FACIAL EXPRESSION RECOGNITION USING CURVELET BASED LOCAL BINARY PATTERNS

Ashirbani Saha and Q.M. Jonathan Wu

Department of Electrical and Computer Engineering, University of Windsor, Windsor, ON, Canada
{sahai, jwu}@uwindsor.ca

ABSTRACT

This paper proposes the use of the combination of digital curvelet transform and local binary patterns for recognizing facial expressions from still images. The curvelet transform is applied to the image of a face at a specific scale and orientation. Local binary patterns are extracted from the selected curvelet sub-bands to form the descriptive feature set of the expressions. The average of the features of a particular class of expression is considered as the representative feature set of that class. The expression recognition is performed using a nearest neighbor classifier with Chi-square as the dissimilarity metric. Experiments show that our method yields recognition rates of 93% and 90% in JAFFE and Cohn-Kanade databases respectively.

Index Terms— Facial Expression Recognition, Curvelets, Local Binary Patterns

1. INTRODUCTION

Similarity in facial expressions is a distinct biological trait observed among human beings. Facial expression recognition has been a very active research area in computer vision. A survey of the research works on facial expression recognition in the last decade can be found in [1].

Despite of the fact that there has been good progress till now, facial expression recognition with high accuracy remains a challenging task due to the subtlety, complexity and variability of facial expressions [2]. The methods for facial expression recognition can be divided into two broad classes: recognition from video and recognition from still images. For both cases, the representation of facial features is crucial. The common methods for facial feature extraction are geometry-based and appearance-based [3]. Geometry-based methods rely on effective and accurate facial feature detection and tracking. Among the appearance-based methods, the potential of Gabor wavelets for recognizing expressions from still images have been established in [4, 5]. Recently, excellent face recognition results were reported in [6, 7] using the new multiresolution analysis method called digital curvelet transform. Generally, face recognition and facial expression recognition are dual problems. Face recognition is made difficult by variety of expressions and expression recognition gets tougher due to the faces varying in age, gender, and ethnicity. The method presented in [6] applied digital curvelet co-efficients to form features for representing the entire face. In order to classify the facial expressions, the local facial information needs to be stored. To obtain the local description of the expressions, local binary patterns (LBPs) are computed using selected sub-bands of image pre-processed by curvelet transform. LBP was proposed by T. Ojala in [8] for texture classification. LBP's have been used extensively for expression recognition with a good rate of success in [2, 9].

The organization of the paper can be summarized as follows. Section 2 discusses the digital curvelet transform briefly. Section 3 describes the LBP operator. Extraction of Curvelet based LBP features is discussed in section 4. The experimental results obtained by applying the proposed technique on JAFFE [10] and Cohn-Kanade [11] databases are summarized in section 5. Section 6 discusses the future scope of the proposed technique.

2. DIGITAL CURVELET TRANSFORM

Curvelet Transform was introduced by Candès and Donoho in 1999 [12]. The transform offers basis elements oriented in various directions thereby providing greater directionality than the wavelets.The basis elements obey the parabolic scaling law of width$\sim$length^2 allowing the transform to be anisotropic. Also, the curvelet transform refines the scale-space viewpoint by adding an extra element, orientation, and operates by measuring information about an object at specified scales and locations but only along specified orientations [13]. As explained by Candes, curvelets can be thought of as obtained by applying parabolic dilations, rotations and translations to a specifically shaped function ψ ; they are indexed using scale $a(0 < a < 1)$, location b and orientation θ as

$$\psi_{a,b,\theta}(x) = a^{-\frac{3}{4}}\psi(D_a R_\theta(x-b)) \tag{1}$$

with

$$D_a = \begin{pmatrix} \frac{1}{a} & 0 \\ 0 & \frac{1}{\sqrt{a}} \end{pmatrix}$$

where D_a is a parabolic scaling matrix, R_θ is a rotation by θ radians.

We know that domain transforms (time-frequency, space-frequency) are greatly affected by the presence of discontinuity. The sparsity of fourier series is affected by discontinuity. The wavelets, being multiscale and localized give a better representation in one-dimension but due to poor orientation selectivity fail to perform better in higher dimensions. The architecture of curvelet transform makes it such that it can display '*curve-punctuated smoothness*' – smoothness except for discontinuity along a general curve with bounded curvature [13]. Curvelets via USFFT (Unequally Spaced Fast Fourier Transform) [14] have been used in the experiments conducted by us. For a 2D function $f[t_1, t_2]$ with $0 < t_1, t_2 < n$ (a specific length), the main steps for wrapping are as follows[14]

1. The 2D Fast Fourier Transform (FFT) of $f[t_1, t_2] \rightarrow \hat{f}[n_1, n_2]$ is calculated.
2. $\hat{f}[n_1, n_2]$ is separated into dyadic subbands using a scale window W_j, j representing j^{th} scale.
3. Each subband is separated into angular wedges using angular windows $W_{j,l}$ with l representing l^{th} wedge.
4. The product $\tilde{U}_{j,l}[n_1, n_2]\hat{f}[n_1, n_2]$ is computed and $\tilde{U}_{j,l}[n_1, n_2]$ is a discrete localizing window.
5. The product is wrapped inside a rectangle W of size $L_{1,j} \times L_{2,j}$ (in east-west) or $L_{2,j} \times L_{1,j}$ (in north-south) around the origin to obtain $\tilde{f}_{j,l}[n_1, n_2] = W(\tilde{U}_{j,l}\hat{f})[n_1, n_2]$. Here, $L_{1,j} \sim 2^j$ and $L_{2,j} \sim 2^{\frac{j}{2}}$ and are constants.
6. The 2D inverse FFT of each $\tilde{f}_{j,l}$ is calculated to obtain curvelet coefficients at scale j and orientation l.

Our experiments have been carried out by applying the curvelet transform in different scales and orientation on the images. In section 5, we use 'curvelet (m, n)' to indicate curvelet co-efficients obtained at scale m and orientation n.

3. LOCAL BINARY PATTERNS(LBP)

The LBP operator, introduced by T. Ojala et al. [8], is a method used in texture description and analysis. Using the basic LBP operator, each pixel of an image is labeled by a binary (0 and 1) representation of its 3×3 neighborhood taking the center value as threshold. Then the histogram of the labels can be used as a texture descriptor. A histogram of the labeled image $f_l(x, y)$ can be defined as

$$H_i = \sum_{x,y} I\{f_l(x, y) = i\}, i = 0, 1, 2, ..., n - 1 \qquad (2)$$

in which n is the number of different labels produced by the LBP operator and

$$I\{A\} = \begin{cases} 1 & \text{if } A \text{ is true} \\ 0 & \text{if } A \text{ is false} \end{cases} \qquad (3)$$

This histogram has information about the distribution of the local micropatterns, such as edges, spots and flat areas of the entire image [8]. The number of labels clearly depends on the number of neighbors used in the LBP operator. For N neighbors, number of labels will be maximum 2^N. The operator was modified to use neighborhoods of different sizes in [15]. In the modified operator, the neighborhood of any pixel was defined as a circular region of radius R. Using bi-linear interpolation, a number of P(depending on R) pixels can be sampled on the circumference of the circular region using the center pixel as threshold. The LBP is therefore represented by LBP(P, R). If the coordinates of the center pixel are (0,0), the coordinates of the pth neighbor are given by ($-R\sin(\frac{2\pi p}{P})$, $R\cos(\frac{2\pi p}{P})$). Since the labels were formed using 8 neighbors, 2^8 labels are available, starting from 0 to 255.

4. PROPOSED METHOD

The images are first cropped to extract the face of the subject. LBP is independent of illumination changes whereas curvelet is not hence normalization is applied. Histogram equalization is then applied to increase the contrast. After that, curvelet transform is applied on the images at a specific scale and orientation. The approximate sub-band obtained from the transform is resized to $a \times b$ and divided into k regions each of size $m \times n$ pixels. From each of these k regions, the LBP histogram of 255 labels is calculated. The histograms from successive regions are concatenated to form the feature set for a particular image. Mathematically, we have k regions $G_1, G_2, ..., G_k$ each element of the feature vector can be expressed as

$$H_{i,j} = \sum_{x,y} I\{f_l(x, y) = i\}I\{(x, y)\epsilon R_j)\} \qquad (4)$$

Here, i=[0, 255] and j=[1, k]. So, effectively from the approximate sub-band, we obtain 3 levels of information. The LBP values have information about the coefficients. Histogram obtained from the LBP values over a region contain the information in a regional level. Finally, all the regional histograms are successively concatenated to obtain a holistic description of the sub-band.

We have the i^{th} feature vector x_i of length k×255. Based on the class labels y of the images, feature vectors of same class label are grouped to form the training set X^c for a particular class of expressions. Thus,

$$X^c = \{x_1^c, x_2^c, x_3^c, ..., x_n^c\} \qquad (5)$$

Table 1: Recognition Rates obtained by different Curvelet (Scale, Orientation) + LBP (Neighborhood, Radius) combinations on JAFFE and Cohn-Kanade Databases

Combination Details	Recognition Rates (%)	
	JAFFE	Cohn-Kanade
curvelet(3,16) + LBP(8,2)	93.69	90.33
curvelet(3,16) + LBP(8,3)	90.99	87.14
curvelet(3,16) + LBP(8,1)	90.86	88.03
curvelet(3,8) + LBP(8,2)	91.98	89.29
curvelet(3,8) + LBP(8,3)	90.11	87.14
curvelet(3,8) + LBP(8,1)	89.07	88.55
curvelet(2,8) + LBP(8,2)	91.75	86.81
curvelet(2,8) + LBP(8,3)	83.41	88.60
curvelet(2,8) + LBP(8,1)	89.74	87.09

where n is the number of training images available for the corresponding class. The representative feature set M^c of the class c is the cluster center of X^c and is calculated as

$$M^c = \frac{1}{n}\sum_{i=1}^{n} x_i^c \ . \tag{6}$$

A simple nearest neighbor classifier using the Chi-square metric (7) is used for classification.

$$\chi^2(S, M^c) = \sum_{i=1}^{N} \frac{(S_i - M_i^c)^2}{(S_i + M_i^c)} \tag{7}$$

Here, S is the feature vector of length N extracted from the test image.

5. EXPERIMENTAL RESULTS

The experiments were carried out on JAFFE and Cohn-Kanade databases. The details of the experiments are discussed in sub-sections 5.1 and 5.2.

5.1. JAFFE Database

The database consists of total 213 images of 7 facial expressions (6 basic facial expressions + 1 neutral, frontal view) posed by 10 Japanese female models. They are pre-processed as described in section 4. The images are randomly divided into five sets of roughly equal images. Five rounds of testing are carried out and at each time, a different combination of four sets are used for training and the remaining set is used for testing. This random division is carried out 3 times. The recognition rates given for this database in the experimental results are therefore the averages of the recognition rates of all tests. The recognition rates for various combinations of curvelet (Scale, Orientation) and LBP (Neighborhood, Radius) are discussed in Table 1. Since the combination of curvelet (3,16) and LBP (8,2) gives best results, the confusion matrix shown in Table 3 is calculated using this combination.

5.2. Cohn-Kanade Database

The Cohn-Kanade database consists of images (frontal view) of 100 university students of different age, gender and ethnicity. The subjects were instructed to perform a series of 23 facial displays, six of which were based on description of basic emotions. The facial displays are converted to basic 6 expressions where possible and a total 484 image sequences are selected. Only the final image of each of the selected sequences are considered for our training and testing. The testing procedure is the same as mentioned in subsection 5.1. The recognition rates given for this database in the experimental results are therefore the averages of the recognition rates of all tests conducted. The recognition rates for various combinations of curvelet (Scale, Orientation) and LBP (Neighborhood, Radius) are discussed in Table 1. In this database also, curvelet (3,16) and LBP (8,2) combination gives the best results. The confusion matrix obtained from this database for the said combination is presented in Table 4.

5.3. Comparative Study

A comparison of the proposed method with Gabor wavelets based method and LBP based method [2] is shown in Table 2. The proposed feature extractor gives better performance compared to Gabor wavelets and LBP based method using the chi-square based nearest neighbor classifier. It is observed that the results improve more in the JAFFE database when compared to that in Cohn-Kanade Database. This may be attributed to the fact that Cohn-Kanade has more variations in ethnicity and gender compared to the JAFFE database.

Table 2: Comparison of proposed method with other methods

Methods	Classification Rates(%)	
	JAFFE	Cohn-Kanade
Gabor Wavelets	82.46	86.9
LBP based method	78.36	84.5
Proposed Method	93.69	90.33

6. CONCLUSION

We present a novel approach of using the digital curvelet transform and local binary patterns for recognition of expressions from still face images. The method has showed promising results in JAFFE and Cohn-Kanade databases and this demonstrates the efficiency of curvelet based LBP as a good feature extractor. There is a scope for detailed study of the proposed feature extractor using various classifiers. Future work may include the application of this method to other specific object recognition tasks like fingerprint recognition, handwriting recognition. Also the effect of using the combination of digital curvelet transform and other advancements of LBP like uniform LBP, rotation invariant uniform LBP and

Table 3: Confusion Matrix obtained by Curvelet (3, 16) and LBP (8, 2) on JAFFE Database

Given Label(%) \ Classified Label(%)	Happy	Sad	Surprise	Anger	Disgust	Fear	Neutral
Happy	95.34	0	0	2.33	0	0	2.33
Sad	0	85.71	0	0	0	14.29	0
Surprise	2.44	2.44	87.8	0	0	4.88	2.44
Anger	0	0	3.45	89.65	0	3.45	3.45
Disgust	0	0	3.57	0	92.86	3.57	0
Fear	5.89	2.94	2.94	0	2.94	85.29	0
Neutral	0	0	6.9	0	0	6.9	86.2

Table 4: Confusion Matrix obtained by Curvelet (3, 16) and LBP (8, 2) on Cohn-Kanade Database

Given Label(%) \ Classified Label(%)	Happy	Sad	Surprise	Anger	Disgust	Fear
Happy	83.87	3.23	4.35	1.02	0	7.53
Sad	3.33	88.33	1.67	1.67	3.33	1.67
Surprise	2.53	0	93.67	0	0	3.8
Anger	0	10.53	0	86.84	2.63	0
Disgust	0	2.38	0	0	95.24	2.38
Fear	6.97	2.33	4.65	0	0	86.05

advanced LBP (ALBP) for similar recognition tasks can be studied.

7. ACKNOWLEDGEMENT

The work is supported in part by the Natural Sciences and Engineering Research Council of Canada.

8. REFERENCES

[1] B. Fasel and J. Luettin, "Automatic facial expression analysis: A survey," *Pattern Recognition*, vol. 36, pp. 259–275, 1999.

[2] C. Shan, S. Gong, and P.W. McOwan, "Facial expression recognition based on local binary patterns: A comprehensive study," *Image and Vision Computing*, vol. 27(4), pp. 803–816, 2008.

[3] Y. Tian, T. Kanade, and J. Cohn, *Handbook of Face Recognition*, chapter 11, Springer, 2005.

[4] M.S. Bartlett, G. Littlewort, M. Frank, C. Lainscsek, I. Fasel, and J. Movellan, "Recognizing facial expression: Machine learning and application to spotaneous behavior," in *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2005, vol. 2, pp. 568–573.

[5] M.J. Lyons, J. Budynek, and S. Akamatsu, "Automatic classification of single facial images," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 21(12), pp. 1357–1362, 1999.

[6] T. Mandal, A. Majumdar, and Q.M.J. Wu, "Face recognition by curvelet based feature extraction," in *International Conference on Image Analysis and Recognition, LNCS*, 2007, vol. 4633, pp. 806–817.

[7] A.A. Mohammed, R. Minhas, Q.M.J. Wu, and M.A. Sid-Ahmed, "A novel technique for human face recognition using nonlinear curvelet feature subspace," in *International Conference on Image Analysis and Recognition, LNCS*, 2009, vol. 5627, pp. 512–521.

[8] T. Ojala, M. Pietikäinen, and D. Harwood, "A comparative study of texture measures with classification based on featured distribution," *Pattern Recognition*, vol. 29 (1), pp. 51–59, 1996.

[9] S. Liao, W. Fan, C.S. Chung, and D.Y. Yeung, "Facial expression recognition using advanced local binary patterns, tsallis entropies and global appearance features," in *IEEE International Conference on Image Processing*, 2006, pp. 665–668.

[10] M.J. Lyons, S. Akamatsu, M. Kamachi, and J. Goba, "Coding facial expressions with gabor wavelets," in *IEEE International Conference on Automatic Face and Gesture Recognition*, 1998, pp. 200–205.

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

[12] D. L. Donoho and M.R. Duncan, "Digital curvelet transform: Strategy, implementation and experiments," Tech. Rep., Stanford University, 1999.

[13] E.J. Candès, "What is... a curvelet?," in *Notices of the American Mathematical Society*, 2003, vol. 50(11), pp. 1402–1403.

[14] E.J. Candès, L. Demanet, D.L. Donoho, and L. Ying, "Fast discrete curvelet transforms," in *Multiscale Modeling and Simulation*, 2006, vol. 5, pp. 861–899.

[15] T. Ojala, M. Pietikäinen, and T. Mäenpää, "Multiresolution gray-scale and rotation invariant texture classification with local binary patterns," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 24(7), pp. 971–987, 2002.