Facial expression recognition based on ASM and Finite Automata Machine

Jiaxin Si [1, a], Yanchen Wan [1, b], Teng Zhang [1, c]

[1]Department of Software Engineering, Jilin University, Changchun, 130000, China

[a]email: sjxmysky@sina.com, [b]email:491946376@qq.com, [b]email: 893696196@qq.com

Keywords: ASM; Finite Automata Machine; Adaboost

Abstract. According to the problem that the facial expression recognition algorithm is easy to be affected by the error of the feature detection, this paper proposes a new algorithm based on Finite Automata Machine and realize the complete facial expression recognition system. The implementation steps include image preprocessing, face detection based on Adaboost, facial features localization based on ASM, normalization and facial expression classification based on Finite Automata Machine. Then experimental results show that this new facial recognition system possesses high accuracy, high stability and fast running speed. This provides a new method for the improvement of facial expression recognition, which promotes the development of artificial intelligence and human-computer interaction as well.

Introduction

The rapid development of Human-Computer Interaction and Artificial Intelligence has put forward higher demand to the intelligence and humanity of computer. As the main part of daily interaction, the expression can convey the human's psychological and emotional information. In order to make the human-computer interaction more intelligent, the facial expression recognition based on image has become a new research hotspot. This technology can be widely used in the field of business, medical care, education, transportation and other industries, and it will bring huge social benefits and economic benefits [1].

At present, the research of facial expression recognition has yielded a rich harvest. However, most of algorithms which use the facial features to classify the expression are easy to be affected by the local errors when the features are detected. In order to solve this problem, this paper proposes a new expression classification algorithm based on Finite Automata Machine, which specifies the criterion of expression judgment to reduce this interference. Meanwhile it separates out the features which are easy to detect to speed up the process of expression judgment.

In this paper, the realization of the optimization of facial expression recognition system is introduced, including the traditional image preprocessing, the face detection based on Adaboost, facial features localization based on ASM; normalization and expression classification based on Finite Automata Machine. Then the optimized expression classification algorithm is described in detail and the superiority and accuracy of it are proved by the experimental results.

Image preprocessing

The original image has lots of interference information with rich color. The conversion from RGB color image to grayscale can not only keep the luminance and chrominance, but also reduce computation load [2]. And histogram equalization makes the intensities better distributed on the histogram, which allows for areas of lower local contrast to gain a higher contrast [3]. Through pre-processing, images become sharper and it lays the foundation for the future processing.

The face detection based on Adaboost Algorithm

The Adaboost algorithm is divided into two parts: the training part and the detection part.

Firstly, Haar features are extracted from the sample images, and the integral graph is used to calculate the eigenvalues. Then, the number and value of each Haar feature are transformed into the corresponding Weak Classifier. Finally, all the weak classifiers are combined to form strong classifiers, and then a number of strong classifiers are graded in series to form a cascade classifier for face detection, which greatly eliminate the error rate [4].

$$h_j(x) = \begin{cases} 1, p_j f_j(x) < p_j \theta_j \\ 0, else \end{cases} \tag{1}$$

In the formula (1), $h_j(x)$ is the value the simple classifier. θ_j is the threshold. p_j is the direction of the inequality sign, and $f_j(x)$ is the eigenvalue.

The facial features localization based on ASM

The facial feature extraction based on ASM includes training and searching.

Firstly, the ASM's training part is executed. N frames of facial images are used as training samples, and K features are calibrated in sequence on each one. Then a base vector is selected, and other vectors are rotated, scaled and translated to correspond with it as possible as they can [5]. Finally, PCA algorithm is used to analyze and reduce dimension of the normalized image, and the local gray model of feature is established.

In the training process, the coordinate of each feature is converted as follows:

$$X^{'} = \frac{1}{n}\sum_{1}^{n}(M(S,\theta)[X]+t)+Pb \tag{2}$$

In the formula (2), X is the original coordinate. $X^{'}$ is the transferred coordinate. $M(S,\theta)[X]$ represents that the rotation angle of X is θ and the scale of X is S . t is the translation vector. $P=(p_1 p_2 ... p_k)$ is the covariance matrix of former K feature vectors, and $b=(b_1 b_2 ... b_k)$ is weight vectors.

When an unlabeled facial image is input, the ASM's searching section is employed to locate the features. Firstly, the initial model is established, and the average shape model is covered on the face contour detected by Adaboost. Then the new coordinate of an optimal shape model is calculated by searching and matching in the examination image. At last, the parameter values are updated iteratively by using the parameter formula until the shape model is not changed obviously.

Searching process is to find the smallest Mahalanobis distance of each feature [6]:

$$d(x, T(\overline{x}+Pb)) = (T(\overline{x}+Pb)-x)^T \sum{}^{-1}(T(\overline{x}+Pb)-x) \tag{3}$$

In the formula (3), d is the Mahalanobis distance. $T(\overline{x}+Pb)$ is the new coordinate of X in the adjustment process. $\sum$ is covariance matrix.

The expression classification based on Finite Automata Machine

Finite Automata Machine describes the state sequence of an object in its life cycle and how it responds to a variety of events from the outside. A FAM can be described as follows:

$M(S, \Sigma, \delta, S_0, F)$

S is the finite state sets. Σ is the sets of input elements. δ is the state transition function. S_0 is the initial state, F is the termination status set [7].

The following part will introduce how to use the Finite Automata Machine (FAM) to realize the facial expression recognition which can reduce the influence of local feature detection errors.

The selection of the state (S):

The states of the FAM in this paper are divided into two categories: the first one is the six major categories of human expression (happiness ,sadness ,anger ,surprise ,disgust ,fear), and the second one is the intermediate state between the different facial expression state, which is established when there are some obvious changes in the feature in the process of expression conversion. For example, in the conversion from neutral expression to happiness, the intermediate state that the mouth opens wider is set up (Figure 1).These states are generally not easy to make mistakes ,and their existence is conducive to jumping out of the state of expression timely. Moreover, they can avoid the repeated computation of each expression recognition process which always begins from the initial state.

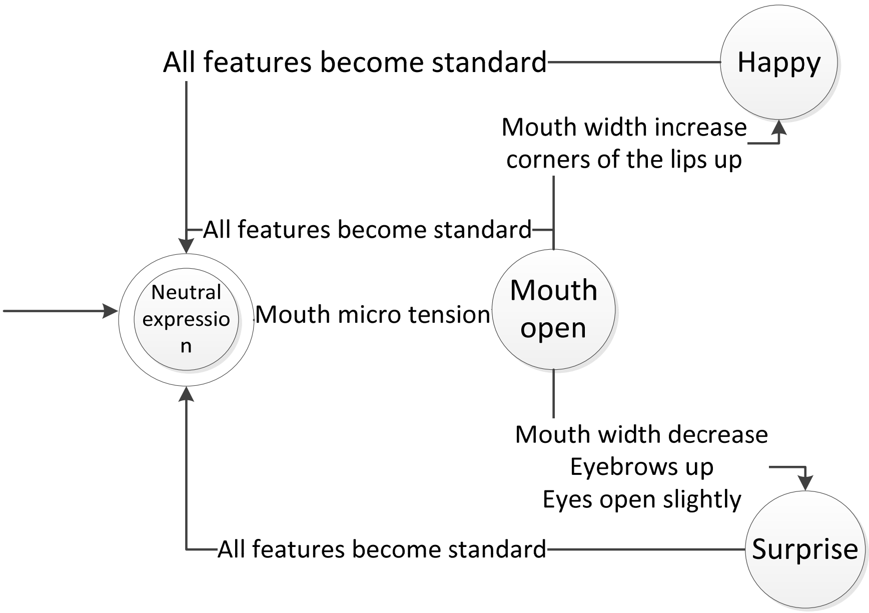

Figure 1.The example of part of Automata

Input elements (Σ):

After obtaining the normalized facial expressions features by ASM, we can determine expression by the different position relations of feature points. For example, the distance between the upper and lower lip in the vertical direction can be employed to identify whether the mouth is open. In this paper, a number of facial features are used in this way to track and identify facial expressions:

To describe the input, $E(a_0, a_1, ..., a_{12})$ is defined as a feature set of each facial frame ,and E_0 is the standard facial feature set which is composed of neutral expression states. $X = E - E_0$ is the input element which eliminates lots of man-made interference.

Initial state and terminal state (S_0 , F)

In this system, initial state S_0 and terminal state F are both set as neutral expression. For S_0 , this is beneficial to confirm the user's standard expression feature set and simplify the program. For F , returning to terminal state (neutral expression) periodically to input current expression can update the set of standard facial states regularly. And if expression states are used as the terminal states, it will not have the advantages described above and may fall into an endless loop.

State transition function (δ)

The state transition function is used to determine whether the input elements can transfer the current state to another one. In this paper, the state transition condition of the FAM is defined as Y , which is measured by expression feature incremental set. That the input element X can meet Y means that for each x_i , if it isn't in the standard fluctuation (this value depends on the quality of the feature point detection and here it is ± 2),then $|x_i| >= |y_i|$ and $x_i \times y_i > 0$.

Automata algorithm

The program accepts the facial features of the current frame and it goes through all outputs of current state to judge whether the input element meets the demand of each edge. If it satisfied the condition, then the next state is achieved and the same steps are performed until entering an expression state. Here the input sets are required not to meet the multiple outputs of the same state at the same time.

```
Input:       Facial feature set of current frame X
Initialize:  stateNum（the sequence number of current status）=neutral expression
             $\text{stateC}_{i,j}$   A set of features to be met by the conversion between state $i$ and state $j$:
Setp1:  foreach  state $\in S$  then
           If（$\text{stateC}_{stateNum,state}$ ！=NULL）then
             foreach  $x_i \in X$ , $y_i \in \text{stateC}_{stateNum,state}$  then
                if（$y_i \notin [-2,2]$ and $|x_i| >= |y_i|$ and $x_i \times y_i > 0$）then
                     stateNum=state
                     goto step2;
                endif
        endif
Setp2:  If (stateNum has no change||state==expression state) then
             goto step3;
        else
             goto step1;
Step3:  If (stateNum is the expression state)
        Output the expression expressed by this state.
```

Test analysis and results

We choose the change of four expression feature increments--Mouth vertical size (MouthVerSize), mouth width (MouthWidth), right mouth Corner (RightMouthCorner) and left mouth height (LeftMouth Corner) in 40 frames to test our system, and the results are showed in the graph below.

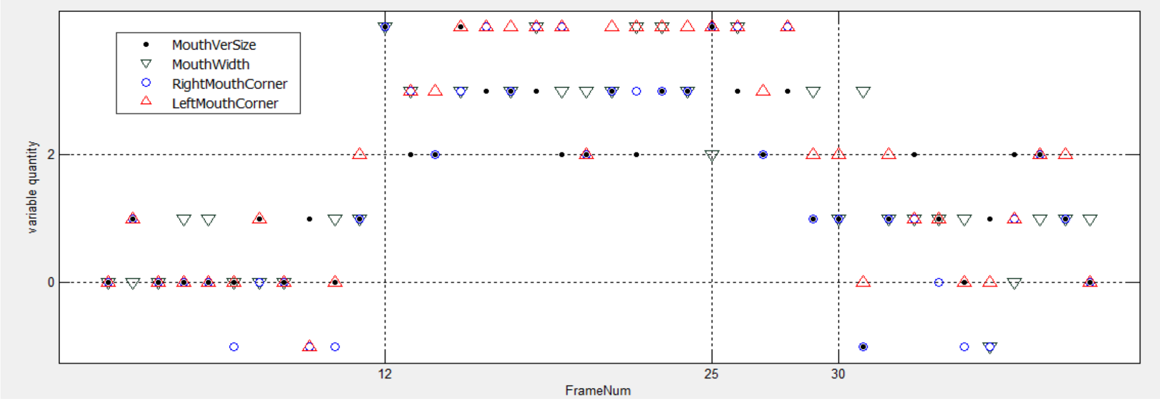

Figure 2 The change of the facial expression feature increment

In accordance with the algorithm above, the expression of 12th frame is judged as happiness, then it stays on this state. Unexpectedly in the 25th frame, the width of the mouth is below the expectation because of the feature detection error. However this error will not change the state, and this is the obvious advantage of this algorithm in contrast to others. Therefore, FAM algorithm in this paper can avoid specific error detection.

In this paper, we only use neutral expression, happiness and surprise to show the results of the test.

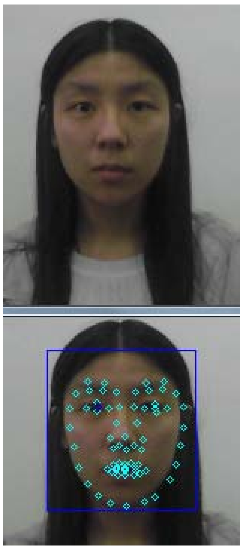

Figure 3 The output of neutral expression

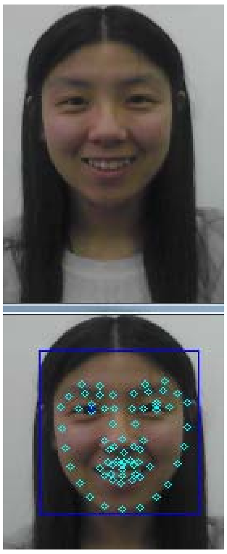

Figure 4 The output of happiness

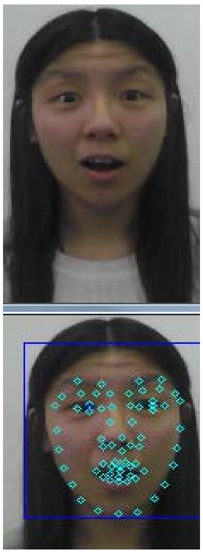

Figure 5 The output of surprise

Based on the experiment data, the correct rate of six expressions are shown as follows:

Table 1 Test result statistics

Expression	number of test/times	number of errors/times	average accuracy
Happiness	100	6	94%
Surprise	250	20	92%
Sadness	150	15	90%
Anger	200	16	92%
Disgust	100	9	91%
Fear	150	12	92%

Conclusion

In this system, on the basis of the traditional image preprocessing, face detection using Adaboost algorithm and the facial features localization with ASM, a new expression recognition algorithm which changes the automaton according to the different changes of features in different expressions is proposed. According to experiment data, it can effectively eliminate the feature detection error, and has high recognition reliability on the continuous expressions with lower time complexity.

References

[1] Lv Zhen.Study on the video-based facial expression recognition technology [D]: Qingdao:2013. 1-5

[2] Xu, Yong; Zhang, Zheng; Lu, Guangming . Approximately symmetrical face images for image preprocessing in face recognition and sparse representation based classification [J], 2016 (54)68-82

[3] Cadik, M. 16th Pacific Conference on Computer Graphics and Applications: Perceptual Evaluation of Color-to-Grayscale Image Conversions [C]. Tokyo: JAPAN, 2008. 27(7)1745-1754.

[4] Zuo Dengyu. A Study of Face Detection Based on Adaboost Algorithm [D]: Beijing, University of Science and Technology of China: 2009:23-36

[5] Sun Yanming. Automatic Facial Feature Points Localization and Application Based on ASM [D].Beijing, Beijing Jiaotong University, 2010:19-29

[6] You Lingli. The video-based facial expression recognition [D].Nanjing: Nanjing University of Science and Technology, 2011:8-22

[7] Zhang Chen; Duan Zhenhua;Tian Cong. Specification and Verification of UML2.0 Sequence Diagrams Based on Event Deterministic Finite Automata [J]. Journal of Software, 2011, 22 (11): 2625-2638.