

FACIAL EXPRESSION RECOGNITION BASED ON IMPROVED LBP OPERATOR AND K-MEANS CLUSTERING

Wang Yunfei ^{1, a}, Ding Hui ^{1,2,b *}, Liu Yi ^{1,c} and Pan Yanyan ^{1,d}

¹ College of Information Engineering, Capital Normal University, Beijing 100048, China

² Collaborative Innovation Center for Mathematics and Information of Beijing, Capital Normal University, Beijing 100048, China

^awangyunfeichn@sina.com, ^bdinghui@ie.cnu.edu.cn, ^cliuyi@ie.cnu.edu.cn, ^dpanyanyan@ie.cnu.edu.cn

Keywords: Feature Extraction, Local Binary Patterns Operator, K-means Clustering, Facial Expression Recognition.

Abstract. For facial expression features extraction and automatic clustering, a method is proposed. This method integrates improved LBP operator and K-means clustering. Firstly, facial features are located by integral and variance projection. Improved LBP operator with changed threshold is figured out. Secondly, K-means clustering produces expression templates. Then K nearest neighbor algorithm recognizes facial expression. Meanwhile, the expression templates will be updated by the result itself, optimizing expression templates dynamically. The algorithm is simple and low computational-complexity. It updates the train results dynamically. Results show that this classification method is accurate.

Introduction

Early face recognition study and analysis focused on psychology and biology. Now they have been applied in fields such as human-computer interaction, artificial intelligence, intelligent monitor, medical diagnosis and social entertainments for large development space and application value.

There are two key-technologies for face recognition: (1) feature extraction. (2) Facial expression classification^[1]. For still images, feature extraction methods include geometric and appearance extraction. Geometric extraction includes Gabor Wavelet transformation algorithm^[2]. Gabor Wavelet transformation is restructured to obtain the feature matrix; Features generated by scale-invariant feature transformation (SIFT) algorithm^[3] have remarkable robustness. But the algorithms above need massive computation. Feature extraction includes Pairwise Reinforcement of Feature Responses (PRFR)^[4], integral projection and variance projection method, etc. The former algorithm extracts the a face edge. The latter extracts facial features' position.

There are several common expression classification methods. Support Vector Machine (SVM) works by training samples^[2], Improved artificial neural network uses neural network to recognize expression^[5], K-means clustering is based on Euclidean distance^[6]. It sorts discrete points, facial expressions, etc. Compared with former methods, K-means clustering is more efficient^[7].

Integral projection and variance projection are combined to locate facial features' position. Then classification is accomplished through improved LBP operator and K-means clustering.

Facial Features' Location Based On Distribution.

Human face obeys oriental face proportion. It means that in vertical direction, a face has the same width of three chambers. And in horizontal direction, a face has the same height of five eyes.

Eyes' location should be accomplished first. Accurate eye location is helpful for following process. Facial features can be normalized better. And more notable preprocess results will be obtained.

Integral and variance projection. Two rules can be found after taking analysis of intensity feature and intensity gradient statics. First, only in eye region, intensity distribution varies severely^[8]. The

eye-region intensity value is low^[9]. Second, the region around eye is bright. So the intensity variance is notable. Taking them into account^[10], eye region can be located quickly and accuracy.

Integral projection is a statistical method based on face proportion. When light changes, the intensity distribution becomes uneven. So the projection result is normalized.

Let the image size be $M \times N$, value of (x, y) is $Y(x, y)$. In discrete context, it is defined in Eq. 1.

$$U_x(x, y) = \sum_{x=1}^M Y(x, y), \quad U_y(x, y) = \sum_{y=1}^N Y(x, y) \quad (1)$$

Integral projection can't show intensity variation where intensities are close. The problem can be solved with variance projection. The variance projection is computed by Eq. 2.

$$S_v^2(x, y) = \sum_{y=y_1}^{y_2} [Y(x, y) - M_v(x)]^2 / (y_2 - y_1), \quad S_h^2(x, y) = \sum_{x=x_1}^{x_2} [Y(x, y) - M_v(y)]^2 / (x_2 - x_1) \quad (2)$$

where $S_v^2(x), S_h^2(y)$ are the variance projection among $[x_1, x_2]$ and $[y_1, y_2]$ in vertical direction and horizontal direction. $M_v(x), M_v(y)$ represent the average value in row and in column.

To simplify the calculation, the variance projection will be normalized^[11] by Eq. 3 and Eq. 4.

$$M_v(x) = (n-1) * (M_v(x) - M_{\min}) / (M_{\max} - M_{\min}) - 1 \quad (3)$$

$$S(x) = (n-1) * (S(x) - S_{\min}) / (S_{\max} - S_{\min}) - 1 \quad (4)$$

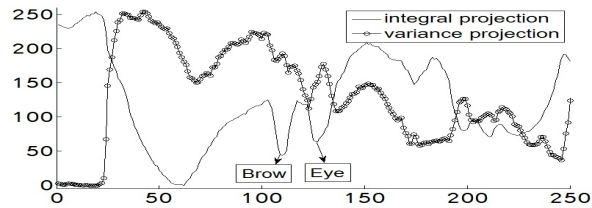
Where max, min represents the maximum, minimum values. $M_{\max} = \max(M_v(1), M_v(2), \dots, M_v(m))$, $M_{\min} = \min(M_v(1), M_v(2), \dots, M_v(m))$, $S_{\max} = \max(S(1), S(2), \dots, S(m))$, $S_{\min} = \min(S(1), S(2), \dots, S(m))$

Figure 1 shows that eye region intensity is lower. And eye region lies in the integral projection trough. For severe changes in eye region, it is in the variance projection crest (Fig. 1(b)). Then eye location can be confirmed by facial feature proportion (Fig. 1(b)).

Eye need to be located accuracy further. Troughs on integral projection graph and crests on variance projection graph have a top-down arrangement, so the first trough-crest meet position is eye's vertical position.



(a) Expression image.



(b) Integral and Variance projection

Fig. 1. Integral and variance projection graph

In integral projection graph, eye and brow are two continuous trough. After eye position being confirmed, the first trough over the eye position must be brow position.

Location Of Nose And Mouth. According to a priori knowledge, the nose width is less than the distance between pupils. And in the vertical direction, the nose lies right below pupils' midpoint. With canny edge detection, nose can be located.

To find mouth, the face needs to be binarized. The operation result depends on how threshold is selected. Directly given threshold is lack of adaptability, so iterative method is used in threshold selection. The iterative method is defined in Eq. 5.

$$T_1 = \sum Y(x, y) \quad (Y(x, y) > T), \quad T_2 = \sum Y(x, y) \quad (Y(x, y) < T) \quad (5)$$

If $|T_1 - T_2| < \epsilon$, T is optimal threshold. Otherwise let $T = (T_1 + T_2) / 2$, repeat until $|T_1 - T_2| < \epsilon$. After close operation, small white areas are deleted. After contour extraction, a rectangular area goes

down along pupils' vertical line. It's the mouth position where the sum of intensity inside the rectangular area is smallest.

Facial Expression Recognition Based On K-means Clustering.

K-means clustering algorithm aims to partition n observations into k clusters in which each observation belongs to the cluster with the nearest mean. Before that, facial expression must be transformed into feature vector. Then texture features are linked to acquire feature vector.

Illumination Equalization. To eliminate illumination change, the image is normalized^[12]. In order to enlarge intensity change, transform and stretch are necessary. Color image is convert into gray-scale image. The image gray scale average can be computed in Eq.6.

$$Aver = \sum_{x=1, y=1}^{x=m, y=n} Y(x, y) / m * n \quad (6)$$

where m and n are the number of rows and columns. The quotient of average and standard deviation is figured out. The standard deviation and illumination equalization are computed in Eq. 7 and Eq. 8.

$$Dev = \sum_{x=0, y=0}^{x=m, y=n} ((Y(x, y) - Aver)^2 / (m * n))^{0.5} \quad (7)$$

$$F(x, y) = Y(x, y) - (Aver / Dev) \quad (8)$$

where $Y(x, y)$ is the intensity of (x, y) . $Aver$ is average intensity. Intensity equalization is $F(x, y)$. Then gray level stretch enlarges the intensity range. The gray level stretch is computed in Eq. 9.

$$G(x, y) = (F(x, y) - \min) / (\max - \min) * 255 \quad (9)$$

where $G(x, y)$ is stretch result in (x, y) . \max is the highest gray level, and \min is the lowest gray level.

After preprocess shadow caused by illumination, the gray level distribution is more homogeneous. Figure 2a shows the the original image. And Figure 2b shows the preprocessed image.

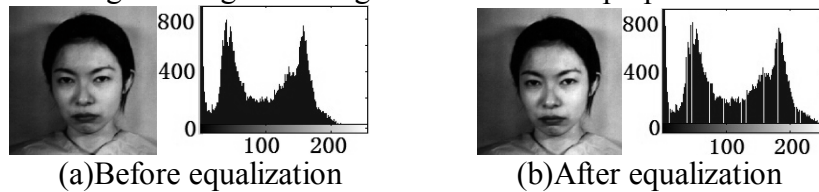


Fig. 2. Original image, preprocessed image and their histograms

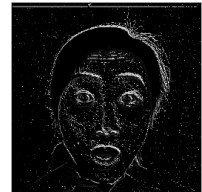
Improved LBP Operator Extraction. Local binary patterns (LBP) is the particular case of the texture spectrum model. The image is divided into cells. Each pixel in a cell is compared with its 8 neighbors. The result is an 8-digit binary number.

LBP contains local micro pattern, including edges, spots, etc. Illumination has little effect on LBP operator. Classic LBP operator detects unobvious change. Figure 3 shows an improved LBP operator. It writes "1" when the center pixel's value is 16 levels greater. Figure 3a is transformed into classic LBP operator (Fig. 3b, value is 248) and improved LBP operator (Fig. 3c, value is 192). The texture is more explicit after improvement. Figure 3d and 3e shows the effect of improvement.

51	78	33
99	80	81
87	86	85

0	0	0
1	80	1
1	1	1

0	0	0
1	80	0
1	0	0



(a)Original Image (b) Classic LBP (c) Improved LBP (d) Classic LBP (e)Improved LBP

Fig. 3. An LBP operator example and effect of improved LBP

Brows, eyes and mouth regions' LBP operators are concatenated as the feature vector.

Improved K-means Clustering Algorithm. K-means clustering algorithm can be described as follows: $X = \{x_1, x_2, \dots, x_i, \dots, x_n\}$ are n observations in R^d space. Given an original sample as centers of k categories, other observations are assigned to the nearest cluster. The distance between x_i and c_i is figured in Eq. 10.

$$d(x_i, c_j) = [(x_i^1 - c_j^1)^2 + \dots + (x_i^k - c_j^k)^2 + \dots + (x_i^d - c_j^d)^2]^{0.5} \quad (10)$$

The observations in category j is $\{x_{j_1}, x_{j_2}, \dots, x_{j_{n_j}}\}$. And n_j is the quantity of the cluster. The center is $c_j = (c_j^1, c_j^2, \dots, c_j^k, \dots, c_j^d)$, where c_j^k is the k th attribute. c_j^k can be acquired in Eq. 11.

$$c_j^k = (x_{j_1}^k + x_{j_2}^k + \dots + x_{j_{n_j}}^k) / n_j \quad (11)$$

Update continues until the cluster centers no longer change. Figure 4 shows the flow chart.

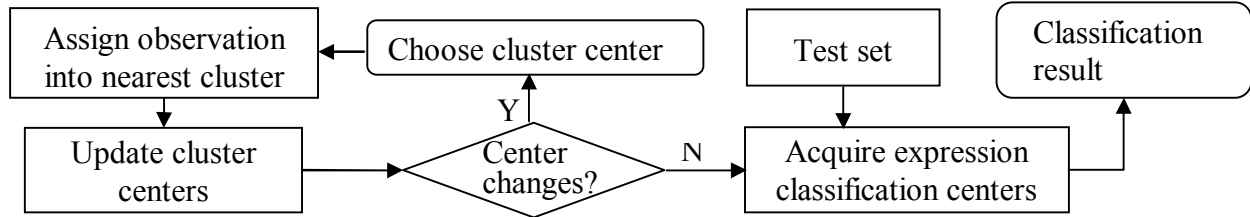


Fig. 4. Classic K-means clustering algorithm flow chart

After classification, cluster-center labels can't be acquired. So cluster centers are recorded. Then images in test set are compared with train set to find out nearest cluster center. The center's label is the expression name. The Japanese Female Facial Expression Database (JAFFE) is used as train set.

Results And Analysis

Experiments were conducted on the JAFFE Database of Japan. This method ran on MATLAB R2013b, Windows 7. The comparisons of the method in this paper, classic LBP classification method, and the method base on irregular block of LBP operator^[13] were recorded in Table 1.

Table 1. Recognition accuracy with method of this paper of 3 facial expressions

Facial expression	Overall recognition accuracy
Happy, surprise, disgust	86.7%
Happy, surprise, fear	80.0%
Happy, fear, sad	80.0%
Surprise, fear, sad	66.7%

The classification test was conducted with improved LBP operator and k-means clustering method in this paper. The result showed that the recognition accuracy was up to 86.7% when there were 3 facial expressions picked out for test. Then results on 5 facial expressions were recorded in Table 2.

Table 2. Recognition accuracy using 3 methods of 5 facial expressions

Recognition method	Method in this paper	Classic LBP	irregular block LBP
Sad	62.1%	34.5%	40.0%
Happy	72.4%	34.5%	90.0%
Fear	43.5%	17.4%	40.0%
Surprise	52.0%	36.0%	76.7%
Disgust	74.1%	18.5%	43.3%

where the recognition accuracy of sad, happy, disgust was about 70%. Compared with other methods, recognition accuracy rised by more than 30.8%. But sad and surprise were easy to be mistook. Fear didn't get satisfied classification accuracy which needed further improvement.

Conclusion

Facial expression recognition has value for many fields including intelligent human-computer interaction, etc. This paper proposes an improved LBP operator to extract facial expression features. Facial features are acquired while the interference information is suppressed. K-means clustering is used for classification, and classification results optimize the classification criteria dynamically.

Acknowledgements

This paper was supported by the National Natural Science Foundation of China (61303104, 61373090, 11178017) and the Beijing Municipal Natural Science Foundation (4132014).

References

- [1]Pantic M, Rothkrantz L J M. Expert system for automatic analysis of facial expressions, J. Image and Vision Computing, 2000, 18 (11) 881-905.
- [2]Guoyun Zhang, Jing Zhang. Multi-scale PCA Support Vector Machine Algorithm Based on Gabor Wavelet for Face Recognition, J. Journal of Data Acquisition & Processing, 2009 (4) 437-442.
- [3]Lowe D G. Distinctive image features from scale-invariant keypoints, J. International journal of computer vision, 2004, 60 (2) 91-110.
- [4]Cristinacce D, Cootes T F, Scott I M. A multi-stage approach to facial feature detection, C //The British Machine Vision Conference (BMVC), Kingston, 2004. Kingston: Kingston University, 2004 1-10.
- [5]Kotsia I, Zafeiriou S, Pitas I. Texture and shape information fusion for facial expression and facial action unit recognition, J. Pattern Recognition, 2008, 41 (3) 833-851.
- [6] Jain A K, Dubes R C. Algorithms for clustering data, M. Englewood Cliffs: Prentice hall, 1988.
- [7]Feng G C, Yuen P C. Variance projection function and its application to eye detection for human face recognition, J. Pattern Recognition Letters, 1998, 19 (9) 899-906.
- [8]Nalwa V S, Binford T O. On detecting edges, J. Pattern Analysis and Machine Intelligence, IEEE Transactions on, 1986 (6) 699-714.
- [9]Xin Geng, Zhihua Zhou, Shifu Chen. Eye Location Based on Hybrid Projection Function, J. Journal Of Software, 2003, 14 (8) 1394-1400.
- [10]Fan Li, Xiaozhen Qin. An Eyes Location Method Based on Integral Projection, J. Computer&DigitalEngineering, 2011, 39 (9) 125-127.
- [11] Jinsong Zhong, Minhui Zhu. Ship Wake Detection Algorithm in SAR Image Based on Normalized Grey Level Hough Transform, J. Journal of Image and Graphics: Series A, 2004, 9 (2) 146-150.
- [12] Khatri N N, Shah Z H, Patel S A. Facial expression recognition: A survey, J. International Journal of Computer Science and Information Technologies (IJCSIT), 2014, 5 (1) 149-152.
- [13]Weifeng Liu, Shujuan Li, Yanjiang Wang. Facial expression analysis using LBP features, J. Computer engineering & application, 2011, 47 (2) 149-152.