Facial expression recognition based on two-step feature histogram optimization

Ling Gan[a], Sisi Si[b]

Chongqing Key Laboratory of Computational Intelligence,

Chongqing University of Posts and Telecommunications,

Chongqing, 40065, China

[a]ganling@cqupt.edu.cn , [b]ssscqupt@163.com

Keywords: Facial expression, feature histogram, wight, support vector machine.

Abstract. The feature histogram is made of different labels containing information about patterns on a pixel-level. This means that pixels of the same label may come from different parts of a face. There must be some errors between the actual feature histogram and the accurate feature histogram which can completely represent the information of a face image. In order to overcome this shortcoming, we propose a facial expression recognition approach based on two-step features histogram optimization. This method requires two steps. In the first step, features histogram based on local binary patterns, uniform local binary patterns and local gradient coding to be extracted. In the second step, a suitable weight to multiply with features histogram extracted in the first step. And these features are classified by the support vector machine. Experiments show that our approach can obtain a higher recognition rate and maintain the time efficiency.

1. Introduction

Facial expression is a significant part in human emotion, mental state and health state. Human facial expression plays an important role in people's communication. Automatic facial expression analysis and classification frame work consist of three modules: face detection, facial feature extraction, and facial expression classification respectively [1]. Feature extraction is the key part of facial expression recognition.Generally, geometric-feature-based methods and appearance-based methods [2] are two common approaches to extract facial features. The geometric-feature-based approaches require separate calculation to localize different facial components before extracting facial features. Comparing with geometric features, the appearance-based methods are easy to implement .The local binary patterns (LBP) [3] and Gabor wavelets transform [4] are the two representative appearance-based methods. The Gabor wavelets transform can detect the multi-scale and multi-direction changes of texture, but this method is very time-consuming and the characteristic dimension is huge.

The LBP algorithm can extract the expression features quickly, and it has strong texture discrimination ability. In 1996, the LBP operator was proposed by Ojala [5] et al and it has strong classification capability and better calculation efficiency. The LBP operator can generate different binary patterns. For example, there are 1048576 binary patterns for 20 sampling points within 5×5 neighborhoods and a great deal of them is not conducive to texture extraction and recognition. Therefore, Ojala [6] proposed 'uniform model' to reduce the dimension for the LBP operator in binary mode. However, these two approaches can not accurately describe the texture of facial muscles, wrinkles and other local deformations in facial expression information. In response to these problems, Ying Tong [7] et al proposed the Local Gradient Code (LGC) algorithm.

Since the critical step of facial expression recognition (FER) is to properly extract the discriminative facial features, this study aims to develop a facial expression recognition approach based on two-step feature histogram optimization. We adopt the local binary pattern (LBP), the uniform local binary pattern (ULBP) and the local gradient code (LGC) to transform a facial image into a feature image, and then we extract the most discriminative features based on block. After

extracting primary features, we will further optimize these features by multiplying a suitable weight. And then these features are classified by the support vector machine (SVM). The experimental results show that our approach can achieve higher recognition rate than traditional recognition methods and maintain the time efficiency.

2. The proposed approach

The feature extraction plays an important role in the whole facial expression recognition. In this study, a method of two-step features histogram optimization is proposed. In the first step, we extract the feature histogram based on the black-based LBP, LGC and ULBP. In the second step, the features vector extracted from the first step are multiplied by different weights. Fig.1 illustrates the proposed feature extraction and optimization framework.
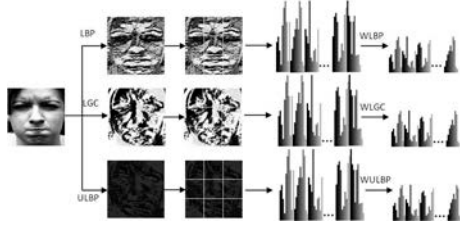

Fig. 1. The proposed two-step facial features extraction and optimization framework.

Fig.2. A 3×3 template of LBP, LGC, ULBP operator

2.1 The first step: get initial feature vector

The facial features based on LBP, LGC and ULBP are extracted. The concept of traditional LBP operator is to label the pixels of an image by a 3×3 neighborhood of each pixel with the center value and considering the results as a binary number. The concept of LGC is to compare the level of eight peripheral pixels, vertical and diagonal gradients, and convert the binary coding to decimal number, then get the LGC histogram statistics [12]. The concept of ULBP operator is that a binary sequence which contains at most twice by a 0-1 or 1-0 of the transformation. The algorithm is defined below in formula (1), (2), (3), using the 3×3 neighborhood template as shown in Fig. 2.

$$LBP_P^R = \sum^{p-1} s\left(g_p - g_c\right)2^p, \quad s\left(x\right) = \begin{cases} 1, x \geq 0, \\ 0, x < 0. \end{cases} \tag{1}\square\square\square$$

$$\begin{aligned} LGC_P^R = {} & s\left(g_1 - g_3\right)2^7 + s\left(g_4 - g_5\right)2^6 + s\left(g_6 - g_8\right)2^5 + s\left(g_1 - g_6\right)2^4 \\ & + s\left(g_2 - g_7\right)2^3 + s\left(g_3 - g_8\right)2^2 + s\left(g_1 - g_8\right)2^1 + s\left(g_3 - g_6\right)2^0. \end{aligned} \tag{2}$$

$$LBP_{P,R}^{riu2} = \begin{cases} \sum_{P=0}^{P-1} s(g_P - g_c) & if\ U(LBP_{P,R} \leq 2), \\ P+1 & otherwise, \end{cases} \tag{3}$$

where

$$U(LBP_{P,R}) = |\ s(g_{p-1} - g_c) - s(g_0 - g_c)\ |$$
$$+ \sum_{p=1}^{p-1} |\ s(g_p - g_c) - s(g_{p-1} - g_c)|.$$

Then we can get different feature vectors based on LBP, LGC and ULBP operator, like, $f_{ULBP}(x_1, x_2, x_3, ..., x_{d-1}, x_d)$ in this step.

2.2 The second step: feature vector optimization

In the second step, the features vector $f_{LBP}(x_1, x_2, x_3, ..., x_{d-1}, x_d)$, $f_{LGC}(x_1, x_2, x_3, ..., x_{d-1}, x_d)$, $f_{ULBP}(x_1, x_2, x_3, ..., x_{d-1}, x_d)$ will be optimized. Feature histogram is a statistic-based method and it represents the features of a face by counting the same pixel. But the same pixel may come from different parts of a face in the statistical results. For example, a pixel of 0 may come from the hair or the eyes. This will generate errors even if the image is divided into pieces. Therefore, we need to reduce this error by the second step. The algorithm formula is defined as below:

$$F_{WLBP}(x_1', x_2', x_3', , x_d' x_{d-1}') = f_{LBP}(x_1, x_2, x_3, ..., x_{d-1}, x_d) \times C \qquad (4)$$

$$F_{WLGC}(x_1', x_2', x_3', , x_d' x_{d-1}') = f_{LGC}(x_1, x_2, x_3, ..., x_{d-1}, x_d) \times C \qquad (5)$$

$$F_{WULBP}(x_1', x_2', x_3', , x_d' x_{d-1}') = f_{ULBP}(x_1, x_2, x_3, ..., x_{d-1}, x_d) \times C \qquad (6)$$

where $F_{WLBP}(x_1', x_2', x_3', , x_d' x_{d-1}')$, $F_{WLGC}(x_1', x_2', x_3', , x_d' x_{d-1}')$, $F_{WULBP}(x_1', x_2', x_3', , x_d' x_{d-1}')$ is the final facial feature vector we need, and C is the weight making the feature vector more exact.

2.3 Face feature representation

In order to capture important features of a face image, we develop a two-step approach to extract the discriminative facial features. Firstly, we transform an original image into the LBP transformed image by using formula (1), the LGC transformed image by using formula (2) and the ULBP transformed image by using formula (3). Secondly, each transformed image is partitioned into M small non-overlapping blocks $\{h_1, h_2, h_3, ..., h_{M-1}, h_M\}$, where each block may have different size commonly. We compute the histogram within each block. The histograms are extracted from each block and then concatenated into a single, spatially enhanced feature histogram which forms the facial feature vector that represents the input image. The feature vector contains facial information on three different levels of locality: the labels for the histogram contain information about the patterns on a pixel-level, the labels are summed over a small region to produce information on a regional level, and the regional histograms are concatenated to build a global description of the face [4]. Thirdly, the feature vectors are optimized by a suitable weight.

3. Experimental results and analysis

We use support vector machine to classify expressions in our experiments. The database used in our simulation experiments is the Cohn-Kanade database [9], one of the most comprehensive databases in the current facial-expression-research community. 540 image sequences are selected from the database which contains six kinds of emotions (Anger, Disgust, Fear, Happiness, Sadness, and Surprise) and each emotion include 90 images. 30 images are used as training images, and lest of the image sequence of each emotion are treated as testing images. The image size is 120×120. In our approach, the weight and the block number are two important factors that recognition quality. So we divide the image into 1×1, 2×2, 4×4, 8×8, 16×16 blocks.

3.1 Influence of weights and block number

In our approach, weight values of C and block number are important to recognition quality. In our experiment, we want to show a truth that no matter what algorithm we use, we can get a higher recognition rate as long as we use Statistic-based method. The maximum recognition rates of the LBP, LGC, and ULBP operator are shown in the Table 1. When C=0.4, 0.5 or 0.6 and images are divided into 4×4 blocks, the RLBP recognition rate will provide better recognition performance as shown in the Table 2. In order to prove a higher recognition performance can be provided with 3×3 blocks, we have done an extra experiment. When C=1.2 or 1.4, the WLBP recognition rate will provide better recognition performance with 3×3 blocks as shown in the Table 3. When C= 0.9 and the images are divided into 4×4 blocks, the WLGC recognition rate will provide better recognition performance as shown in the Table 4. When C= 0.6 and the images are divided into 3×3 blocks, the WULBP recognition rate will provide better recognition performance shown in the Table 5.

Table 1 The LBP, LGC and ULBP operators' recognition rate (%) with different block number size.

Method	block number				
	1×1	2×2	4×4	8×8	16×16
LBP	39.72	87.78	97.78	93.89	17.5
LGC	41.17	85.56	96.94	88.06	16.67
ULBP	43.06	90.56	97.78	93.61	17.22

Table 3 The WLBP's recognition rates (%) with the weights between 0.92 and 1.8 and 3×3 blocks.

	Weights							
	0.92	0.94	0.96	0.98	1.2	1.4	1.6	1.8
Recognition rate (%)	98.06	98.06	98.06	98.06	98.33	98.33	97.78	97.5

Table 2 The WLBP's recognition rates (%) with different weights and block number.

Weights	block number				
C	1×1	2×2	4×4	8×8	16×16
0.1	24.44	30.56	86.39	97.5	96.39
0.2	24.72	59.44	95.28	97.5	93.61
0.3	26.67	66.67	97.22	96.94	91.67
0.4	28.33	71.94	98.06	95.56	88.89
0.5	30.28	76.39	*98.33*	95	88.06
0.6	33.33	79.17	*98.33*	94.72	68.89
0.7	36.39	80.56	*98.33*	95	43.89
0.8	38.61	83.06	98.06	94.44	28.33
0.9	37.5	85.83	97.5	94.17	20.28
2	52.22	95.56	96.67	85.83	16.67
3	56.67	95.28	93.89	33.33	16.67
4	63.06	95.56	96.39	16.94	16.67
5	69.17	95	96.94	16.67	16.67
6	71.39	95	96.94	16.67	16.67
7	73.61	94.44	94.17	16.67	16.67
8	75	93.61	88.06	16.67	16.67
9	75.56	92.78	77.22	16.67	16.67
10	77.5	92.5	64.72	16.67	16.67

Table 4 The WLGC's recognition rates (%) with different weights and block.

Weights	black number				
C	1×1	2×2	4×4	8×8	16×16
0.1	36.39	55	83.06	97.5	96.11
0.2	36.39	59.17	92.5	96.94	91.67
0.3	36.39	64.44	95.28	94.72	87.78
0.4	36.67	70.83	96.39	92.5	87.22
0.5	36.39	73.61	96.39	90.83	84.72
0.6	38.06	74.72	96.61	89.44	62.22
0.7	39.72	78.33	96.39	89.44	36.94
0.8	40.56	80.83	96.94	88.61	22.5
0.9	42.78	83.33	97.22	88.06	18.33
2	47.78	93.39	96.94	76.67	16.67
3	50.83	95	96.11	32.22	16.67
4	55.56	95.56	96.11	16.94	16.67
5	59.17	96.94	95	16.67	16.67
6	66.94	96.39	92.5	16.67	16.67
7	70.83	96.67	93.33	16.67	16.67
8	71.94	96.11	86.94	16.67	16.67
9	73.61	95.56	78.33	16.67	16.67
10	75	95.28	66.39	16.67	16.67

Table 5 The WULBP's recognition rates (%) with different weights and block.

Weights C	black number				
	1×1	2×2	4×4	8×8	16×16
0.1	25	32.22	86.67	98.33	96.39
0.2	25.28	56.39	95	97.5	93.33
0.3	28.61	66.11	97.22	97.22	90.83
0.4	30	72.78	98.33	96.11	89.44
0.5	31.39	76.94	98.33	94.72	86.39
0.6	36.39	81.11	98.33	93.61	67.5
0.7	40.28	84.44	98.06	93.61	43.33
0.8	40.06	86.39	97.78	93.06	25.56
0.9	38.89	88.06	98.06	93.89	19.44
2	52.22	95.83	96.39	85	16.67
3	56.94	96.67	95.83	29.44	16.67
4	63.89	96.67	95.56	16.67	16.67
5	70.83	96.67	95.28	16.67	16.67
6	73.89	95.83	95.28	16.67	16.67
7	75.56	95.83	93.06	16.67	16.67
8	77.22	95	88.06	16.67	16.67
9	78.61	94.17	75.56	16.67	16.67
10	76.94	94.17	63.61	16.67	16.67

Table 6 The operators' recognition rate and run-time with 3×3 blocks.

	Methods							
	LBP	WLBP	LGC	WLGC	ULBP	WULBP	LGC-HD	WLGC-HD
Recognition rate (%)	98.06	98.33	95.83	96.11	98.33	98.89	83.06	87.78
Classification time (s)	12.028	11.011	11.243	11.208	8.976	8.377	8.522	8.505

3.2 The effect of our method on execution time and the comparison with other methods

We do not want to see that the recognition rate is improved but the time efficiency is declined. So an experiment using 3×3 blocks to research the execution time is done and the results are shown in the Table 6. Overall, the execution time is not increased in the case of improving recognition rate. We can also see the WULBP recognition rate is the highest and its execution time is the shortest comparing with other methods from the Table 6. In addition, the comparative results of each facial expression by using the aforementioned methods are shown in Fig. 3. These comparisons reveal that the recognition rate of the original algorithm can be improved by two-step exact features histogram by weights.
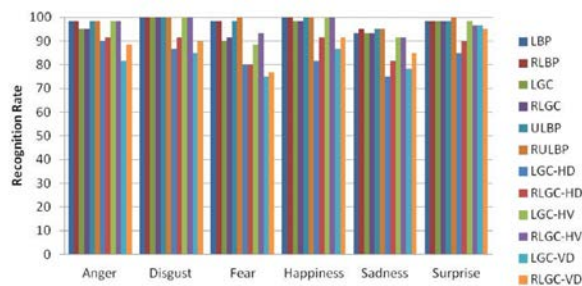

Fig.3. Recognition performance comparison of six basic facial expression.

4 Summary

Facial expression recognition is one of the most challenging subjects at present which pursues faster and more accurate recognition performance. In this paper, we propose a method of two-step exact features histogram by weights. In the first step, we extract the feature histogram, which contains the basic texture information of the face image, but there are some errors. In the second step, we use the appropriate weight to reduce the error of the feature histogram. Even more amazing is that the time efficiency remain stable. Experimental results show that the proposed method can improve the recognition rate and the time efficiency remains stable. Further work, we will find a way to automatically obtain the optimal configuration so that the recognition rate can be further improved.

References

[1] B. Fasel, J. Luettin, Automatic facial expression analysis: a survey, Pattern Recogn. 36 (1) (2003) 259–275.

[2] C. Shan, S. Gong, P.W. McOwan, Facial expression recognition based on local binary patterns: a comprehensive study, Image and Vision Computing 27 (6)(2009) 803–816.

[3] T. Ahonen, A. Hadid, M. Pietikä inen, Face description with local binary patterns: application to face recognition, IEEE Trans. Pattern Anal. Mach. Intell. 28 (12) (2006) 2037–2041. A. Rosenfeld and A. Kak. Digital Image Processing (2nd Edition, Vol. 2 ed.),Academic Press, Orlando (1982)

[4] M.S. Barlett, G. Littlewort, I. Fasel, J.R. Movellan, Real time face detection and facial expression recognition: development and applic ations to human computer interaction, in: 2003 Conference on Computer Vision and Pattern Recognition Workshop, 2003, p. 53.

[5] T. Ojala, Pietikainen M, D. Harwood. A comparative study of texture measures with classification based on feature distributions[J]. Pattern Recognition,1996, 29:51-59.

[6] W.-c. Zhang, S.-g. Shan, H.-m. Zhang, et al., Histogram sequence of local Gabor binary pattern for face description and identification, J. Softw. 17 (12) (2006) 2508–2517.

[7] Y. Tong, R. Chen, Y. Cheng. Facial expression recognition algorithm using LGC based on horizontal and diagonal prior principle[J]. Optik - International Journal for Light and Electron Optics,2014.

[8] V. Vapnik, The Nature of Statistical Learning Theory, Springer, 1995.

[9] T. Kanade, J. Cohn, Y. Tian, Comprehensive database for facial expression analysis, in: Fourth IEEE International Conference on Automatic Face and Gesture Recognition, 2000, pp. 46–53.