL

2.1. The Active Shape Model

Changes in the expression of the human face tracking feature points is a dynamic feature extraction method, this method is to select the most representative of changes in the expression of the feature points, feature points through these changes may reflect the movement trends of expression on the feature points of the face are generally selected organs by tracking feature points can ignore the other without the necessary background and irrelevant information to extract the expression of motion information.

Active Shape Model (Active Shape Models, ASM) [4] is a feature matching method based on statistical model, it needs to calibrate the shape of the feature points of the target object is constructed as a training sample an active shape model. The main core of the algorithm is two sub-models: the global model and local texture model. First, by way of manual calibration calibrate the shape of the feature points of the target object, as a set to generate a training set; then the sample statistics, and establish a statistical model. This model is a model of the approximate location of the feature point has, therefore, after the statistical model, ASM methods still have to deal with the use of local texture model feature point detection target search to find the best match position of the feature point, and then through the feedback establishing the statistical model parameter adjustment, so that the true



Fig. (1). ASM algorithm for tracking facial expression feature points.

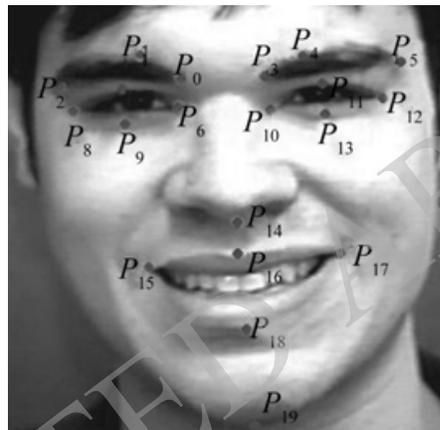


Fig. (2). The definition of feature points.

contour model and a little closer to the target, after completing the adjustment, can be the target feature points for precise positioning.

This section uses the algorithm to extract human face image sequences extracted based on the improved ASM [5] of the feature points in the face Feature points, through improved methods, can be more precise positioning of facial feature points.

Fig. (1) is a dynamic expression Cohn-Kanade Gallery [6] in the resulting image use ASM movement tracking feature points 8, can be seen in the face calm and face the highest point of the case, people can accurately locate facial feature points.

2.2. Geometric Feature Extraction

When the feature point extraction, the last chapter of the first to use face detection algorithm to the first frame image face position detection and location, and then use face detection results of the ASM initialization, each frame will be the next before the results of the previous frame as the value of the initialization, and then tracking the results of the update to the ASM model. Calculate the distance parameter for each frame of each individual face feature points, then the distance parameter after a frame image is subtracted from the

previous frame parameters, obtained the geometric characteristics of this expression vector. The aim is to extract the motion can better correlation between the expression of the moving image sequence, expression of the better use of the motion information.

Firstly, the ASM model of human face location of 68 feature points, so you can easily extract the coordinates of 68 facial feature points of this, but if these 68 points, then all the feature extraction, vector dimension will high, will bring a lot of redundant information, it will lead to a decline recognition rate, including many parts of feature points on the outer contour of the face, including the change in expression when there will not be a great location the change, which for our geometrical feature extraction is meaningless. Changes in the expression of the human organs such as the face feature point eyes, eyebrows and mouth, the position data will be a corresponding change, so that these need to extract large contribution to expression recognition feature point motion information.

The paper selected (facial characteristic points, FCP) [7] of the feature point set in the entire set of feature points locate feature points altogether 20 people face, are the most able to reflect changes in the expression of the point, with the when changes in the expression variation points organs also varies, as shown in Fig. (2).

Table 1. Geometric distance definition of facial feature point.

v_i	Geometric Distance	Features	v_i	Geometric Distance	Features
V_1	$(P_0, P_1)_y$	Left eyebrow	V_{10}	$(P_{11}, P_{13})_y$	Right eyebrow
V_2	$(P_0, P_2)_y$	Left eyebrow	V_{11}	$(P_{10}, P_{12})_y$	Right eyebrow
V_3	$(P_3, P_4)_y$	Right eyebrow	V_{12}	$(P_{10}, P_{13})_y$	Right eyebrow
V_4	$(P_3, P_5)_y$	Right eyebrow	V_{13}	$(P_{14}, P_{16})_y$	Mouth
V_5	$(P_0, P_{14})_y$	Left eyebrow	V_{14}	$(P_{15}, P_{18})_y$	Mouth
V_6	$(P_3, P_{14})_y$	Right eyebrow	V_{15}	$(P_{14}, P_{15})_y$	Mouth
V_7	$(P_7, P_9)_y$	Left eyebrow	V_{16}	$(P_{14}, P_{17})_y$	Mouth
V_8	$(P_6, P_8)_y$	Left eyebrow	V_{17}	$(P_{15}, P_{17})_y$	Mouth
V_9	$(P_6, P_9)_y$	Left eyebrow	V_{18}	$(P_{14}, P_{19})_y$	Chin

Since each frame image of the face can not be completely identical, so when the distance parameter extraction, the need for the location of these feature points are normalized, so that the site can be eliminated since the head or face sizes posture changes due to differences deformation of facial feature points. Use P6, P10 and P14 of the three coordinates of points, the affine transformation can be by each frame image 20 is aligned feature point coordinates [8], then the use of this 20 face feature point position information 18 is constructed dimensional geometric features. Facial expression changes caused by facial muscle strain is usually in the vertical direction, so the distance between the extracted parameters are mainly concentrated in the vertical direction, while the horizontal direction only calculate the horizontal distance between two points outside the mouth. 18-dimensional geometric distance parameter defines these 20 points as shown in Table 1.

P is a face feature points, $(P_i, P_j)_x$ as a feature point P_i , P_j , and the horizontal distance, $(P_i, P_j)_y$ as a feature point P_i , P_j and the vertical distance. For a video sequence, we passed the distance of each frame image geometry geometric distance minus one frame can be calculated per frame motion change facial feature points, the feature vectors are stored in the matrix.

$$x_i = (dv_1, dv_2, dv_3, \dots, dv_{18})^T, i = 1 \dots n-1 \quad (1)$$

Wherein x_i is a 18-dimensional feature vector, representing the motion of feature points geometric distance, n is the length of the image sequence. So that we can pass the image displacement vector for each frame in the sequence of feature points, feature points obtained displacement matrix $X = [x_1, x_2, \dots, x_{n-1}]$, apparently, the feature dimension feature matrix X is $18 \times (n-1)$.

3. DYNAMIC FEATURE EXTRACTION BASED ON THREE-DIMENSIONAL FACE MODEL CANDIDE3

3.1. Candide3 3D Face Model

Candide3 is a parameterized model [9]. Candide3 model is composed of 113 points $P_i (i=1,2,\dots,113)$ composition, and these points by a certain order link into triangle meshes, where each triangle is called a patch, a total of 184 surface sheet (see Fig. 3). This model can be described as:

$$g = sR(\bar{g} + AT_a + ST_s) + t \quad (2)$$

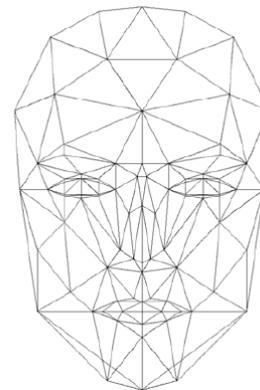


Fig. (3). Candide3 3D face model.

Wherein s is the amplification factor, $R = R(r_x, r_y, r_z)$ is the rotation matrix, \bar{g} is the standard model, A is the movement unit, S is shaped unit, T_a, T_s respectively, change their corresponding parameter. $t = t(t_x, t_y)$ for model

transformation vector in space, \bar{g} is the desired face model obtained.

In the above parameters, the shape of the shape of the module contains 12 cells [8]. There are 11 sports unit, including the Upper lip raiser, Lower lip depressor, Inner brow lower, Outer brow raiser, lip corner depressor, Lip stretcher, Eyes closed, Lid tightener, Nose wrinkler, Lip presser, Upper lid raiser.

Candide3 g mesh model can be expressed as:

$$\bar{g} = \bar{g} + AT_a + ST_s \quad (3)$$

Wherein \bar{g} is a standard model, A is the motion module, S is the shape of the module, T_a , T_s respectively, change their corresponding parameter. By T_a , T_s you can get change frontal face different expression changes.

Since it is desirable to different head pose, size and location tracking human faces, the model and the introduction of the six parameters, i.e. the above mentioned amplification factor S , the rotation matrix $R = R(r_x, r_y, r_z)$, and the conversion vector $t = t(t_x, t_y)$, the finally obtained:

$$\bar{g} = sR(\bar{g} + AT_a + ST_s) + t \quad (4)$$

In the tracking process, the shape parameter once identified, will no longer change to the composition of the remaining parameters vector b , that is $b = [s, r_x, r_y, r_z, T_a, t_x, t_y]$, b can be reflected in changes in the tracking process. Candide3 model based tracking algorithm is to quickly update to b , so that the face of the current model of facial expression match.

3.2. Expression Feature Extraction Motion Parameters

Candide3 model is a very detailed model of the face, changes in the control model is given the shape of 12 units and 11 units of motion simultaneously, Candide3 model also gives the corresponding AUs unit according to the respective movement means, for expression analysis provides a convenient.

Emotes unit (Facial action units, FAU) [9], the expression refers to the basic elements of a common organ changes in facial expression changes in the composition, and other emotions can be represented by a combination of FAU, so typically by computer vision tracking FAU to analyze changes in the expression [10].

Candide3 model motion control unit and motor unit face correspondence, research to identify the corresponding seven sports unit, which includes tracking applications Upper lip raiser, Lower lip depressor, Inner brow lower, Outer brow raiser, Upper lip outside raiser, Lip stretcher six sports unit also includes a Nose wrinkler motor unit.

Candide3 model gives control model changes in the shape of 12 units and 11 units, while the movement,

Candide3 model is also given according to various sports unit of the corresponding AUs unit provides a convenient expression analysis [11, 12].

Combined emotes element analysis, dynamic characteristics presented here is based on a model parameter extraction Candide3 program. That application Candide3 tracking algorithm based on the model, change tracking head pose human face image sequences and internal expression, the number of consecutive frames to get updated motion parameters b constitute the dynamic characteristics, T_a represents the change in the motion parameters, that is, on behalf of the intensity of expression motor unit, so each frame to get a representative expression vector motion unit T_a motion tracking features of the model in Candide3 each face image sequences extracted to can be expressed as follows:

$$f = [T_a(1)^T, T_a(2)^T, T_a(3)^T, \dots, T_a(L-1)^T, T_a(L)^T]^T \quad (5)$$

Wherein, L represents the length of the image sequence. Since we use the seven sports unit, so you can come and f is a $L \times 7$ dimensional feature matrix.

Algorithm for one frame of image sequences were localized, seven motion parameters extracted motion feature, the tracking results of the experiment shown in Fig. (4):

Based on the dynamic characteristics Candide3 model parameter extraction method is advantageous in that [13]:

(1) At the same time face tracking, the characteristic parameters can be extracted directly, not as ASM algorithm as geometric feature points tracing calculations.

(2) That the model is not affected head rotation, motion parameters characterizing changes in the expression of the human face attitude change has a strong robustness.

(3) The model can be updated only by the number of low-dimensional motion parameters can reflect changes in facial expression, so the extracted image in each frame in the sequence of motion parameters that reflect changes in the expression of the dynamic feature vectors.

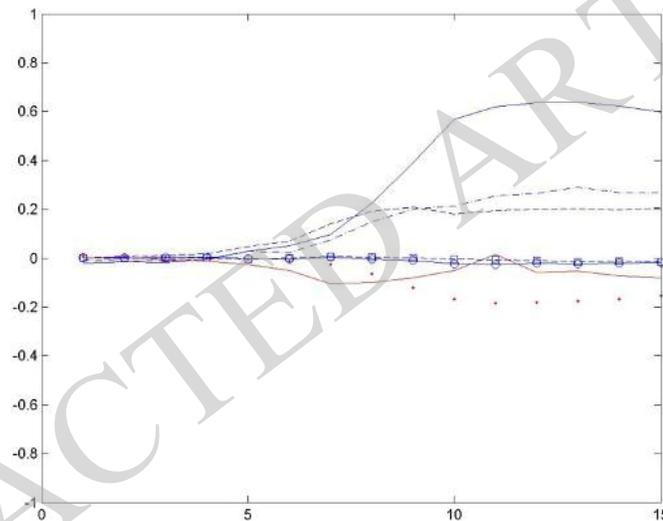
4. DYNAMIC TIME WARPING (DTW)

Dynamic image sequence chronologically Gallery expression facial expression images will line up, but even with a look that everyone will not have time to complete each exactly the same as the length of a sequence of images in the gallery of course, will not do exactly the same, which led directly we acquired the time corresponding to the range of features is not the same. In classifier usually requires a unified feature dimension, which requires extracting a feature in a sequence of images having a uniform length, which requires the image sequence library according to the number of frames set normalized. In this paper, a dynamic time warping (DTW) algorithm to solve the problem of inconsistent image sequence length.

DTW algorithm is based on dynamic programming, non-linear time again normalized pattern matching algorithm used in gait recognition on time is very sensitive to changes



(a) Sequence image tracking



(b) Changes in seven motion parameters

Fig. (4). The dynamic characteristics of seven model parameters extraction.

in the pattern recognition problems [14]. The main idea is as follows: For the two different time ranges, using time warping function difference modeling them on the time axis, in order to eliminate the difference between two time ranges, DTW by varying one of a time axis, so that with the other a possible overlap.

We say DTW is based on dynamic programming, is that it will be a complex problem into the overall best for many simple local optimization problems. If we have a dynamic feature vector time series $A = \{a_1, a_2, \dots, a_i\}$, there is another one waiting to identify the sequence $B = \{b_1, b_2, \dots, b_j\}$, where $i \neq j$, then the DTW algorithm would need to find a time warping function, making the timeline sequence B j can be mapped to a non-linear sequence a timeline of i, and

minimize distortion. Assuming time warping function $C = \{c(1), c(2), \dots, c(N)\}$ (N is the path length), $c(n) = (i(n), j(n))$ represents an n-th matching points, the matching point is from the first sequence A i (n) a j-th feature vector and the sequence B (n) of feature vectors constitute the distance $d(a_{i(n)}, b_{j(n)})$ between the two feature vectors, which we call the local matching distance.

By constantly looking for local matching distance [15], to find a path through the path can be made by a weighted sum of the distances of all the matching point on this path of feature vectors is minimized:

$$D = \min \sum_{n=1}^N [d(a_{i(n)}, b_{j(n)})] \tag{6}$$

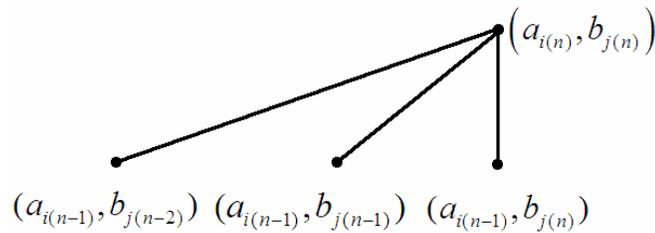


Fig. (5). A local path constraints.

Prior to the use of structured function to calculate the optimal path, the need for regular function plus constraint, otherwise inappropriate regular function might cause some problems. In our problem, when a moving image sequence expression library dynamically structured, regular function should satisfy the following two conditions:

(1) Continuity. Because the expression is a dynamic continuous process, in order to effectively preserve the characteristic information, to ensure the accuracy of identification, which requires regular function does not skip any match point in the sequence;

(2) Limit the monotony. Obviously the image sequence expression is calm over time by the state gradually to the climax of the state, which requires regular function to calculate the optimal route to get to keep the change of time, can not appear jump. Namely: $c_{i+1} \geq c_i$.

To achieve these two conditions, in the path of the selected time necessary to design appropriate constraints, the path you first need to meet the requirements of continuity and monotonicity. Can also be based on the actual situation, add different local path constraints, Fig. (5) represents a path constraint.

4.1. Based on Design Process and Production Process of Product Components Decomposition

The current ceramic enterprises in ceramic product design process and production process of detailed study, the design process of ceramic products for further analysis and decomposition, Exploring the project design of ceramic needs human interaction steps in the 3d CAD system, and needs to be done by the system automatically in order to better improve enterprise's key steps in the efficiency of product design and in the decomposition on the basis of the design process to classify modeling of ceramic products, and finished components decomposition of complex products. Ceramic products decomposition is different from the mechanical parts and components industry and other industry products, ceramic products decomposition lies mainly in the design of the components in the process of decomposition, and the final product in general is not an integral whole, for example, in the design process for more complex products such as "pot" will be broken down parts into the pot body, the pot, a spout, the lid and lid knob, a girder of the pot, an ear piece [10]. And set up all kinds of parts in the three-dimensional model of the material library.

As shown in Fig. (6), figure, a solid line indicates a complete path, this path has a continuous and monotonic characteristic, satisfies the constraint path shown in Fig. (6).

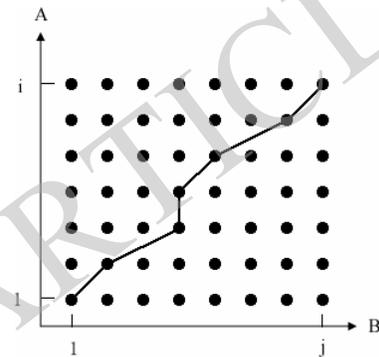


Fig. (6). The DTW search path.

DTW algorithm searches the following methods will be about the best path description:

For path constraints Fig. (5), the point $(a_{i(n)}, b_{j(n)})$ before the grid is one of only three points below where:

$$(a_{i(n-1)}, b_{j(n)}), (a_{i(n-1)}, b_{j(n-1)}) \text{ or } (a_{i(n-1)}, b_{j(n-2)})$$

Then $(a_{i(n)}, b_{j(n)})$ will select the minimum distance to the point of these three points as a preamble to the grid, the cumulative distance of:

$$D(a_{i(n)}, b_{j(n)}) = d(a_{i(n)}, b_{j(n)}) + \min(D(a_{i(n)}, b_{j(n)}), (a_{i(n)}, b_{j(n-1)}), (a_{i(n)}, b_{j(n-2)})) \tag{7}$$

With the above method of calculation can be from (1,1) to start the search starting, through repeated recurrence until the optimal path.

DTW algorithm principle is very simple, but because the algorithm in finding the best path back to the time required for repeated recursive, resulting in the amount of computation becomes large, a great impact on the computational efficiency. By analyzing its causes can be found in that search space is too large, too much of various branch path, and many of the searched path is often not required. Therefore, we consider adding a constraint for the global path selection,

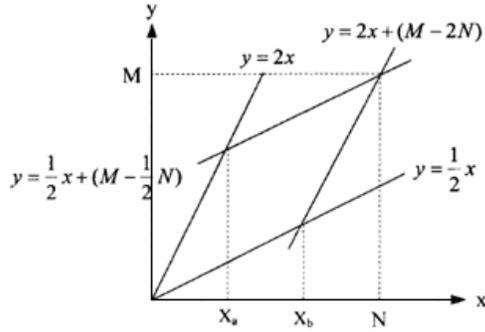


Fig. (7). DTW global path constraint.

article selects a slope in the range of constraint 1/2 to 2 (Fig. 7), if the slope is too large, the search path will end prematurely, do the aim is both to ensure the full path can be searched, you can reduce the mismatch, the most important thing is given a certain range of the search path, you can make a significant reduction in the amount of calculation.

When $x_a = x_b$ is relatively divided into two segments:

$$\begin{cases} \frac{1}{2}x \sim 2x, & x < x_a \\ 2x + (M - 2N) \sim \frac{1}{2}x + (M - \frac{1}{2}N), & x > x_a \end{cases} \quad (8)$$

When $x_a < x_b$, relatively divided into the following three sections:

$$\begin{cases} \frac{1}{2}x \sim 2x, & x < x_a \\ \frac{1}{2}x \sim 2x + (M - \frac{1}{2}N), & x_a < x \leq x_b \\ 2x + (M - 2N) \sim \frac{1}{2}x + (M - \frac{1}{2}N), & x > x_b \end{cases} \quad (9)$$

DTW algorithm also has the following two characteristics: first, the calculation process can be seen as a cycle of cumulative matrix generated problems; secondly, DTW algorithm based on dynamic programming, in the accumulation of matrix generation process, each point is calculated only with a number of points concerning this point before. Therefore, when using DTW algorithm does not require calculation Fig. (7) point diamond outside, but do not have to keep a distance and accumulate matrix matching distance matrix and only needs to be saved locally in the calculation process can be constantly updated.

According to the previous analysis, the use of DTW algorithm to calculate the cumulative distance of the next frame in the X-axis direction, just in front of a column of accumulated distance, so in the implementation process of the algorithm, we do not need to save the entire matrix, and only need two D and d variables to save the current column and on a cumulative distance. When performing the variable D is

the last element of the test template and reference template matching between the template when the last frame of the test distance. By this method, can greatly reduce the amount of storage space and calculation, thereby increasing the speed of recognition.

5. FEATURE SELECTION

Feature selection is the key feature of useful features to choose from, all of them, the removal of the original features of the redundant features, thus leaving the greatest contribution to the classification features.

Feature by feature selection algorithm to select and then perform pattern classification, can bring many benefits to the pattern recognition system:

- Reduce the feature dimension;
- Reduce the data acquisition time;
- Reduction of time training the classifier;
- Improve the recognition rate of the classifier.

In face recognition based on image sequences, the expression of different frames in an image, for feature extraction expressions used are not the same, so we want to extract all the features, select the most advantageous feature classification, feature selection is to complete this task.

5.1. Selection Based on Fisher Criterion

The main task is to remove the section of each frame to extract the motion features of the non-redundant information related to that feature selection. Using a single feature of Fisher discriminant method, after the screening of the motion feature to get a group of sub-optimal subset of features, so you can remove poor classification performance characteristics.

Within the feature class scatter smaller, the inter-class scatter, the stronger its categorical, Fisher criterion [10] is based on this idea for feature extraction. In this paper, a single feature of the Fisher discriminant rate as the criterion, the basis for calculating the value of each feature, and then ordered from these features, select a category and strong features, remove the classification of weak features, so as to achieve a better face recognition results .

The definition of the training set there are n samples that belong to class C: w_1, w_2, \dots, w_c samples per class contains n_i ,

μ_i represents the mean of the i-type samples. The formula for calculating the dispersion of the between-class and within-class scatter as follows:

$$S_b = \sum_{i=1}^c P_i (\mu_i - \mu_0) (\mu_i - \mu_0)^T \quad (10)$$

Wherein, μ_0 is the global mean vector, $\mu_0 = \sum_{i=1}^c P_i \mu_i$, $\{S_b\}$ is a measure of the average distance between the mean of each class and the global mean.

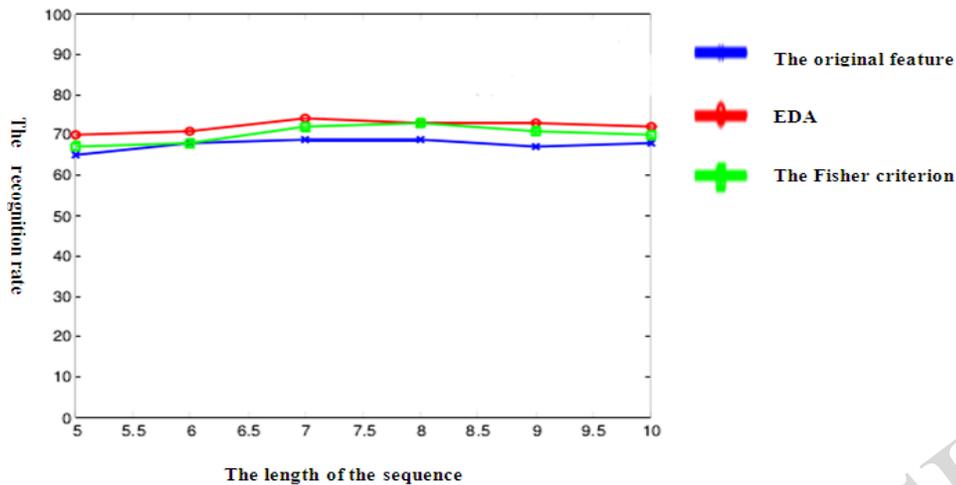


Fig. (8). Identification of different image sequence length ratio.

$$S_w = \sum_{i=1}^C P_i S_i \quad (11)$$

Wherein, S_i is the covariance matrix of class $S_i = E[(\mu_i - \mu_0)(\mu_i - \mu_0)^T]$, P_i is the a priori probability of w_i , $w_i \approx n_i/n$, S_w is an average measure of the variance of the characteristics of all the classes.

5.2. Estimation of Distribution Algorithm Based on Feature Selection

Estimation of Distribution Algorithm (EDA) [15] is the field of evolutionary computation research frontier is being proposed in 1996, which proposed a new algorithm evolutionary algorithm improves the traditional evolutionary genetic algorithm (GA).

In GA, in order to optimize the solution of the problem of the candidate, they are represented by the population, the population of each individual has their own fitness, and then simulate the natural evolutionary process, carried out other options, crossover and mutation, so repeatedly solving complex operation. In EDA, the longer the traditional selection, crossover and mutation of such genetic manipulation, but the probability model for sampling and study, EDA probabilistic model to describe the spatial distribution of candidate solutions, from a macro point of view of the statistics of the candidate solution probability distribution, and the establishment of the model, based on the probability model, a random sampling of the population get a new, EDA is the use of such methods continue to be repeated to achieve the evolution of the population, until the solution of the problem is obtained.

In the feature selection process EDA, we use the EDA algorithm is relatively simple models - learning (Population-based Increased Learning, PBIL) algorithm based on the incremental population.

6. SIMULATION AND ANALYSIS

6.1. Motion Feature Extraction Based on Active Appearance Model

Motion feature extraction facial expression in the sequence of images, different sequence lengths will bring a different feature dimensions, the first sequence of images in order to compare the impact of different lengths of expression recognition in image sequences for different lengths of facial expressions after the feature points localization, we use a different feature of the sequence of images, respectively, of different lengths were extracted and the motion information of the feature point recognition result shown in Fig. (8).

As can be seen from Fig. (8) the identification rate, and not using a feature selection algorithm to improve the recognition rate greatly, because the feature points based on geometric features for classification contributions are more evenly. We found in the image when the sequence length of 7 experiments achieved the highest recognition rate. It is worth mentioning that, with the increase in the recognition rate in the sequence length instead reduced, because feature points are extracted when the ASM, the positioning of each frame and absolute accuracy is not enough, even after affine arithmetic, may still be pixel shift, which is extracted geometric features will result in errors in the calculation of the motion feature will feature point position error as the geometrical distance change is counted together, and with the increase of the number of frames, expression changes between frames become smaller with increasing dimension of feature vectors, and will bring some redundant information is not conducive to the classification, feature point extraction truly reflect the movement of information will become more difficult to influence the final recognition rate. Conversely, when the sequence length is short, the positioning error of the facial feature points for each frame geometric motion feature extraction will greatly reduce the impact, as shown in Fig. (9).



Fig. (9). The 6 frame positioning effect expressions in image sequences.

Table 2. The recognition results.

Test Sample		Correct Identification of Samples	
Expression	Test Set	Correct Recognition	Recognition Rate
Angry	15	11	73%
Hate	15	10	67%
Fear	15	12	80%
Happy	15	13	87%
Sadness	15	9	60%
Surprised	15	12	80%
Average	90	67	74%

Table 2 shows the image sequence 7, the ASM algorithm with facial expression motion feature extraction and use of estimation of distribution algorithms for feature selection, support vector machine classifier to obtain recognition results of each test set expression. As can be seen from the results, the method for tracking the movement of feature points, for the pleasure and surprise of this relatively exaggerated facial expressions, identify more accurate, but the magnitude of change in the operation of a relatively small effect on the expression recognition is not ideal, this is because this is because the method is based on the limitations of face tracking feature points lead.

6.2. Feature Extraction Based on Dynamic Models Candidate3

The same sequence of still images to different lengths to extract motion parameters affect the expression of Candidate3 model experiments, for different lengths of sequence image feature points of facial expression to locate, respectively, using the six model parameters and seven sports model

motion parameters for image sequences of different lengths were extracted expression motion motion information unit. By comparing the use of ASM motion feature extraction algorithm, we found Candidate3 model based on a more accurate face tracking, and motion parameters as characteristic facial feature is not directly on point positioning obtained, it is possible in the case of a longer sequence length, got more conducive to expression feature classified information to improve the recognition rate. Fig. (10) is a picture of the 10 track results Candidate3 models.

Fig. (11) different classifier 9 image to extract the seven models utilize motion parameter estimation of distribution algorithms for feature selection. From the figure can be seen in different image frames, for different classifiers, the contribution of each frame of the motion parameters of the classification is different.

CONCLUSION

This paper describes the extraction based on dynamic image sequence characteristic facial expression and active



Fig. (10). The model of Candide3 10 image frames for tracking results.

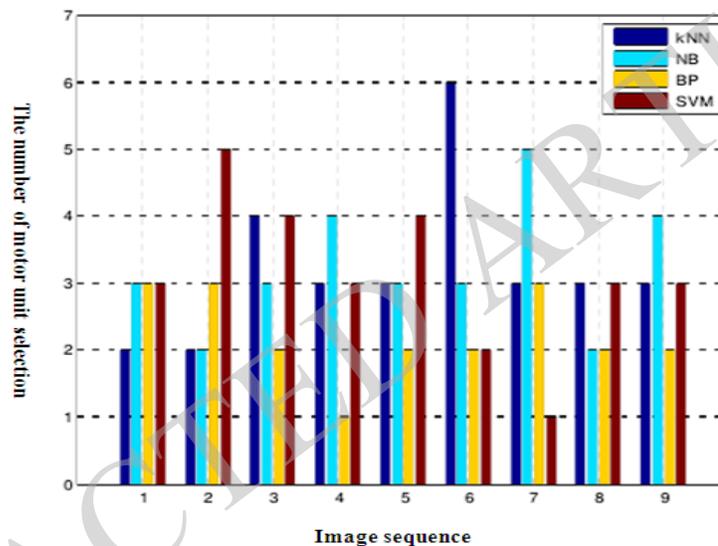


Fig. (11). Extraction of motion parameters of each frame image.

appearance model Candide3 model algorithms, and the two algorithms were used to sequence the human facial expression image extracted motion feature, Active Appearance Model method of calculating speed. However, given the feature points is not very accurate, has some influence on the results. Candide3 feature extraction method based on the dynamic model is the model for the extraction of changes in the expression of motion parameters, more accurate motion information extraction. While for dynamic image sequences of varying sequence length feature dimension cause problems ranging training and recognition, and were normalized using an improved method of dynamic time warping. Then feature extraction is performed on the motion feature selection, each frame image by selecting the classification ability of the characteristics of the final classification, can be seen from the experimental results, a different image sequence length different motion feature extraction algorithm will bring different effects, but also through the feature selection algorithm to select the feature vector will improve the recognition rate of the algorithm.

CONFLICT OF INTEREST

The authors confirm that this article content has no conflict of interest.

ACKNOWLEDGEMENTS

This work was financially supported by the Henan Science and Technology Key Project (122102210563 and 132102210215).

REFERENCES

- [1] P. Ekman, W. V. Friesen, *Facial Action Coding System: a Technique for the Measurement of Facial Movement*, CA: Wadsworth, 1978, pp.120-135.
- [2] M.H. Yang, D. J. Kriegman, and N. Ahuja, "Detecting faces in images: a survey," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 7, pp. 156-165, 2002.
- [3] Y. Xue, G. Hair, G. Ye, and S. Lv, "Recognition of human-computer interaction in the facial expression progress," *Chinese Journal of Image and Graphics*, vol. 5, pp. 764-772, 2009.

- [4] T.F. Cootes, A. Hill, and C.J. Taylor, "The use of active shape models for locating structures in medical images," *Image and Vision Computing*, vol. 12, no. 6, pp. 355-366, 1994.
- [5] C. Peng, S. Liu, C. Wan, and T. Tao, "ASM local texture model based on facial expression recognition," *Journal of Intelligent Systems*, vol. 3, pp. 231-238, 2011.
- [6] T. Kanade, J. Cohn, and Y. Tian, "Comprehensive database for facial expression analysis," In: *Proceedings of IEEE international Conference on Face and Gesture Recognition*, New York, 2000, pp. 46-53.
- [7] H. Kobayashi, and F. Hara, "Facial interaction between animated 3D face robot and human beings," In: *Proceedings of IEEE International Conference on System, Man and Cybernetics*, Orlando, 1997, pp. 3732-3733.
- [8] J.J.J. Lien, *Automatic Recognition of Facial Expressions Using Hidden Markov Models And Estimation of Expression Intensity*. Pittsburgh: Carnegie Mellon University, 1998.
- [9] Z. Ji, M. Hu, and J. Yin, "Summary of feature selection algorithm," *Electronic Design Engineering*, vol. 9, pp. 46-51, 2011.
- [10] S. Wang, and L. Zheng, "Based on fisher criterion and feature clustering," *Computer Applications*, vol. 27, no.11, pp. 2812-2840, 2007.
- [11] W.S. Yao, W.L. Chen, Xu Ye, "Estimated distribution algorithm research progress," *Control and Decision*, vol. 7, pp. 961-966-974, 2012.
- [12] Z. De, and S.Z. Qi, "Estimation of distribution algorithms summary," *Institute of Automation*, vol. 2, pp. 113-124, 2007.
- [13] Q. Zheng, and X. Lv, "Integrated classifier based on multi-feature facial expression recognition," *Journal of China University of Petroleum (Natural Science)*, vol. 1, pp. 174-178, 2011.
- [14] Y. Zhang, *Face Recognition Based on Subspace*, Beijing: Tsinghua University Press, 2009, pp. 50-92.
- [15] N. Wang, X. Li, and G. Liu, "Local feature-based subspace neighborhood keep manifold learning algorithms," *Application Research of Computers*, vol. 4, pp. 1318-1321, 2012.

Received: September 16, 2014

Revised: December 23, 2014

Accepted: December 31, 2014

© Xinfeng and Shan; Licensee Bentham Open.

This is an open access article licensed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/3.0/>) which permits unrestricted, non-commercial use, distribution and reproduction in any medium, provided the work is properly cited.

RETRACTED ARTICLE