RESEARCH ARTICLE OPEN ACCESS

Facial expression recognition based on completed LBP

Zicheng Lin[1], Yuanliang Huang[2]

[1](College of Science and Engineering, Jinan University, Guangzhou, PR China)

[2](Institute of Electrical Automation, Jinan University, Zhuhai, PR China)

--************************----------------------------------

Abstract:

Because of the limitation for the traditional local binary pattern (LBP) operator in facial expression recognition, an effective facial expression recognition algorithm based on completed local binary pattern (CLBP) is proposed. Firstly, the method performs face segmentation and illumination normalization on original image. Secondly, CLBP operator is used to extract facial expression features from the preprocessed image. CLBP histograms could be combined by CLBP-Center (CLBP_C), CLBP-Magnitude (CLBP_M) and CLBP-Sign (CLBP_S) features in two ways, jointly or hybridly. Lastly, the nearest neighbour classifier is applied for classifying and recognizing the CLBP histograms. The experiments on JAFFE and Oulu-CASIA facial expression databases demonstrate that the proposed algorithm can achieve better performance than using the traditional LBP facial expression method. In addition, our algorithm can achieve good performance in complex illumination conditions.

Keywords **—Completed local binary pattern, facial expression recognition, illumination normalization.**

--************************----------------------------------

I. INTRODUCTION

Facial expression is a significant component of interpersonal communication which provides rich and powerful emotional information. The majority of the existing expression recognition techniques focus on classifying 7 basic expressions, namely: anger, happiness, neutral, fear, surprise, disgust, sad. Facial expression recognition is of great significance in human-computer interaction, thus attracting more and more scientists' attention.

Facial expression recognition system includes three parts: image preprocessing, feature extraction and classification. Among them, extracting expression features effectively is the most important step in the system. Commonly used facial feature extraction techniques include LBP feature-based methods [1], Gabor wavelet-based methods [2] and geometric feature-based methods [3]. Gabor wavelets can extract multi-scale and multi-orientation features effectively, but it is inefficient in both time and memory for the redundancy of Gabor wavelet features. Most geometric-based approaches use the Active Shape Model (ASM) and Active Appearance Model (AAM) to extract reliable features. However, ASM and AAM are sensitive to initial shape and may easily be stuck in local minima. LBP-based methods are computationally efficient and robust to monotonic illumination variation, thus LBP operator is widely applied in facial expression recognition. Although LBP operator provides a theoretically simple and efficient approach to facial expression analyses, it merely uses the sign of the difference between center pixel and neighbor pixel to represent the local pattern, instead of the magnitude. In the view of this shortcoming, Guo *et al*. [4,5] proposed completed local binary pattern (CLBP) operator, which was successfully applied in texture classification.

In this paper, a facial expression recognition method based on CLBP is proposed. Firstly, face segmentation and illumination normalization are applied to the face images, and then we divide the original image into 4×4 block. Secondly, CLBP operator is utilized to extract facial features from each block. In CLBP, a local region is represented by local different sign-magnitude transform (LDSMT) and its center pixel. The LDSMT decomposes the image local structure into two complementary components: the difference magnitudes and the difference signs. Then two operators, CLBP-Magnitude (CLBP_M) and CLBP_Sign (CLBP_S) are proposed to code them. The center pixel is simply coded by binary code after global thresholding, and the binary map is named as CLBP_Center (CLBP_C). The CLBP histograms extracted from each block could be combined by CLBP_C, CLBP_M and CLBP_S features into joint or hybrid distributions, and then the histograms of the blocks are combined into the CLBP histogram of the whole image. Lastly, the nearest neighbour classifier is used to identify the CLBP features. The experiments with the JAFFE and Oulu-CASIA facial expression databases demonstrate that the CLBP operator can extract more comprehensive facial information than LBP operator and has strong illumination robustness. It can achieve a good recognition under complex illumination conditions like dark, weak and strong illuminations.

II. BASIC THEORY

In order to improve the accuracy of facial expression recognition, face image preprocessing is performed before extracting facial expression features.

A. *Image Preprocessing*

In this paper, we use the Viola-Jones face detection algorithm to determine the position of the face in the image and crop the face gray images from the original images to remove the background information which is not useful for recognition. After that, we use bilinear interpolation to normalize the cropped images to 128×128 pixels. In the practical application of facial expression recognition, different illumination conditions have great influence on the accuracy. Therefore, illumination normalization is performed to reduce the influence of illumination before extracting features. In this paper, Gamma Correction, Difference of Gaussian filtering (DoG) and histogram equalization [6] are used to normalize the face image.

1) Gamma Correction

Gamma Correction is a nonlinear gray-level transformation which replaces gray-level I with $\log(I)$ (for $\gamma = 0$). Here $\gamma \in [0,1]$ is a user-defined parameter. This optimizes the local dynamic range of the image in shadowed or dark regions while compressing it at highlights and in bright regions. The images are transformed by the following formula:

$$I \to \log(I), \gamma \in [0,1] \quad (1)$$

2) Difference of Gaussian Filtering (DoG)

Gamma correction does not eliminate the influence of overall intensity gradients such as shading effects. Shading induced by surface structure is a potentially useful visual cue but it is predominantly low spatial frequency information that is hard to separate from effects caused by illumination gradients. High-pass filtering removes both the incidental and the useful information, thus simplifying the recognition problem and increasing the overall system performance in many cases. DoG is used to correct the non-uniform illumination in the image. Besides, it is a bandpass filter to filter high-frequency information and retains low-frequency information to eliminate noise. Its transfer function is the difference between two different parameters Gaussian function [7]:

$$G(s) = A_1 e^{-s^2/2\alpha_1^2} - A_2 e^{-s^2/2\alpha_2^2}, A_1 \geq A_2 > 0, \alpha_1 > \alpha_2 \quad (2)$$

3) Histogram equalization

Histogram equalization increases the global contrast of images and makes the intensities better distribute on the histogram. Let N be the number of pixels in the image, n_i is the number of gray-value of i, L is the number of gray levels, then the probability of occurrence of gray scale i is

$$p(x=i) = \frac{n_i}{n}, 0 \leq i < L \quad (3)$$

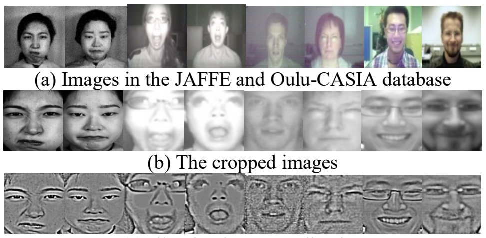
Histogram equalization processing formula is

$$H(i) = \sum_{j=0}^{i} P(x=i) = \sum_{j=0}^{i} \frac{n_i}{N} \quad (4)$$

Then the pixel i after histogram transformation, the gray level is

$$I = L \times H(i) \quad (5)$$

Some images before preprocessing and after preprocessing in the JAFFE and Oulu-CASIA facial expression database are shown in Fig. 1.



(a) Images in the JAFFE and Oulu-CASIA database

(b) The cropped images

(c) Illumination normalization

Fig.1 JAFFE and Oulu-CASIA database

B. LBP

In order to describe the texture features of images effectively, Ojala [8] proposed local binary pattern (LBP) operator and LBP was successfully applied in texture recognition. LBP is a gray-scale and rotation invariant texture primitive which can achieve impressive classification results in texture recognition. LBP is a simple and efficient operator to describe local image pattern, and it has been adapted to many applications like dynamic texture recognition, fingerprint recognition and face recognition [9,10]. The basic LBP is originally defined for 3×3 neighbourhoods, giving 8 bit codes based on the pixels of the neighbourhoods around the central pixel. The operator labels the pixels of an image by thresholding a 3×3 neighbourhood of each pixel with center value and considering the result as a binary number, and the 256-bin histogram of the LBP labels computed over a region is used as a texture descriptor. The basic LBP operator encoding process is illustrated in Fig. 2.

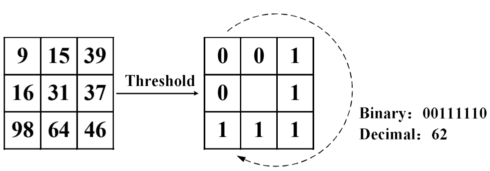


Fig. 2 The basic LBP operator

In order to extract different scale texture features, Ojala [11] extended the basic LBP operator to use different sizes of neighborhood to extract dominant features.

The extended LBP operator can be described as:

$$LBP_{P,R} = \sum_{p=0}^{P-1} s(g_p - g_c)2^p \qquad (6)$$

where $s(g_p\text{-}g_c)=\begin{cases}1, & g_p\text{-}g_c \geq 0\\ 0, & g_p\text{-}g_c<0\end{cases}$, g_c is the gray value of the center pixel, g_p the gray value of the neighbourhood, R the radius of the neighbors and P the number of neighbours. The LBP operator with different P and R is shown in Fig. 3.

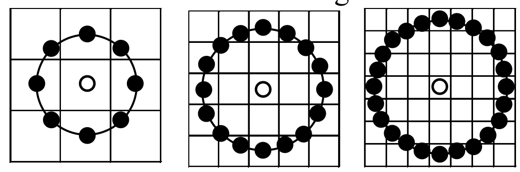


Fig. 3 Examples of the extended LBP operator: the circular(8,1), (16,2) and (24,3) neighborhoods.

C. CLBP

Because the traditional LBP operator is only represented by the information of differences between the center pixels and the neighbour gray values and the magnitude information is ignored, some information may be lost when using LBP operator to extract texture feature. As shown in Fig. 4, the two different local patterns also can get the same results when using LBP to extract features which show that tradition LBP operator can not extract comprehensive texture features.

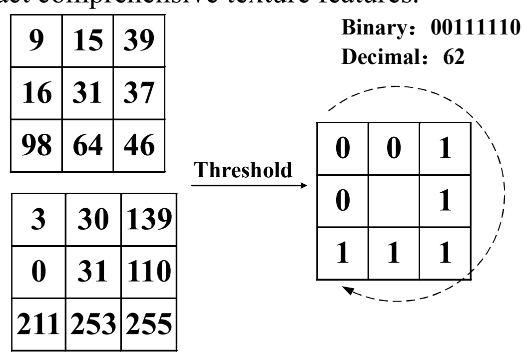


Fig. 4 Two examples of computing LBP

Thus, Guo *et al* proposed completed local binary pattern for describing the local spatial structure of the image texture. In CLBP operator, the center pixel and local difference sign-magnitude transform (LDSMT) can present the local spatial structure of the image texture. The center pixel is coded by binary code after global thresholding, and the binary map is named as CLBP-Center (CLBP_C). The LDSMT decomposes the image local structure into two complementary components: CLBP-Sign (CLBP_S) and CLBP-Magnitude (CLBP_M). Referring to Fig. 3, given a central pixel g_c and its P circularly and evenly spaced neighbours g_p with radius R, $p = 0,1,\ldots,P\text{-}1$, we can simply calculate the difference between g_c and g_p as $d_p=g_p - g_c$. d_p can be further decomposed into two components:

$$d_p=s_p\times m_p \text{ and } s_p=\begin{cases} s_p=\text{sign}(d_p) \\ m_p=|d_p| \end{cases} \qquad (7)$$

where $s_p=\begin{cases} 1\ , & d_p\geq 0 \\ -1\ , & d_p<0 \end{cases}$ is the sign and m_p is the magnitude of d_p.

9	15	39
16	31	37
98	64	46

(a)A 3× 3 sample block

-22	-16	8
-15		6
67	33	15

(b)the local differences

-1	-1	1
-1		1
1	1	1

(c)the sign components

22	16	8
15		6
67	33	15

(d)the magnitude components

Fig. 5 Completed LBP operators

Thus, three operators are defined as:

$$CLBP\_S_{P,R}=\sum_{P=0}^{P-1} t(s_p,0)2^p \qquad (8)$$

$$CLBP\_M_{P,R}=\sum_{P=0}^{P\text{-}1} t(m_p,c_m)2^p \qquad (9)$$

$$CLBP\_C_{P,R}=t(g_c,c_I) \qquad (10)$$

where t is a threshold function $t(x,c)=\begin{cases} 1, & x\geq c \\ 0, & x<c \end{cases}$. c_m is the mean value of m_p for the image and c_I is the average gray level of the image.

D. Facial expression recognition based on CLBP

Aiming at facial expression recognition under complex illumination conditions, a facial expression recognition algorithm based on completed LBP is proposed in this paper. CLBP can comprehensively describe texture features of the expression images and it can achieve strong robustness to illumination change. Before feature extraction, face segmentation and illumination normalization were used to remove the background information and the influence of the illumination changes. In order to effectively describe the facial expression features, the face image is divided into 4×4 sub-blocks. The CLBP operator is used to extract features from each sub-block, and then we can get the CLBP_S, CLBP_M and CLBP_C histograms of the sub-blocks. These histograms could be combined in two ways, jointly or in concatenation. In the first way, for example, we separately calculate the histograms of the CLBP_S and CLBP_M, and then concatenate the two histograms together. We can get the CLBP histogram named CLBP_S_M in this way. In the second way, we calculate a joint 2D histogram of the CLBP_S and CLBP_M codes. This CLBP scheme is represented as "CLBP_S/M". The advantage of this approach is to keep expression features richer. Besides, the three operators, CLBP_S, CLBP_M and CLBP_C, could be combined in the above two ways. A 2D joint histogram, The CLBP_M/C or CLBP_S/C is built first, and then the histogram is converted to a 1D histogram, which is then concatenated with CLBP_S or CLBP_M to generate a joint histogram, denoted by "CLBP_S_M/C" or "CLBP_M_S/C". As shown in Fig. 6, (a) is the histogram of CLBP_S of a sample and (b) is the histogram of CLBP_M of a sample, (c) is the histogram of CLBP_S_M and (d) is the histogram of CLBP_S/M.

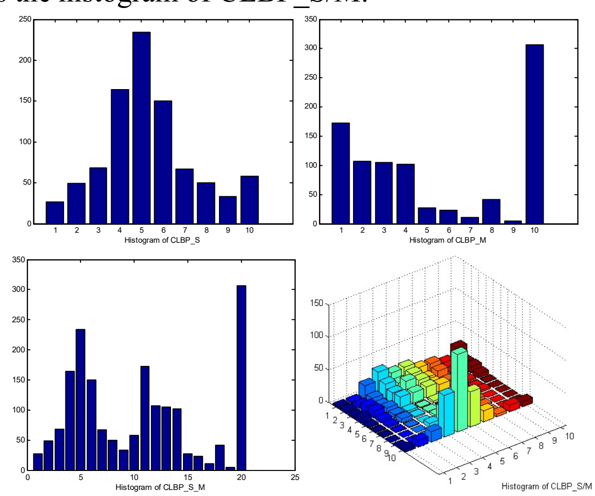


Fig. 6 Two fusing methods

The CLBP histogram of each sub-block is fused into the CLBP histogram of the whole image. Lastly, the CLBP features are classified by the nearest neighbour classifier.

The concrete steps of our algorithm are as follows:

Step1: Facial expression image preprocessing according to Section 2.1

Step2: 4×4 sub-blocks processing after preprocessing

Step3: The CLBP_S, CLBP_M and CLBP_C histograms are obtained by using the multi-scale CLBP operators to extract expression features from each sub-block.
Step4: CLBP_M/C, CLBP_S_M/C and CLBP_S/M histograms of the sub-blocks are built by The CLBP_S, CLBP_M and CLBP_C histograms.
Step5: Fusing the whole image's CLBP_M/C, CLBP_S_M/C and CLBP_S/M histograms.
Step6: Classified by the nearest neighbour classifier.
The framework of our algorithm is illustrated in Fig .7.

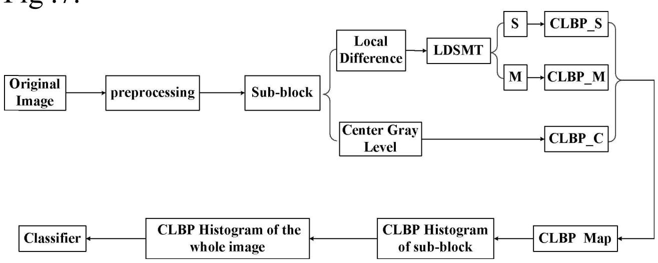


Fig. 7 The framework of our algorithm

III. EXPERIMENTAL RESULTS

To evaluate the effectiveness of the proposed algorithm, a series of experiments are conducted on JAFFE database [12] and Oulu-CASIA VIS database [13]. The JAFFE database contains 213 images of seven facial expressions (six basic facial expressions and one neutral expression) from 10 Japanese female participants. The Oulu-CASIA VIS database consists of six basic facial expressions (anger, disgust, fear, happiness, sadness, surprise) from 80 persons (50 Europeans and 30 Chinese) between 23 to 58 years old. The Oulu-CASIA database was taken under three difficult lighting conditions: normal, weak and dark.

In the experiment of JAFFE database, 144 images are the training set and the remaining 69 images are the testing set, the experiment result is shown in Table. 1

Table 1 Classification rate on JAFFE database

Histograms	(P,R)=(8,1)	(P,R)=(16,2)	(P,R)=(24,3)
LBP^{u2}	76.81%	79.71%	82.60%
CLBP_S	78.26%	85.51%	85.51%
CLBP_M	76.81%	79.71%	81.16%
CLBP_M/C	85.51%	85.51%	86.89%
CLBP_S_M/C	86.96%	88.41%	89.86%
CLBP_S/M	78.26 %	85.51%	88.41%

In the experiments of Oulu-CASIA database, we evaluate our algorithm in dark, weak and strong illumination conditions. In the experiment of each illumination condition, 240 samples are chosen as the training set and the remaining 240 are testing samples.

Table 2 Classification rate on Oulu-CASIA database under dark illumination

Histograms	(P,R)=(8,1)	(P,R)=(16,2)	(P,R)=(24,3)
LBP^{u2}	43.75%	45.83%	47.50%
CLBP_S	87.92%	89.17%	90.42%
CLBP_M	86.67%	87.50%	88.75%
CLBP_M/C	91.25%	92.08%	93.33%
CLBP_S_M/C	92.92%	93.75%	95.42%
CLBP_S/M	96.25%	97.50%	98.75%

Table 3 Classification rate on Oulu-CASIA database under weak illumination

Histograms	(P,R)=(8,1)	(P,R)=(16,2)	(P,R)=(24,3)
LBP^{u2}	44.58%	46.67%	50.83%
CLBP_S	93.33%	95.83%	96.67%
CLBP_M	92.08%	94.58%	96.25%
CLBP_M/C	95.42%	96.67%	97.50%
CLBP_S_M/C	96.67%	97.92%	98.33%
CLBP_S/M	97.92%	98.75%	97.91%

Table 4 Classification rate on Oulu-CASIA database under strong illumination

Histograms	(P,R)=(8,1)	(P,R)=(16,2)	(P,R)=(24,3)
LBP^{u2}	50.42%	55.00%	58.33%
CLBP_S	97.92%	99.17%	100%
CLBP_M	97.50%	98.33%	99.58%
CLBP_M/C	99.17%	99.58%	100%
CLBP_S_M/C	99.58%	100%	100%
CLBP_S/M	100%	100%	100%

It can be seen from Table 1 to Table 4, the recognition rate of CLBP method is higher than traditional LBP method, which indicates that CLBP features contain more discriminant information than LBP features, and CLBP features can describe expression information comprehensively. Besides, CLBP_S can achieve better results than CLBP_M because the sign component is more informative than the magnitude component. The CLBP_M/C could get higher recognition rates than CLBP_M because CLBP_M/C contains center pixel information which represents the gray level of the local patch. CLBP_S and CLBP_M/C contain complementary features, while fusing them could improve classification accuracy. It can be seen from Table. 2 to Table.4 that the traditional LBP operator is less effectiveness in complex illumination conditions while CLBP always has better results, which shows that CLBP has stronger robustness under complex illumination conditions. The superiority of the CLBP is owing to the utilization information of the magnitude and center pixel, which preserves some significant texture features discarded by the original LBP operator.

IV. CONCLUSIONS

In this paper, a facial expression recognition method based on completed LBP is proposed. Through using normalization illumination and face segmentation, we could extract pure expression features effectively. Besides, completed LBP operator could extract more comprehensive reliable features than the traditional LBP operator. The experiments on JAFFE and Oulu-CASIA facial expression database demonstrate that the proposed algorithm can achieve better performance than using the traditional LBP facial expression method. In addition, our algorithm can achieve good performance in complex illumination conditions.

ACKNOWLEDGMENT

The authors sincerely thank Psychology Department of Kyushu University and CMVS of the University of Oulu for sharing JAFFE and Oulu-CASIA facial expression database.

REFERENCES

[1] C. F. Shan, S. G. Gong, P. W. McOwan. Facial expression recognition based on Local Binary Patterns: A comprehensive study. Image And Vision Computing, 2009, 27:803-816.

[2] W. F. Gu, C. Xiang, Y. V. Venkatesh, D. Huang, H. Lin. Facial expression recognition using radial encoding of local Gabor features and classifier synthesis. Pattern Recognition, 2012, 45:80-91.

[3] X. Pu, K. Fan, X. Chen, L. Ji. Facial expression recognition from image sequences using twofold random forest classifier. Neurocomputing, 2015, 168: 1173-1180.

[4] Z. H. Guo, L. Zhang, D. Zhang. A Completed Modeling of Local Binary Pattern Operator for Texture Classification. IEEE Transactions on Image Processing, 2010, 19: 1657-1663.

[5] Z. H. Guo, X. Z. Wang, J. Zhou, J. You. Robust Texture Image Representation by Scale Selective Local Binary Patterns. IEEE Transactions on Image Processing, 2016, 25: 687-699.

[6] X. Y. Tan, B. Triggs. Enhanced Local Texture Feature Sets for Face Recognition Under Difficult Lighting Conditions. IEEE Transactions on Image Processing, 2010, 19: 1635-1650.

[7] Wang Zhaozhong, Zhou Fugen. Nonuniform illumination correction based on difference of Gaussian filter. Infrared and Laser Engineering，2010, 29:64-64

[8] *I. H. T Ojala. A comparative study of texture measures with classification based on featured distributions. Pattern Recognition, 1996, 29(1):51-59.*

[9] Wang Xian, Zhang Yan, Mu Xin, Zhang Fangsheng. The Face Recognition Algorithm Based on Improved LBP. Opto-Electronic Engineering， 2012, 39:109-114.

[10] Yuan Baohua, Wang Huan, Ren Mingwu. Face recognition based on completed local binary pattern. Application Research of Computers，2012,29: 1557-1559.

[11] T. O. M. P. T. Mäenpää. Multiresolution Gray-Scale and Rotation Invariant Texture Classification with Local Binary Patterns. IEEE Transactions on Pattern Analysis And Machine Intelligence, 2002, 24(7): 971-987.

[12] M. Lyons, S. Akamatsu, M. Kamachi, J. Gyoba. Coding facial expressions with Gabor wavelets. Automatic Face And Gesture Recognition - Third IEEE International Conference Proceedings, 1998:200-205.

[13] G. Y. Zhao, X. H. Huang, M. Taini, S. Z. Li, M. Pietikainen. Facial expression recognition from near-infrared videos. Image And Vision Computing, 2011, 29:607-619.