Facial Expression Recognition Based on Local Directional Pattern Using SVM Decision-level Fusion

Juxiang Zhou[1], Tianwei Xu[2], Jianhou Gan[1]

1. Key Laboratory of Education Informalization for Nationalities, Ministry of Education,Yunnan NormalUniversity,Kunming,China2. College of Information,Yunnan Normal University, Kunming, China
{zjuxiang@126.com,xutianwei@ynnu.edu.cn,kmganjh@yahoo.com.cn}

Abstract. This paper presents a novel expression recognition method based on global and local features with decision-level fusion. We first extract Local Directional Pattern (LDP) global features of the whole face which can guarantee basic expression difference and decrease the influence of no-facial region meanwhile, then the Local Directional Pattern Variance (LDPv) descriptor is used to extract local features of regions of eyes and mouth to extrude their contribution on expression changes. After feature extraction, we don't use feature fusion with simple concatenation, a decision-level fusion for global LDP feature and local LDPv feature by Support Vector Machine (SVM) is selected to recognition respectively. Furthermore, we also research the optimal parameters for regions-dividing and weight of LDPv. Extensive results from two standard databases indicate the effectiveness of our proposed method.

Keywords: Facial expression recognition, Local directional pattern, SVM

1 Introduction

Automatic facial expression analysis is an interesting and challenging problem, and impacts important applications in many areas such as human-computer interaction, data-driven animation, human psychology theory and emotional simulation study [1]. For successful facial expression recognition, deriving an effective facial representation from original face images is a crucial step [2]. Recently Taskeed Jabid and others [3] proposed Local directional pattern (LDP) method which was a valid appearance-based feature extraction method and applied successfully in face recognition, object description, gender recognition and facial expression recognition [4],[5],[6].

LDP feature extraction method can represent facial curve, edge and texture characteristic well. Like most existing appearance-based method, they only consider the whole facial features to classify expressions. In the changes of facial expression, it is undeniable that the regions of eyes and mouth have powerful influence. Therefore, how to weaken the influence of no-facial region and highlight the local region which has more contribution on expression changes is particularly important. Paper [7] introduced contrast information to LDP operator, and proposed a weighted LDP method (LDPv) for facial expression recognition. It adjusts the different contributions of LDP coding using the variance of local structure and account that texture with significant contract should impact more such as eyes and mouth that are more sensitive to high contract regions. So in this paper, we first extract global LDP features which can guarantee basic expression difference and decrease the influence of no-facial region meanwhile, then LDPv descriptor is used to extract local regions of eyes and mouth to extrude the distinction between expressions. After feature extraction, we don't use feature fusion with simple concatenation, we select decision-level fusion for global LDP feature and local LDPv feature by SVM to recognition respectively. Extensive results from two standard expression databases demonstrate the effectiveness of our proposed method.

The structure of the paper is organized as follows. In section 2, we will introduce some related work, feature extraction and expression recognition with SVM decision-level fusion. In section 3, we will do extensive experiments and analyze the results, and the conclusions are given in section 4.

2. Facial feature extraction and expression recognition

2.1 LDP and global feature extraction

The LDP descriptor is an eight bit binary code assigned to each pixel of an input image that can be calculated by comparing the relative edge response value of a pixel in different directions. So that eight directional edge response values $\{m_i\}$, $i = 0,1, 7$ of a particular pixel are computed using Kirsch masks [3] in eight different orientations M_i centered on its own position. Figure 1(a) and (b) shows eight directional edge response positions and LDP binary bit positions. Because different importance of the response values, the k most prominent directions are considered to generate the LDP. So the top k values $|m_j|$ are set to 1, and the other positions are set to 0. Finally, the LDP code is derived by formula (1), where m_k is the k-*th* most significant directional response value. Figure 1(c) shows an exemplary LDP code with k=3.

$$LDP_k = \sum_{i=0}^{7} b(m_i - m_k)\, 2^i, \quad b(a) = \begin{cases} 1, & a \geq 0 \\ 0, & a < 0 \end{cases} \qquad (1)$$

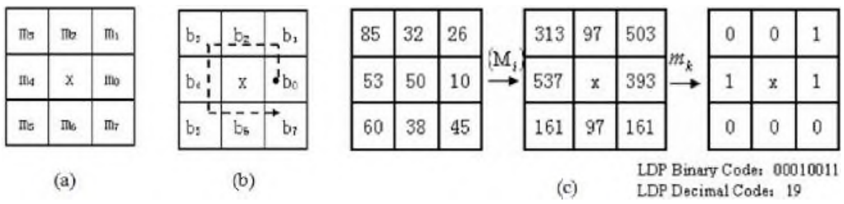

Fig.1.(a) Eight directional edge response positions; (b) LDP binary bit positions; (c) LDP Code with k=3

The input image of size M x N can be represented by an LDP histogram H using (2) after computing all the LDP code for each pixel (r, c), where i is the LDP code value.

$$H(i)=\sum_{r=1}^{M}\sum_{c=1}^{N} f(LDP_k(r,c),i),\quad f(a,i)=\begin{cases}1 & a=i\\ 0, & a\neq i\end{cases} \qquad (2)$$

For a particular value k, there has a different number of bins for the histogram H. In essence, a resulting histogram vector size of $1 \times c_i^f$ is produced for the image. Whereas computing LDP over the whole face image only considers the occurrences of micro-pattern without any information of their location and spatial relationship which usually represents the image content better. Hence, the image is divided into g regions $R_0, R_1, \cdots, R_{g-1}$ as shown in Figure 2 when using LDP, so that there will be a LDP_i histogram for every region R_i. Consequently, the resulting LDP descriptor is obtained via concatenating all the LDP_i histograms that is the global feature extraction in this paper.

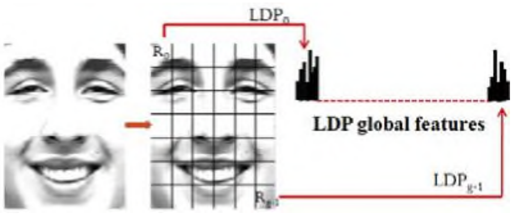

Fig.2. Expression image is divided into small regions from which LDP histograms are extracted and concatenated into LDP global features

2.2 LDPv and local feature extraction

As described in section of introduction, local directional pattern variance (LDPv) method was proposed which considered contrast information to LDP operator. Therefore, the variance a is introduced as an adaptive weight to adjust the contribution of the LDP code in the histogram generation. The proposed LDPv descriptor is computed as:

$$LDPv(\tau)=\sum_{r=1}^{M}\sum_{c=1}^{N} w(LDP_k(r,c),\tau) \qquad (3)$$

$$w(LDP_k(r,c),\tau)=\begin{cases}\sigma(LDP_k(r,c)) & LDP_k(r,c)=\tau\\ 0 & \text{otherwise}\end{cases} \qquad (4)$$

$$\sigma(LDP_k(r,c))=\frac{1}{8}\sum_{i=0}^{7}(m_i-\bar{m})^2 \qquad (5)$$

Where, $\bar{m}$ is the average of all directional responses $\{m_i\}$ calculated for a position (r, c). Same as LDP descriptor, in order to represent spatial relationship, the image is divided into many regions. In this paper we want to extract LDPv feature on local regions of eyes $(R_0 \sim R_{g-1})$ and mouth $(R_g \sim R_{7})$, as shown in Figure 3 shows, the formation of LDPv local feature. The final LDPv feature is obtained via concatenating all the $LDPv^i$ histograms which is built for each region R_i.

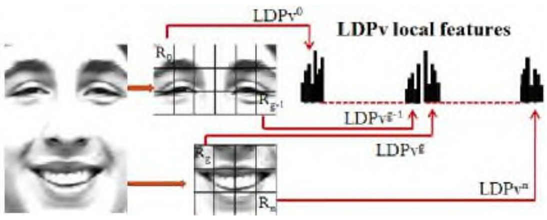

Fig.3. LDPv histograms are extracted from the regions of eyes and mouth and concatenated into LDPv local features

3.3 Feature dimensionality reduction and expression recognition

After feature extraction an inadequate number of feature vectors will be got. Then for global LDP features and local LDPv features we utilize PCA technique to reduce dimension respectively. SVM is a well-founded statistical learning theory that has been successfully applied in various classification tasks in computer vision. SVM makes binary decisions, and multi-class classification can be achieved by adopting the on binary classification: “one-against-all,” “one-against-one,” and DAGSVM. Some research [8] indicated that the “one-against-one” and DAG methods are more suitable for practical use than the other methods. So in our work we used the one-against-one technique, which was constructs $k(k-1)/2$ binary classifiers where each one is trained on data from two class expressions (anger-happiness, anger-fear, etc.). Then voting strategy was adapted and the output is the “MaxWins” expression.

In the stage of recognition, we adopt decision-level fusion. That’s, we use global LDP features and local LDPv features to train SVM model and predict classification result respectively, and then the respective voting result are added to decide “MaxWins” one as the final expression. The benefit of this is that we not only consider the effect of global facial characteristic but also highlight the important contribution of local facial region which influence the changes of expression greatly.

3. Experimental results and analysis

3.1 Databases and experiment setup

We will evaluate the proposed method on two benchmark databases: the JAFFE databases [9] and the Cohn-Kanade (CK) databases [10]. The JAFFE database contains 213 gray images (256x256) of individual human subjects with a variety of facial expressions. In this database, 10 different Japanese females performed seven prototypical expressions: anger, disgust, fear, joy, sadness, surprise and neutral. We choose three samples per facial expression for each subject, and a total of 210 images. The C-K database includes video sequences of 97 subjects displaying distinct facial expression. We create a subset with 10 subjects for our experiments. All the subjects selected have six basic expressions. From six sequences of expressions for a selected subject, we select the last four frames as six basic expressions and the first frame as neutral expression four of which are used. So there are 280 total images (640x490) in all. After choosing the images, they were cropped from the original one using the position of two eyes according to the distribution proportion of facial organs and resized into 150×110 pixels.

In our experiments 20 images per expression were selected randomly from JAFFE database for training and the rest images are used for testing, and it is the same for the Cohn-Kanade databases. For SVM the selection of kernel function is very important, here we used polynomial kernel with the degree of 1. In order to verify the effectiveness of various methods and get persuasive conclusion for every method the designed experiment is repeated 10 times, and the average recognition accuracy with best PCA dimension was picked.

4.2 Result and analysis

It can be found that there have two parameter k and g (the number of blocks) will affect the forming of LDP features. In related research of LDP, k=3 and g=7x6 is proved best for Cohn-Kanade. Here we first affirm the best k and g again for Cohn-Kanade and JAFFE databases before comparing our proposed method with others. So we investigate LDP features with SVM classifier when k=3 with different number of blocks as shown in Table 1.The results reflect that in Cohn-Kanade the best number of blocks is really g=7x6 but on JAFFE database g=9x8 is best with 88.43%. The reason may be the facial structure of JAFFE is different of Cohn-Kanade, and they belong to different human species.

Table 1. Recognitionperformance (%) for different number of blocks with k=3

blocks	C-K	JAFFE
g=3x3	90.43±4.4	76.71±4.2
g=5x5	92.64±3.4	80.86±5.0
g=7x6	**94.93**±3.5	87.29±2.8
g=9x8	94.42±4.4	**88.43**±5.0
g=12x9	93.78±5.1	87.29±4.8

With best parameter of LDP we can get our global features, and then we will find optimal weight for LDPv to obtain local features. Table 2 lists the result of the recognition performance of our proposed method with different weight $\sigma$for LDPv local features. From the result we can see that when weight is standard variance the best performance is achieved. Meanwhile, the recognition rate also improved when there has no weight or with forth root of variance but when the weight is variance. This only shows that overlarge weight will reduce the recognition rate instead.

Table 2. Recognition performance (%) with different weight

Weight	C-K	JAFFE
no weight	95.93±2.5	89.57±4.2
Variance	92.14±2.6	88.29±4.9
Standard variance	**96.86**±2.2	**92.14**±4.1

Forth root of variance	94.43±3.8	91.71±4.1

In order to verify the effectiveness of our proposed method with optimal parameter, we carry out experiments comparing with other existing methods on C-K, JAFFE and across-dataset (training on C-K and testing on JAFFE).Table 3 gives out the performance of various methods. The results indicate that the proposed method achieved better recognition accuracy. Furthermore, the recognition accuracy on across-dataset is all relatively low that still is the challenge of the future.

Table 3. Recognition performance (%) with different methods on different databases

	C-K	JAFFE	Train:CK,Test:JAFFE
LBP	89.17	83.92	41.90
LDP	94.93	88.43	42.86
LDPv	95.12	89.28	43.33
Proposed method	**96.86**	**92.14**	**45.71**

5. Conclusion

In this paper, we present a novel method for expression recognition based on global LDP features and local LDPv features with SVM decision-level fusion, which can retain the influence of global facial face and while highlight the local region with more contribution on expression changes. Extensive experimental results demonstrate the effectiveness of our proposed method. In future work we plan to increase the number samples on Cohn-Kanade and enhance the robustness and applicability.

References

1. Zhiguo Niu and Prof.Xuehong Qiu,"Facial expression recognition based on weighted principal component analysis and support vector machines" Int.Conf.on Advanced Computer Theory and Engineering.2010 IEEE.
2. Fasel, B., Luttin, J.: Automatic Facial Expression Analysis: a survey. Pattern Recognition36(1), 259–275 (2003)
3. Jabid, T., Kabir, M.H., Chae, O.S.: Local Directional Pattern (LDP) for Face Recognition.In: IEEE Int. Conf. Consum. Electron. 2010, pp. 329–330 (2010)
4. Jabid, T., Kabir, M.H., Chae, O.S.: Local Directional Pattern(LDP)-A robust Descriptor for Object Recognition. In: IEEE Int. Conf. on AVSS (2010) 978-0-7695-4264-5/10.
5. Jabid, T., Kabir, M.H., Chae, O.S.: Gender Classification using Directional Pattern (LDP).IEEE Pattern Recognition (2010)
6. Jabid, T., Kabir, M.H., Chae, O.S.: Robust Facial Expression Recognition Based on LocalDirectional Pattern. ETRI Jounal 32(5) (October 2010)
7. Kabir, M.H., Jabid, T., Chae, O.: A Local Directional Pattern Variance (LDPv) based FaceDescriptor for Human Facial Expression Recognition. IEEE (2010),doi:10.1109/AVSS.2010. 9978 -0-7695-4264-5/10 © 2010 IEEE

8. C.W. Hsu and C.J. Lin, "A Comparison on Methods for Multiclass Support Vector Machines," IEEE Trans. NeuralNetworks, vol. 13, no. 2, 2002, pp. 415-425.
9. Jaffe dataset, http://www.kasrl.org/jaffe.html
10. Kanade, T., Cohn, J., Tian, Y.: Comprehensive Database for Facial Expression Analysis.In: Proc. IEEE Int'l. Conf. Face and Gesture Recognition (AFGR 2000), pp. 46–53 (2000)