Available online at http://www.cnki.net

The Journal of China Universities of Posts and Telecommunications

February 2018, 25(1): 48 – 53

http://jcupt.bupt.edu.cn

Facial expression recognition based on fusion of extended LDP and Gabor features

Luo Yuan [1], Yu Chaojing[1] (✉), Zhang Yi [2], Wang Boyu [2]

1. Key Laboratory of Optoelectronic Information Sensing and Transmission Technology, Chongqing University of Posts and Telecommunications, Chongqing 400065, China
2. Engineering Research Center for Information Accessibility and Service Robots, Chongqing University of Posts and Telecommunications, Chongqing 400065, China

Abstract

The local directional pattern (LDP) is unsusceptible to random noise which is widely used in texture extraction of face region. LDP cannot encode the central pixel thus the important information will be lost. Thus a new feature descriptor called extended local directional pattern (ELDP) is proposed for face extraction. First, the mean value of the eight directional edge response values and the gray value of center pixel are calculated. Second, the mean value is taken as the threshold. Then, the expression image is encoded using nine encoded values. In order to reduce redundant information and get more effective information, the Gabor filter is used to obtain the multi-direction Gabor magnitude maps (GMMs), and then the ELDP is used to encode the GMMs. Finally, support vector machine (SVM) is applied to classify and recognize facial expression. The experimental results show that the feature dimensions is greatly reduced and the rate of facial expression recognition is improved.

Keywords facial expression recognition, local directional pattern, ELDP, Gabor

1 Introduction

Facial expression recognition is a research hotspot in the field of human-computer interaction, computer vision, human psychology and emotion simulation [1]. Extracting an effective facial representation from human face images is a vital component of any successful facial expression recognition system. Recognition performance is influenced by the information contained in the expression representations heavily [2]. The methods of facial expression feature extraction are broadly divided into two types: global feature extraction and local texture feature extraction. Compared with global feature extraction, local feature extraction is more robust to the changes of illumination and posture. Local binary pattern (LBP) is a traditional method for local feature extraction, which is computationally efficient and robust to monotonic illumination changes, but the LBP operator produces very long histograms and thus in the case of a region descriptor, it is very difficult to use. To overcome these problems, center-symmetric LBP (CS-LBP) [3] has been introduced which produces very small and compact binary patterns. Then Taylor feature pattern (TFP) [4] based on the LBP and Taylor expansion has been presented to obtain an effective facial feature

Received date: 14-06-2017
Corresponding author: Yu Chaojing, E-mail: 1027096592@qq.com
DOI: 10.19682/j.cnki.1005-8885.2018.0005

from the Taylor feature map. A more robust facial descriptor, named as LDP [5], has been devised by Jabid et al., where the LDP representation of face demonstrated better recognition performance than LBP. In order to further improve the LDP operator and obtain better recognition performance, Local directional texture pattern (LDTP) [6] and dimensionality reduced local directional pattern (DR-LDP) [7] have been introduced. Local directional ternary pattern [8] has been devised which efficiently encodes information of emotion-related features by using the directional information and ternary pattern. However, the above improved methods ignored the importance of central pixels to image texture features, which affects the recognition rate.

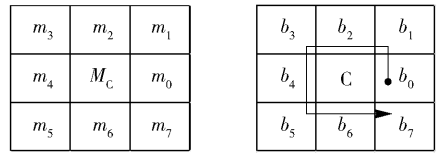
ELDP feature extraction method is proposed, which encodes the central pixel and increases the difference of various expressions to classify and recognize. Next, the fusion of ELDP and Gabor [9] is used to extract the facial expression features for enhancing the local features and extracting more detailed and effective texture information. Experimental results show that compared with the existing local texture feature extraction algorithm, the method proposed can greatly reduce the feature dimension and improve the recognition rate.

2 LDP and ELDP

2.1 LDP

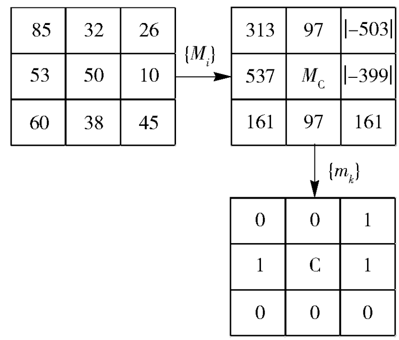
A LDP operator computes the edge response values in all eight directions at each pixel position and generates a code from the relative strength magnitude. Given a central pixel in the image, the eight directional edge response values $\{m_i\}$, $i = 0, 1, \ldots, 7$ are computed by Kirsch masks M_i [10] in eight different orientations centered on its position.

The response values are not equally important in all directions. The presence of a corner or edge causes high response values in some particular directions. Here, the top k directional bit responses, b_i, are set to 1. The remaining $(8-k)$ bit of the 8 bit LDP pattern are set to 0. The best recognition rate is achieved when $k=3$ [5]. Fig. 1 shows the mask response and LDP bit positions, and Fig. 2 shows an exemplary LDP code with $k=3$.



(a) Eight directional edge response positions (b) LDP binary bit positions

Fig. 1 LDP code



LDP binary code is 0001 0011 LDP decimal code is 19

Fig. 2 LDP code with $k=3$

2.2 ELDP

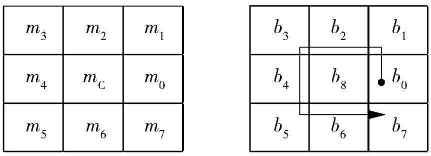
The central pixel provides more information than its neighborhood under some certain circumstances. Though LDP features have great discriminative power, LDP ignores the importance of central pixels to image texture features, which may lead to the loss of important information. In view of the above problems, the ELDP operator is proposed, which adds encoding to the central pixel.

In order to encode the central pixel, the mean value of the 3 × 3 region including the central pixel is obtained by using the Eq. (1) after calculating the eight directional edge response values. And then, the mean value is taken as the threshold, and the values of the top threshold are set to 1, and set the other $(9-k)$ values to 0. Finally, the ELDP code is derived using Eq. (2):

$$m = \frac{1}{9}\left(\sum_{i=0}^{7} m_i + m_C\right) \tag{1}$$

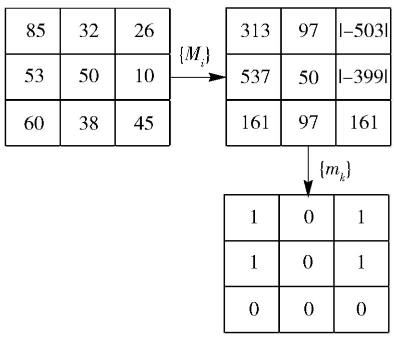
$$\left.\begin{aligned} E_{\mathrm{ELDP}} &= \sum_{i=0}^{7}\left(b_i(m_i - m)2^i + b_i(m_C - m)2^8\right) \\ b_i(a) &= \begin{cases} 1; & a \geqslant 0 \\ 0; & a < 0 \end{cases} \end{aligned}\right\} \tag{2}$$

where m_i is edge response values, m_C is the gray value of center pixel, and m is the mean value of the 3×3 region. Fig. 3 shows a ELDP code. The process of encoding one pixel of the image by using ELDP operator is shown in Fig. 4.



(a) Eight directional edge response positions (b) ELDP binary bit position

Fig. 3 ELDP code



LDP binary code is 0 0001 1011 LDP decimal code is 27

Fig. 4 Encoding process of ELDP

The ELDP code of each pixel (x, y) produces the corresponding ELDP coding pattern according to the original coordinate. After computing the ELDP code, the input image I of size $M\times N$ is represented by a ELDP histogram H_{ELDP} using Eq. (3). The resultant histogram H_{ELDP} is the ELDP descriptor of that image.

$$\left.\begin{aligned} H_{\text{ELDP}}(r) &= \sum_{x=1}^{M}\sum_{y=1}^{N} f(E_{\text{ELDP}}(x,y),r) \\ f(a,r) &= \begin{cases} 1; & a=r \\ 0; & a\neq r \end{cases} \end{aligned}\right\} \qquad (3)$$

where r is the ELDP code value. Fig. 5 shows the ELDP expression feature extraction process.

Fig. 5 The ELDP expression feature extraction process

Although ELDP effectively extracts the expression feature information, its feature extraction speed is slower and the feature dimension is higher than other local texture feature extraction algorithm. Therefore, in order to solve the above problems, the following research is performed in this paper.

3 Fusion of ELDP and Gabor for feature extraction

Gabor filter is a popular linear filter used by image processing community for edge detection. Gabor functions is discovered that it can model simple cells in the visual cortex of mammalian brains [11]. Hence, a Gabor wavelet filter with different frequency and direction can be obtained by setting the frequency and direction of the Gabor filter, which can realize multi-resolution and multi-direction analysis. A Gabor filter can be represented by the following equation:

$$g_{v,u}(x,y) = \frac{\| k_{v,u} \|}{\sigma^2} e^{\frac{\| k_{v,u} \|^2 \| z \|^2}{2\sigma^2}} \left[e^{(ik_{v,u}z)} - e^{-\frac{\sigma^2}{2}} \right] \qquad (4)$$

where v and u define the orientation and scale of the Gabor filters, $z=(x,y)$, σ is the standard deviation of the Gaussian function and determines the radius of the Gaussian envelope, and $\|\cdot\|$ denotes the norm operator, and the wave vector $k_{v,u} = k_v e^{i\varphi}$ where $k_v = k_{\max}/f^v$ and $\varphi_u = \pi u/n$. f is the spacing factor between filters in the frequency domain, $k_{\max}$ is the maximum spatial frequency of the filter bank and n is total number of filter banks. The Gabor representation of a face image is derived by convolving the face image with the Gabor filters [12]. Let $I(x,y)$ be the face image, its convolution with a Gabor filter $g_{v,u}(x, y)$, is defined as follows:

$$O_{v,u}(x,y) = I(x,y) * g_{v,u}(x,y) \qquad (5)$$

where $*$ denotes the convolution operator. $v=0, 1,\dots,4$ represents scales and $u=0, 1,\dots,7$ represents orientation of the Gabor filters. This results in 40 different filters with different scales and orientation. Convolving the image with each of the 40 Gabor filters can then generate the Gabor features. The phase information of the transform is time-varying, generally, only its magnitude is explored, thus one magnitude value will be computed, resulting in 40 response values at each pixel position for each Gabor filter. Therefore, the magnitude value of Gabor filter can be represented by the following equation:

$$G_{v,u}(x,y) = \sqrt{\mathrm{Re}(O_{v,u}(x,y))^2 + \mathrm{Im}(O_{v,u}(x,y))^2} \tag{6}$$

The features extracted by the Gabor filter contain a large amount of high dimensional data and redundant information. Based on this observation, we argue that combining Gabor filter with ELDP can enhance the local features, and thus more detailed information can quickly be extracted from ELDP operation. The ELDP operator encodes the texture features of GMMs in different directions and scales, denoted as Ψ_{LDPG}:

$$\Psi_{\mathrm{LDPG}} = \sum_{i=0}^{7} (b_i(G_i - G)2^i + b_i(G_c - G)2^8) \tag{7}$$

where G_i is ith directional features extracted by the Gabor filter after the ELDP code, G is features extracted by the Gabor filter after the ELDP code for the average pixel value, and G_c is features extracted by the Gabor filter after the ELDP code for center pixel.

In order to reduce the number of Gabor images, and to retain the original Gabor image multi-scale multi-directional local detail information. In this paper, the Gabor feature of five scales and eight directions is obtained by Gabor filter, and then the Gabor features of different scales in the same direction are superimposed, and the Gabor feature graphs in the eight directions are superimposed as superimposed feature graph (SFG), as shown in Fig. 6.

Therefore, in this paper, the SFG are first obtained from the Gabor filter, and then extract the ELDP feature of the above eight SFG features using the ELDP operator and the ELDP feature histogram of each SFG is obtained. Finally, concatenating all the feature histograms to get the fusion feature based on ELDP and Gabor.

4 Facial expressionimage descriptor

The fusion feature generated from the whole expression image loses some location information. But for facial expression images, some degree of locations and spatial relationship represents the expression image content better. Consequently, following a similar approach as Ref. [13], the histogram is modified to an extended histogram, where the expression image is divided into 8 × 8 sub-regions. The histogram sequences of each region are concatenated to describe the overall information of facial expressions, thus not only preserving the overall description of the facial expression image, but also highlights the changes in the facial detail and reduces the feature data dimension, and the extended histogram is built as follows

$$H = \sum_{x=1}^{7} \sum_{y=1}^{6} f(\psi_G(x,y), r) \tag{8}$$

where $f(a,r) = \begin{cases} 1; & a = r \\ 0; & \text{other} \end{cases}$, r is the ELDP code value. Finally, concatenating all the sub-regions distributions yields the feature descriptor shown in Fig. 7.

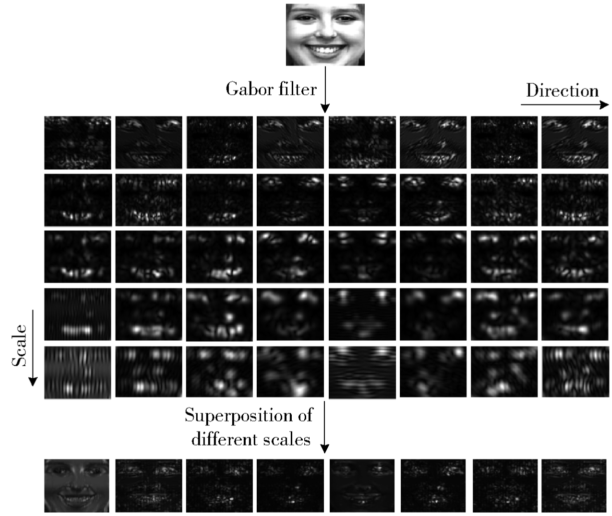


Fig. 6 SFG

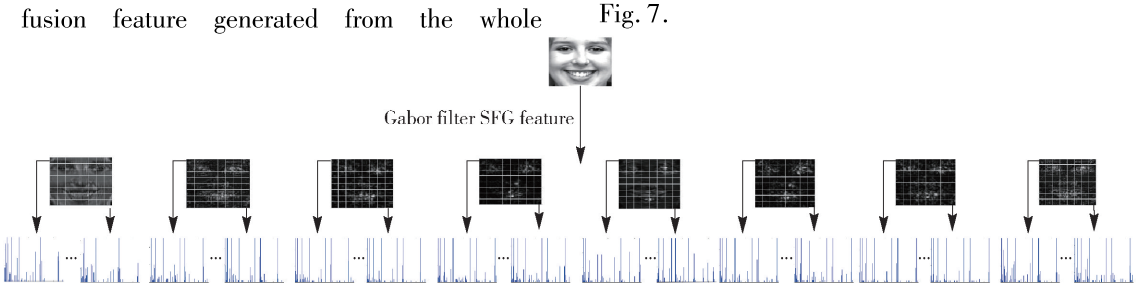


Fig. 7 Facial expression representations using the combined fusion feature histogram

5 Experimental results and analysis

SVM [14] is a well-established statistical learning theory that has been successfully applied in various classification tasks in computer vision. So we adopt SVM to classify facial expressions.

Most facial expression recognition systems attempt to recognize a set of prototypic emotional expressions like anger, disgust, fear, happy, sadness, and surprise [15]. The experiments are conducted on the Japanese female facial expression (JAFFE) [16] and the extended Cohn-Kanade (CK+) database [17]. JAFFE database contains 213 images of female facial expression expressed by 10 subjects. The CK+ database consists of 123 subjects and 593 image sequences. Among 593 video sequences, 327 sequences have emotion labels. All the 327 sequences are categorized into the following seven emotions: anger, happy, sad, contempt, fear, surprise, and disgust. The image sequences change from onset (the neutral frame) to peak (expressive frame). For focusing on the single image facial expression analysis, the peak frames should be picked out from each image sequence as the face image used. In order to verify the effectiveness of the proposed algorithm, all the samples were cropped from the original one using the positions of two eyes and resized into 64×64 pixels, and partitioned the images into 8×8 regions.

In the experiment, the facial expression features of JAFFE and CK+ database were extracted by weighted-projection-based LBP (WPLBP) [18], dynamic Gabor volume feature (DGVF), multiscale cell local intensity increasing patterns (MC-LIIP) [19], spatio-temporal manifold (STM) [20], LDP and the proposed method, respectively. Finally, SVM was used for classification. The proposed protocol in Ref. [13] is followed and the leave-one-out is taken cross validation strategy. The samples from one subject are used for testing and the samples from the remaining subjects are used for training for each time. The above process is repeated 118 times and the overall accuracy is computed as the performance. The experimental comparison result of facial expression recognition rate is given in Table 1, which is compared with WPLBP, DGVF, STM, MC-LIIP, LDP and the proposed method. Table 1 shows that the average recognition rate of the proposed method is 96.57% for CK+ database, which is 1.07% higher than that of WPLBP with the highest recognition rate in the above algorithms, and the average recognition rate of the proposed method is 91.43% for JAFFE database, which is the best recognition rate in the above algorithms. At the same time, the average time of the proposed method in this paper carried out a complete facial expression recognition for CK+ database is 60.27 ms, which is the shortest computing time compared with several other algorithms.

Table 1 Facial expression recognition rate and time of different algorithm

Method	Recognition rate		Time/ms
	JAFFE/%	CK+/%	
WPLBP	89.41	95.50	74.68
DGVF	87.20	93.60	92.35
STM	88.35	94.13	78.49
MC-LIIP	89.30	93.70	103.26
LDP	85.40	93.40	87.79
ELDP + Gabor	91.43	96.57	60.27

To get a better picture of the recognition accuracy of individual expression types, the confusion matrice (CM) for 6-class expression recognition using the CK+ database are given in Table 2. It can be observed that, anger, disgust, happy and surprise can be recognized with high accuracy, while the recognition rates of fear and sadness are lower than others due to their facial expression features are not obvious enough.

Table 2 CM of 6-class facial expression recognition using SVM

Expression/%	Anger	Disgust	Fear	Happy	Sadness	Surprise
Anger	96.4	–	–	–	3.6	–
Disgust	2.1	96.8	–	–	–	1.1
Fear	1.6	–	94.1	–	4.3	–
Happy	–	–	–	100	–	–
Sadness	1.8	–	5.7	–	92.5	–
Surprise	–	–	0.4	–	–	99.6

6 Conclusions and future work

A novel local image descriptor based on ELDP code is proposed in this paper, which is a available way to solve the problem that LDP ignores the importance of central pixels to image texture features. Meanwhile, in order to enhance the local features to improve the recognition rate, a new fusion method of ELDP and Gabor is presented. Experimental results prove that the proposed method not only greatly reduces the feature dimension, but also improves the recognition rate while maintaining better real-time performance compared with the current local feature extraction methods. The most obvious changes of facial expressions are eyebrows, eyes and mouth areas. Therefore, the enhancement of these three regions will be the further study in the subsequent works, and thus further improves the recognition rate of facial expression recognition.

Acknowledgements

This work was supported by the Science and Technology Commission Project of Chongqing (cstc2015jcyjBX0066).

References

1. Iglesias F, Negri P, Buemi M E, et al. Facial expression recognition: a comparison between static and dynamic approaches. Proceedings of the International Conference on Pattern Recognition Systems (ICPRS'16), Apr 20-22, 2016, Talca, Chile. Piscataway, NJ, USA: IEEE, 2016: 6p
2. Tarannum T, Paul A, Talukder K H. Human expression recognition based on facial features. Proceedings of the 5th International Conference on Informatics, Electronics and Vision (ICIEV'16), May 13-14, 2016, Dhaka, Bangladesh. Piscataway, NJ, USA: IEEE, 2016: 990-994
3. Neeru N, Kaur L. Face recognition based on LBP and CS-LBP technique under different emotions. Proceedings of the 2015 IEEE International Conference on Computational Intelligence and Computing Research (ICCIC'15), Dec 10-12, 2015, Madurai, India. Piscataway, NJ, USA: IEEE, 2016: 4p
4. Ding Y Y, Zhao Q, Li B Q, et al. Facial expression recognition from image sequence based on LBP and Taylor expansion. IEEE Access, 2017, 5: 19409-19419
5. Jabid T, Kabir M H, Chae O. Robust facial expression recognition based on local directional pattern. ETRI Journal, 2010, 32(5): 784-794
6. Rivera A R, Castillo J R, Chae O. Local directional texture pattern image descriptor. Pattern Recognition Letters, 2015, 5(1): 94-100
7. Srinivasa P R, Chandra Mouli P V S S R. Dimensionality reduced local directional pattern (DR-LDP) for face recognition. Expert Systems with Applications, 2016, 63: 66-73
8. Ryu B, Rivera A R, Kim J, et al. Local directional ternary pattern for facial expression recognition. IEEE Transactions on Image Processing, 2017, 26(12): 6006-6018
9. Ithaya Rani P, Muneeswaran K. Recognize the facial emotion in video sequences using eye and mouth temporal Gabor features. Multimedia Tools and Applications, 2017, 76(7): 10017-10040
10. Luo Y, Zhang T, Zhang Y. A novel fusion method of PCA and LDP for facial expression feature extraction. Optik: International Journal for Light and Electron Optics, 2016, 127(2): 718-721
11. Nouyed I, Poon B, Ashraful Amin M. A study on the discriminating characteristics of Gabor phase-face and an improved method for face recognition. International Journal of Machine Learning and Cybernetics, 2016, 7(6): 1115-1130
12. Al-Sumaidaee S A M, Dlay S S, Woo W L, et al. Facial expression recognition using local Gabor gradient code-horizontal diagonal descriptor. Proceedings of the 2nd IET International Conference on Intelligent Signal Processing, Dec 1-2, 2015, London, UK. Piscataway, NJ, USA: IEEE, 2015: 6p
13. Chen J K, Chi Z R, Fu H. Facial expression recognition with dynamic Gabor volume feature. Proceedings of the 18th IEEE International Workshop on Multimedia Signal Processing (MMSP'16), Sept 21-23, 2016, Montreal, Canada. Piscataway, NJ, USA: IEEE, 2016: 5p
14. Liu Y, Bi J W, Fan Z P. A method for multi-class sentiment classification based on an improved one-vs-one (OVO) strategy and the support vector machine (SVM) algorithm. Information Sciences, 2017, 394/395: 38-52
15. Luo Y, Wu C M, Zhang Y. Facial expression recognition based on fusion feature of PCA and LBP with SVM. Optik: International Journal for Light and Electron Optics, 2013, 124(17): 2767-2770
16. Kanade T, Cohn J F, Tian Y L. Comprehensive database for facial expression analysis. Proceedings of the 4th IEEE International Conference on Automatic Face and Gesture Recognition, Mar 28-30, 2000, Grenoble, France. Piscataway, NJ, USA: IEEE, 2000: 46-53
17. Kamarol S K A, Jaward M H, Parkkinen J, et al. Spatiotemporal feature extraction for facial expression recognition. IET Image Processing, 2016, 10(7): 534-541
18. Kumar S, Bhuyan M K, Chakraborty B K. Extraction of informative regions of a face for facial expression recognition. IET Computer Vision, 2016, 10(6): 567-576
19. Zhou A L, Wang H. Person-independent facial expression analysis by fusing multiscale cell features. Optical Engineering, 2013, 53(3): 254-260
20. Liu M Y, Shan S G, Wang R P, et al. Learning expressionlets on spatio-temporal manifold for dynamic facial expression recognition. Proceedings of the 2014 IEEE Conference on Computer Vision and Pattern Recognition (CVPR'14), Jun 23-28, 2014, Columbus, OH, USA. Piscataway, NJ, USA: IEEE, 2014: 1749-1756

(Editor: Wang Xuying)