Curvelet Entropy for Facial Expression Recognition

Ashirbani Saha and Q.M. Jonathan Wu

University of Windsor, Ontario, Canada

Abstract. This paper proposes the use of curvelet entropy for classifying facial expressions from still images. The idea behind this work is that the expressions impose non-rigid motions on the face thereby changing the orientations of facial curves occurring due to different types of expressions. Hence a multiresolution transform like curvelet which refines its domain by using orientation information may be applied for the task of expression classification. Since similarity of facial expressions has earlier been studied using Gabor wavelet which uses filters oriented in different directions on specific feature points in images, the orientation selectivity and information content of curvelet subbands at specific facial points are used here. The information at selected facial points are gathered using the entropy of the corresponding pixel at various subbands. The proposed method is evaluated in the JAFFE and Cohn-Kanade databases without and with cross-validations. Experimental results show that the curvelet subband entropy at selected points may be used to form effective features for classifying facial expressions.

1 Introduction

Facial expression recognition is an active research field for almost 20 years now. The field has achieved various breakthroughs till now in terms of research success and on availability of different databases for algorithm testing and verification. Surveys on facial expression recognition can be found in [7,10].

There had been many researches on facial expression recognition from still images, image sequences and videos. The approaches used are classified as geometry-based and appearance based. In appearance based approaches, a dichotomy is evident- the use of entire face [11] and use of selected facial points [8,14]. Methods based on both of the approaches have achieved success in the past years.

Recently, facial images were decomposed by curvelet transform to form effective feature set for face recognition with high success rate [9]. Also, dual-tree complex wavelet transform was used for extracting facial features [13]. Both of these multiresolution analysis methods have direction selectivity. The subbands offered by these multiresolution representations (MRs) have better orientation details. Motivated by these approaches of applying MRs for extracting facial features, this paper attempts to use all subbands of the curvelet representation of a facial image at selected points for classifying the expressions.

G. Qiu et al. (Eds.): PCM 2010, Part II, LNCS 6298, pp. 617–628, 2010.

The points selected on the face are those points affected due to the six basic expressions of happiness, sadness, surprise, anger, disgust and fear. The coefficients at a specific point of the original signal in any subband of its MR are the measure of the strength of the point under the parameters of the corresponding subband. It has been established already in [3,12] that curvelets are very effective in representing edges and ridges. Hence, points around which edges and ridges occur due to expressions will have significant energy in specific subbands. Thus, the entropy at the selected point at any subband will differ due to different expressions. Earlier, wavelet energy and entropy obtained from all available subbands have been used for facial feature extraction [4,5]. Since, the entire face has been considered for energy calculations in those methods, the global facial energy which has the details of a person's identity as well, is being used to generate the features. In order to eliminate that problem, the proposed approach is based on selecting 36 facial points from a face and analyzing the subband entropies at those points to form an effective feature set. Different neighbourhoods of a chosen point in curvelet domain are used to get the entropy. The extracted entropies are projected to principal component analysis (PCA) space. The vectors in the PCA space are subjected to linear discriminant analysis (LDA). A simple nearest neighbour classifier is used for classification. Experiments are carried out without cross-validation and with cross-validation in the JAFFE and Cohn-Kanade database and interesting results are obtained.

The rest of the paper is arranged as follows. Section 2 discusses the curvelet transform briefly. The proposed method is elaborated in section 3. Details of the experiments carried out in two databases and the summary of their results are presented in section 4. Conclusion and scope for future work are discussed in section 5.

2 Curvelet Transform

Curvelet transform was introduced by Candès and Donoho. Mathematical details of the transform can be found in [6]. The basis elements of the transform are anisotropic and they abide by the parabolic scaling law of $width \sim length^2$. This property takes a major role in obtaining sparse representation of smooth functions and straight edges by the curvelets. For obtaining curvelet transform of an image, at first, the image is represented in scale-space domain. Then the frequency plane obtained is divided into dyadic coronae and each corona is partitioned into angular wedges which abide by the parabolic aspect ratio. Hence, orientation is taken into account in the scale-space description of the image. Thus, curvelets provide coefficients by measuring information about an object at specified scales, locations as well as orientations. As explained by Candes, curvelets can be thought of as obtained by applying parabolic dilations, rotations and translations to a specifically shaped function Φ which is indexed using scale $s(0 < s < 1)$, location l and orientation β as

$$\Phi_{s,l,\beta}(x) = s^{-\frac{3}{4}}\Phi(D_a R_\beta(x - l)) \quad (1)$$

with

$$D_s = \begin{pmatrix} \frac{1}{s} & 0 \\ 0 & \frac{1}{\sqrt{s}} \end{pmatrix}$$

where D_s is a parabolic scaling matrix, R_β is a rotation by β radians.

There are some digital implementations of curvelets and *curvelets via wrapping* [2] is used for this work. For a 2D function $g[x_1, x_2]$ with $0 < x_1, x_2 < n$ (a specific length), the curvelets via wrapping can be performed according to the following steps[2]

1. 2D Fast Fourier Transform (FFT) for $g[x_1, x_2] \rightarrow \hat{g}[w_1, w_2]$ is computed.
2. $G[w_1, w_2]$ is divided into dyadic subbands using a scale window V_i, i representing i^{th} scale.
3. Each subband is separated into angular wedges using angular windows $V_{i,k}$ with k representing k^{th} wedge.
4. The product $\tilde{P}_{i,k}[w_1, w_2]G[w_1, w_2]$ is calculated where $\tilde{P}_{i,k}[w_1, w_2]$ is a discrete localizing window.
5. The product is wrapped inside a rectangle R of size $H_{1,i} \times W_{2,i}$ (in east-west) or $H_{2,i} \times W_{1,i}$ (in north-south) around the origin to obtain $\tilde{G}_{i,k}[w_1, w_2] = R(\tilde{U}_{j,l}G)[w_1, w_2]$. Here, $H_{1,i} \sim 2^i$ and $H_{2,i} \sim 2^{\frac{i}{2}}$ and are constants.
6. The curvelet coefficients at scale i and orientation k are calculated from the 2D inverse FFT of each $\tilde{G}_{i,k}$.

Curvelet transform is applied in different scales and orientations for this work. The performance of the proposed method improves when the decomposition is carried out in more orientations for a fixed scale.

3 Proposed Method

As this work is based on facial features using dimension reduced curvelet entropy of selected points, the proposed method comprises of three main parts- facial feature points selection, calculation of subband entropy followed by dimension reduction and classification. All the images are subjected to tilt correction before selecting the points.

3.1 Facial Feature Point Selection

36 specific facial points are chosen manually from each of the faces for processing. 34 of these points are similar to those used in [14], 2 other points are used in this work. All the 36 points are shown in Fig. 1. Points marked in red are two additional points considered for this work, as it is found by visual inspections that curves around these points vary significantly with different expressions. Also, the green point is a shifted form of the point at the tip of the nose owing to its changing proximity with the mouth for different expressions. The locations of these points are stored for further processing.

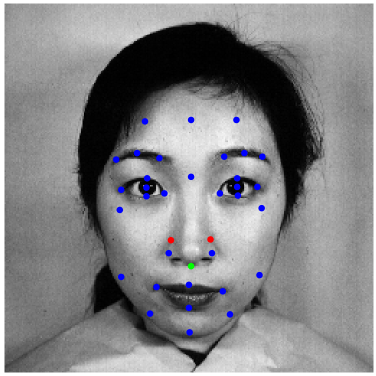


Fig. 1. 36 facial points selected for feature extraction

3.2 Curvelet Subband Entropy

After the selection of 36 specific facial points, the face images are represented in curvelet domain with specific scale and orientation. Figure 2 shows the subbands of an image obtained by curvelet transform. It is also evident from the

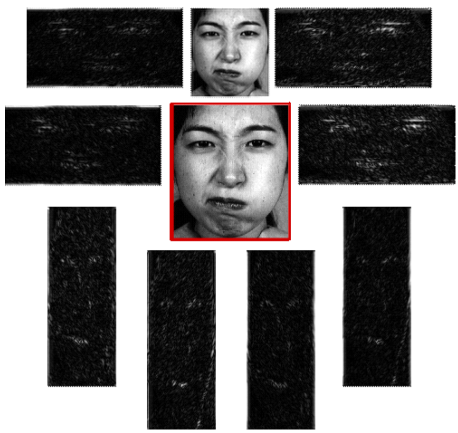


Fig. 2. Curvelet subbands obtained from the image (bordered by red) at scale = 2 and orientation = 8

figure that each subband has different information about the face. The number of subbands obtained depend on the scale and orientation of the curvelet transform. The corresponding points of the selected or interest points are found out in each subband. All the curvelet subbands do not have same dimensions, hence the location of the corresponding point of an interest point varies in each subband. The corresponding points are identified in each subband. In each subband, centering each interest point, four energy values are calculated by taking square regions of sizes 3×3, 5×5, 7×7 and 9×9 according to eqn. 2.

$$E_W = \sum_W |S_n(x, y)|^2 \quad \text{for } W \ = \ 3, 5, 7 \text{ and } 9 \tag{2}$$

Here, $S_n(x, y)$ the coefficient at the point (x, y) in subband S_n. The energy values E_W are normalized by dividing them by the total energy E_T of the original image I where

$$E_T = \sum_I |I(p, q)|^2. \tag{3}$$

The normalized energies are denoted by E_{WN} where

$$E_{WN} = \frac{E_W}{E_T} \quad \forall\ W. \tag{4}$$

The normalized energies thus become higher as the W increase. This is obvious for all the interest points. In order to have an idea about the neighbouring coefficients with the increase of W, normalized energy per point E_{WNP} is calculated for all window sizes by eqn. 5.

$$E_{WNP} = \frac{E_{WN}}{W^2} \quad \forall\ W \tag{5}$$

Here the distribution of the energy of an interest point is obtained by E_{WNP}. E_{WNP} will have higher values if the corresponding point in any subband is surrounded by more number of significant coefficients. The nature of this energy distribution is formulated by entropy which highlights uniform energy variation in all windows when large in value. Same expressions will yield similar orientation curves which imply similar variation in energy distribution surrounding the interest points in same subband. Therefore entropies of interest points are used here. Before calculating the entropy, the sum of all energy values obtained for a point is made one by eqn. 6 given as

$$E_{WNPN} = \frac{E_{WNP}}{\sum_{W=3,5,7,9}(E_{WNP})}. \tag{6}$$

Finally, the entropy for the i^{th} interest point in the j^{th} subband is calculated by the following equation

$$ENT_{i,j} = \sum_W E_{WNPN} \log_2 E_{WNPN} \tag{7}$$

Thus from each subband 36 values of entropy will be obtained. If N subbands are present, the feature vector for an image is obtained by concatenating the entropy values from all subbands. Hence, the size of the feature vector is $36 \times N$.

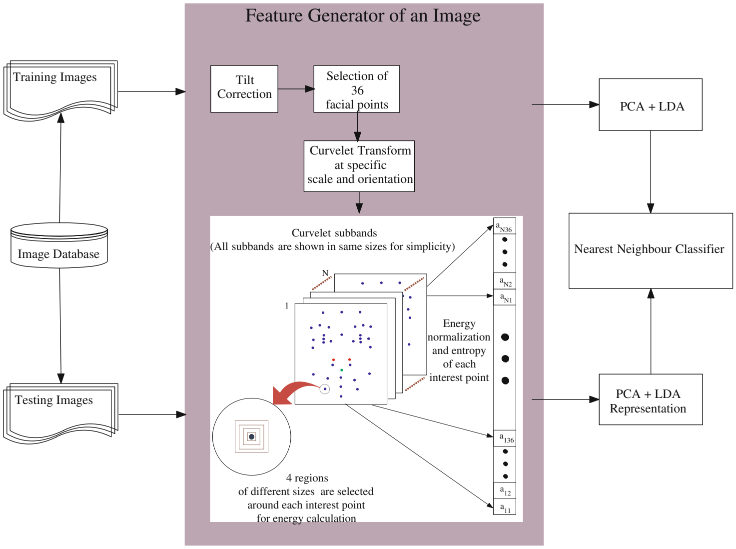


Fig. 3. The proposed method

3.3 Dimension Reduction and Classification

The feature vectors obtained are of large size and hence dimension reduction methods are applied. The expression recognition task performed here classifies images into one of the 7 (6 basic expressions + 1 neutral) classes. Since LDA works by discriminating between the classes we preferred to use them. As the number of training samples are lesser than the length of the vector, a singular $within-class\ scatter\ matrix$ will be generated. Thus, the vectors are projected to PCA space at first. The data in PCA space is projected to LDA space to be classified by a nearest neighbour classifier. The process diagram of the proposed method can be summarized as shown in Fig. 3.

4 Experiments and Results

The experiments are carried out in the JAFFE and Cohn-Kanade databases. Experiments are performed with and without cross-validations to evaluate the proposed method. The JAFFE database has 213 images of 10 female subjects. Each image shows a neutral face or one of the 6 basic expressions. All of the images of this database are used for training and testing. The Cohn-Kanade database has image sequences of different expressions of 100 university students, both male and female and of different ethnicity. We have chosen 303 sequences from 86 subjects with only condition that the sequences can be labeled as one of the six basic expressions. We take the last image of every sequence. Also, neutral faces are chosen and total number of images we worked with is 355 and each subject chosen does not have all six expressions. Curvelet transform is applied for four different combinations of scale (s) and orientation (o) values for both types of experiments in the two databases.

4.1 Experiments without Cross-Validation

For JAFFE database, each subject has at least one image of each expression. Therefore, by taking 7 different expressions (including neutral) from each subject a test set 70 images is formed. Rest of the images are used for training. This testing scheme is the same as used in [1]. This procedure of generating the test image set is done five times and it is ensured that every image of the database becomes a part of the testing set for at least once. Fig. 4 shows the recognition results in JAFFE database. We find an increase in accuracy when the number of orientations are increased while keeping the scale constant. The recognition rates are plotted against the number of principal components and though the rates increase with the number of components, the intermediate fluctuations are evident.

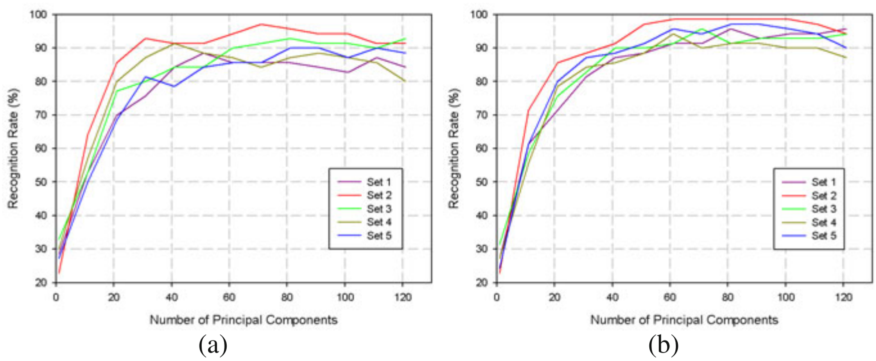


Fig. 4. Recognition rates in JAFFE database for (a) $s = 3$, $o = 32$, (b) $s = 4$, $o = 16$

The testing set of Cohn-Kanade database is formed by taking one image of every subject such that each test set has roughly same number of images of each of 7 expressions. Rest of the images are used for training. Total five different test sets are formed to ensure that each image of the chosen set falls in the testing set for at least once. The results are shown in Fig. 5. The accuracy is remarkably lower than that of JAFFE database. It may be attributed to the fact that since instances of all expressions of a subject are not present in the training set (unlike JAFFE), the training set needs to be increased for better results. By looking at this, the experiments are carried out with cross-validation which shows that our claim here regarding the these results in Cohn-Kanade database holds valid.

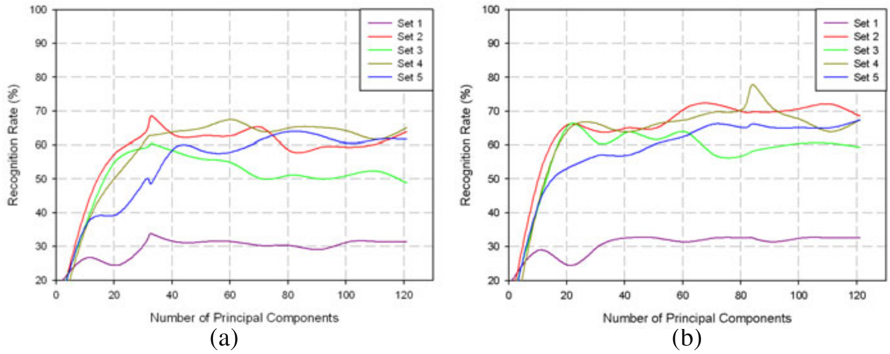


Fig. 5. Recognition rates in Cohn-Kanade database for (a) $s = 3$, $o = 32$, (b) $s = 4$, $o = 16$

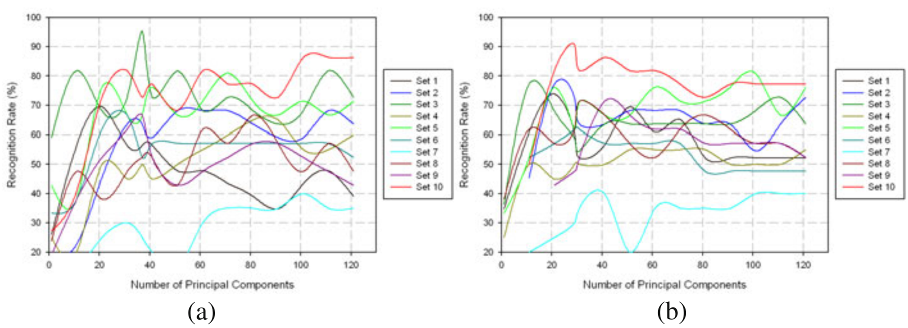


Fig. 6. Cross-validated recognition rates in JAFFE database for (a) $s = 3$, $o = 32$, (b) $s = 4$, $o = 16$

4.2 Experiments with Cross-Validation

For the JAFFE database, 10-fold cross-validations are used and it is similar to *leave one out* strategy since only 10 subjects are present. In each round, all images of a subject are used for testing and images of 9 other subjects are used for training. The recognition rates are shown in Fig. 6 and it is clear that the results depend greatly on the subject under testing. The recognition rates are quite irregular in this database. Thus the average recognition rate is pretty low which is also shown in [11].

A 10-fold cross-validation is conducted in Cohn-Kanade database as well. In each round, the testing set consists of 8-9 subjects and number of test images are below 50 always. The results are shown in Fig. 7. The results obtained are better than that of the non-cross-validated ones. This supports that with enough training images, the proposed method is going to perform better. Also, recognition rates for different sets are quite regular compared to that in the JAFFE database. The cross-validated results indicate the how well the method can perform to recognize face independent expressions. This implies that regularity or uniformity in the expressions is also very important. Therefore, more number of training images can improve results significantly. As already expressed in [11], the small size of the JAFFE database may be responsible for low and less regular recognition rates. More training images in Cohn-Kanade database give better cross-validated results.

4.3 Comparison

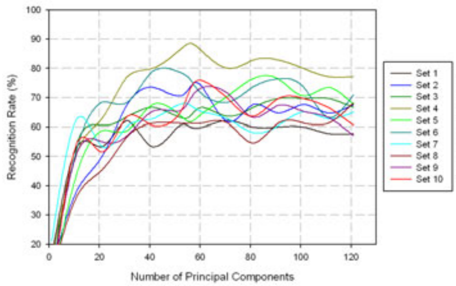
In this part the curvelet based entropy is compared with Gabor wavelet based entropy and LBP + PCA + LDA [11] method which is one of the popular methods in literature. For Gabor wavelets, 8 orientations of Gabor filter with wave numbers $\pi/4$, $\pi/8$ and $\pi/16$ is used and this is similar to [14]. For LBP, uniform LBP(8,2) operator is used. Both cross-validated and non-cross-validated experiments are carried out in JAFFE and Cohn-Kanade database using the same sets of train and test images as mentioned in sub-sections 4.1 and 4.2. The results are presented in Fig. 4.3 and 4.3. It is observed that with Gabor wavelet entropy and LBP, the trend is similar to what we got with curvelets.

It is found from Fig. 4.3, that directional multiresolution entropy performs better than Gabor wavelet entropy, though the general trend of the recognition rates are similar. For non-cross-validated experiments in JAFFE database, the recognition rates are higher and uniform. The cross-validated rates in JAFFE are irregular and much lower with Gabor wavelets. For Cohn-Kanade database the recognition rates for cross-validated and non-cross-validated experiments are irregular though cross-validated recognition rates are comparatively more uniform. Similar trend is found using LBP, though the rates are much better than that obtained using Gabor wavelets. Curvelet entropy is comparable to LBP based recognition. For non-cross-validated experiments in JAFFE database, directional multiresolution performs better than LBP whereas for cross-validated experiments in Cohn-Kanade, LBP performs better. The maximum recognition rates offered by curvelet entropy for in JAFFE database is better than that of LBP + PCA + LDA and in Cohn-Kanade database the reverse is true. These maximum recognition rates are summarized in Tables 1 and 2 for clarity.

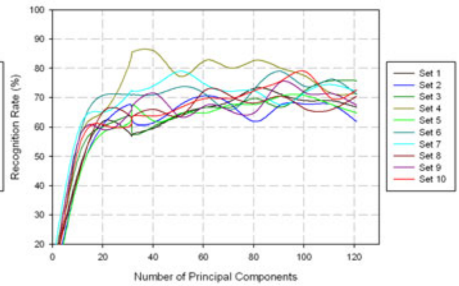
5 Conclusion and Scope for Future Work

A novel approach is proposed for classifying facial expressions from still facial images using curvelet subband entropy at selected points. The approach gives better results in JAFFE database for experiments without cross-validation. For cross-validated experiments, the performance in Cohn-Kanade database is superior. The results also reveal that curvelet based entropy can prove to be a better feature extractor than Gabor wavelet based entropy.

This research contributes in two aspects. The first aspect is the exploration the curvelet subbands by entropy for generating effective facial features used in

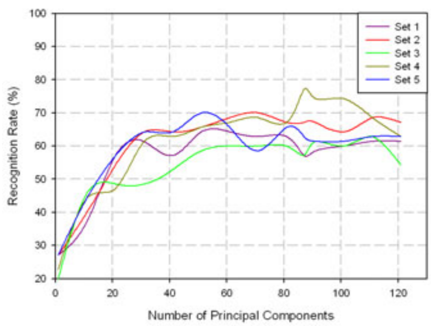


(a)

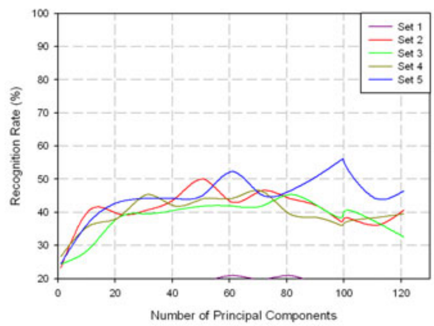


(b)

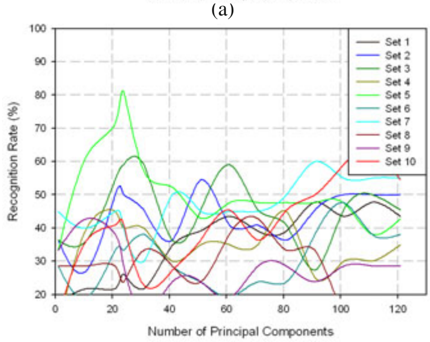
Fig. 7. Cross-validated recognition rates in Cohn-Kanade database for (a) $s = 3$, $o = 32$, (b) $s = 4$, $o = 16$



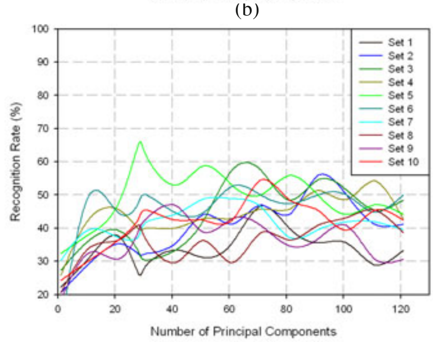
(a)



(b)



(c)



(d)

Fig. 8. Gabor entropy based recognition in (a) JAFFE non-cross-validated, (b) Cohn-Kanade non-cross-validated (c) JAFFE cross-validated and (d) Cohn-Kanade cross-validated

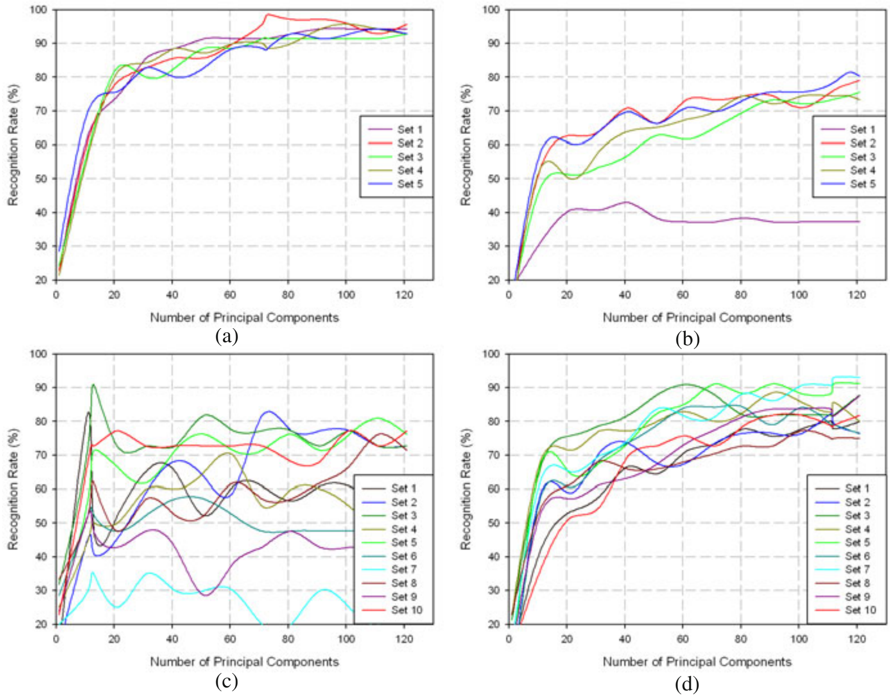


Fig. 9. LBP + PCA + LDA based recognition in (a) JAFFE non-cross-validated, (b) Cohn-Kanade non-cross-validated (c) JAFFE cross-validated and (d) Cohn-Kanade cross-validated

Table 1. Comparison maximum recognition rates (%) for experiments without cross-validation

Database	Curvelet entropy	Gabor Wavelet Entropy	LBP + PCA + LDA
JAFFE	98.57	77.14	97.17
Cohn-Kanade	77.91	55.81	81.41

Table 2. Comparison of maximum recognition rates (%) for experiments with cross-validation

Database	Curvelet entropy	Gabor Wavelet Entropy	LBP + PCA + LDA
JAFFE	95.45	80.95	90.91
Cohn-Kanade	88.57	64.81	93.42

expression recognition. The second part of the contribution lies in conducting experiments. The experiments are carried out with and without cross-validations. Therefore, the analysis and comparison of results brought out by both of these approaches can be made simultaneously. Thus, this study indicates that further research can be directed towards using the curvelet representation of facial images effectively with other classifiers for robust facial expression recognition.

References

1. Bashyal, S., Venayagamoorthy, G.K.: Recognition of facial expressions using gabor wavelets and learning vector quantization. Engineering Applications of Artificial Intelligence 21(7), 1056–1064 (2008)
2. Candès, E., Demanet, L., Donoho, D., Ying, L.: Fast discrete curvelet transforms. Multiscale Modeling and Simulation 5(3), 861–899 (2006)
3. Candès, E.J., Donoho, D.L.: Curvelets, multiresolution representation and scaling laws. Wavelet Applications Signal Image Processing 4119, 1–12 (2000)
4. Chen, C., Jhang, Z.: Wavelet energy entropy as a new feature extractor for face recognition. In: International Conference on Image and Graphics, pp. 616–619 (2007)
5. Chen, F., Wang, Z., Xu, Z., Xiao, J.: Facial expression recognition based on wavelet energy distribution feature and neural network ensemble. In: Global Conference on Intelligent Systems, vol. 2, pp. 122–126 (2009)
6. Donoho, D., Duncan, M.: Digital curvelet transform: Strategy, implementation and experiments. Technical report, Stanford University (1999)
7. Fasel, B., Luettin, J.: Automatic facial expression analysis: A survey. Pattern Recognition 36, 259–275 (1999)
8. Lyons, M.J., Akamatsu, S., Kamachi, M., Gyoba, J.: Coding facial expressions with gabor wavelets. In: International Conference on Automatic Face and Gesture Recognition, pp. 200–205 (1998)
9. Mandal, T., Wu, Q.M.J., Yuan, Y.: Curvelet based face recognition via dimension reduction. Signal Processing 89(12), 2345–2353 (2009)
10. Pantic, M., Rothkrantz, J.: Automatic analysis of facial expressions: State of the art. IEEE Transactions on Pattern Analysis and Machine Intelligence 22(12), 1424–1445 (2000)
11. Shan, C., Gong, S., McOwan, P.W.: Facial expression recognition based on local binary patterns: A comprehensive study. Image and Vision Computing 27(4), 803–816 (2008)
12. Starck, J., Candès, E.J., Donoho, D.L.: The curvelet transform for image denoising. IEEE Transactions on Image Processing 11(6), 670–684 (2000)
13. Celik, T., Ozkaramanli, H., Demirel, H.: Facial feature extraction using complex dual-tree wavelet transform. Computer Vision and Image Understanding 11(1), 229–246 (2008)
14. Zhang, Z., Lyons, M., Schuster, M., Akamatsu, S.: Comparison between geometry-based and gabor wavelets-based facial expression recognition using multi-layer perceptron. In: International Conference on Automatic Face and Gesture Recognition, pp. 454–459 (1998)