



CS-ONPP: Class Similarity Based ONPP for the Modular Facial Expression Recognition

Sujata^(✉), Purvi A. Koringa, and Suman K. Mitra

Dhirubhai Ambani Institute of Information and Communication Technology,
Gandhinagar, India
{201521003,201321010,suman_mitra}@daiict.ac.in

Abstract. Facial expression recognition is a big problem in the field of Human behavioral analysis. Much work has been done in this field where local texture, features have been extracted and used in the classification. Due to the very local nature of this information, the dimension of the feature vector achieved for the full image is very high, posing computational challenges in real-time expression recognition. In recent times, Dimensionality Reduction methods have been successfully used in image recognition tasks. Though being high dimensional data, natural images such as face images lie in low dimensional subspace, and Dimensionality Reduction methods try to learn this underlying subspace to reduce the computational complexity involved in classification stage of image recognition task. This paper proposes the Orthogonal Neighborhood Preserving Projection with Class Similarity-based neighborhood for expression recognition. The extensive experiment has been performed on three well-known Facial Expression Databases. The proposed techniques achieve similar recognition performance at very low dimensions compared to local feature extraction based methods.

Keywords: Facial expression recognition · Dimensionality reduction · ONPP

1 Introduction

Facial expressions, poses, and gestures are the non-verbal cues, which are an essential part of our daily communication where the face is the most prominent. Facial expression through expressing feelings common means of human communication, it indicates their agreement or disagreement, express intentions, interacting with others as well as environment [1]. Thus, many research is being done in the field of automated expression recognition. Numerous approaches for facial expression recognition have been suggested in the literature, which can be broadly classified into feature-based techniques and appearance-based techniques [2]. Local image descriptors like SIFT(Scale Invariant Feature Transform) [3] has been extensively examined and has assumed a predominant job in FER applications. SIFT has achieved the best overall performance as local image descriptors in the literature. It improves effectiveness while diminishing

the memory prerequisites. Impressed by the SIFT [4] presented Speed Up Robust Features (SURF). Instead of using the gradient information like SIFT descriptor does, SURF computes responses of Haar Wavelet and exploits integral images to save computational cost. As a result, it runs faster than SIFT. Dalal and Triggs [5] introduced the Histogram of Oriented Gradient (HOG) descriptor. For a long time, Gabor wavelet [6] have been comprehensively utilized in face feature extraction. It is capable of extracting the multi-direction and multi-scale statistics from the given face image. It is able to extract the multi-orientation and multi-scale information from the given face image. Gabor features has the huge feature dimensions, so it is not useful for the real time applications. [7] proposed Local Binary Patterns (LBP) for texture classification, and such a descriptor encodes texture information by comparing the central pixel and its neighborhood. It was successfully applied in the task of face recognition [8] and object categorization [9]. LBP is sensitive to local illumination variations. It does not correctly identify the texture of facial muscles and other types of local deformations, thus not very robust for FER applications. Another local feature based method based on Histogram of Second order Gradient (HSOG) [10].

Above mentioned work basically works based on local texture information, the local texture of face image is extracted with various image processing procedures, followed by histogram pooling by various block sizes. These approaches usually results in the feature vector length of order 10^5 making classification step computationally complex. To overcome problem of high dimensionality, in recent times many Dimensionality Reduction (DR) techniques are proposed, these techniques are proved to be very efficient in image recognition tasks. In this article, we are using one state of the art DR method - ONPP for facial expression recognition. The paper mainly proposes a variant of ONPP suitable for facial expression recognition. The main advantage achieved in the DR based method is computational complexity reduction in the classification. Using ONPP, the image is represented using feature vector 100s of dimensions instead of the feature vector having 10^5 dimensions. The performance of proposed method is tested on three bench-mark databases: CK+, Jaffe and Video database. Proposed method achieves similar recognition accuracy with comparatively very low dimensions than that of local feature based methods.

Remaining part of the study is held as follows. Section 2 explains dimensionality reduction methods employed successfully in image recognition tasks, followed by detailed discussion on Orthogonal Neighborhood Preserving Projection in Sect. 3. Section 4 documents expression recognition experiments performed on well-known facial expression databases, followed by the conclusion in Sect. 5.

2 Dimensionality Reduction for Image Recognition

Natural images are considered as very high-dimensional data, but as proved in [11], they usually reside on relatively low dimensional linear or non-linear manifold. The fundamental idea of dimensionality reduction is to seek this linear/non-linear manifold where data representation is compact and more meaningful, in

the sense that the learned low dimensional representation reveals the variability of the data more clearly. During last two decades many dimensionality reduction methods are employed in image recognition.

Among these methods appearance based DR methods consider the image of size $m \times n$ as a mn -dimensional data point in \mathcal{R}^{mn} and learn the low dimensional image manifold $\mathcal{M} \in \mathcal{R}^d$ where $d \ll mn$. DR methods are broadly classified in two categories: Non-linear DR methods and Linear DR methods. Linear Embedding (LE) [11], Laplacian Embeddings (LE) are non-linear DR methods. The mapping of data points achieved by these methods are non-explicit thus making recognition of an *out-of-sample* image data impossible. The solutions to overcome non-explicit mapping are proposed in Locality Preserving Projections (LPP) [12], Neighborhood Preserving Projections (NPP) where training samples are used to learn a projection matrix which can be used to project *out-of-sample* test data on the learned projection space. These linear methods have performed very well in image recognition tasks, but in recent times, orthogonal variants of these DR methods are proposed in OLPP [13] and ONPP [14] have been proved superior to their non-orthogonal counterparts and considered state-of-the-art in appearance based image recognition. Over the time, various discriminating and/or 2-dimensional variants of OLPP and ONPP are proposed and used in the image recognition tasks, too.

In this article, ONPP technique with class similarity based neighborhood is used to perform expression recognition and the performance is compared with conventional ONPP approach [14] and Modified ONPP approach [15]. The article also compares the performance of DR based expression recognition approach with local feature descriptor based state-of-the-art approaches.

3 Expression Recognition Using ONPP and Class Similarity Based Neighborhood

3.1 Data Preparation for Modular Expression Recognition

Appearance based DR methods consider an $m \times n$ image as a data point in mn -dimensional space where each image is vectorized either by columns or by rows. Let $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N$ be the N data points of given training images, so the data matrix \mathbf{X} can be defined as $\mathbf{X} = [\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N] \in \mathbf{R}^{mn \times N}$. As discussed in Sect. 2 it is expected that the similar images will be projected in the same neighborhood in the projected space. It is proven in that area between eyebrows, eyes, area containing nose and lips plays major role in expressing emotions while the rest of the facial area does not provide any significant information for expression recognition. The modular approach considers only these areas while recognizing expressions, thus the each data vector \mathbf{x}_i corresponding to image i is prepared considering above mentioned four areas from the image. Each significant area is cropped from the whole image and vectorized, these four vectors of facial region are then concatenated to make l -dimensional vector \mathbf{x}_i as shown in Fig. 1. Note that the size of these regions across all face images should be same so that the resulting vector representation be a point in \mathcal{R}^l .

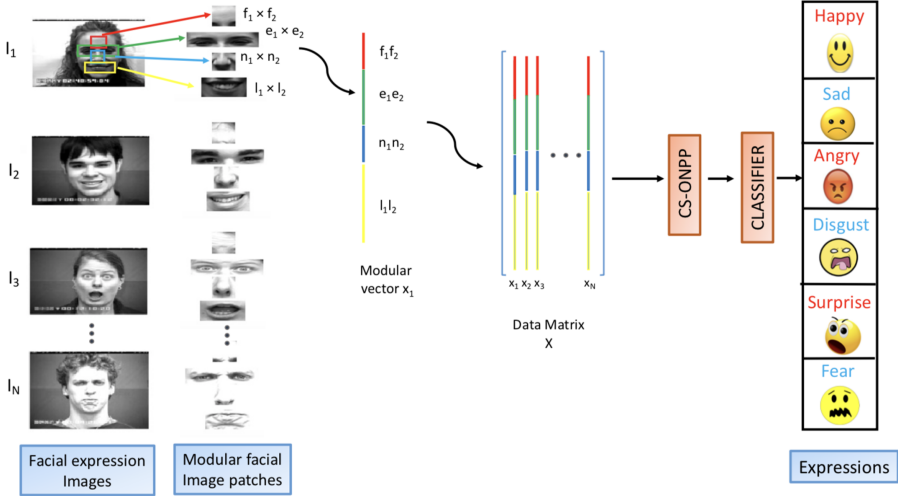


Fig. 1. Conversion of Facial expression image data into modular vectorized data: each l -dimensional vector $\mathbf{x}_i \in \mathbf{X}$ represents face image I_i

3.2 ONPP with Class Similarity Based Neighborhood

Let X be the data matrix of training images such that $\mathbf{X} = [\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N] \in \mathbf{R}^{l \times N}$ for N modular vector representing face images. ONPP is a three step procedure where in the first step defines the neighborhood of each data point \mathbf{x}_i , the next step expresses each data point as a linear combination of its neighbors \mathbf{x}_j s. The last step finds projection bases $\mathbf{V} \in \mathcal{R}^{l \times d}$ of an ONPP space using training data \mathbf{X} that can be used to embed test image \mathbf{x}_t in the ONPP space for recognition step.

Step 1: Finding Nearest Neighbors: For every data point \mathbf{x}_i , neighborhood \mathcal{N}_{x_i} is defined. The choice of neighborhood can be defined in several ways, the simple choice can be made using k neighbors are chosen by Nearest Neighbor (NN) method where k is a reasonably picked parameter. In another way, neighbors could be chosen which are inside ε radius from the data point. Let \mathcal{N}_{x_i} be the set of k neighbors of \mathbf{x}_i . Both these approaches do not consider the available knowledge of class labels, thus known as unsupervised approach. In an alternative approach - supervised approach - data points belonging to same class are considered neighbors of each other.

Both the methods have their own shortcomings. Unsupervised methods consider only euclidean distance which essentially describes similarity between data points regardless of discriminating information based on class labels. Supervised methods decide the neighborhood neglecting the similarity between data points those are not in the same class. Both of these approaches are suitable for face recognition application where the difference between two images stems only from occlusions, illumination, expression and pose variations. On the other hand, in

expression recognition task, images of an expression coming from two person show large variation which cannot be represented by only euclidean distance or only label knowledge.

Instead of claiming a data point \mathbf{x}_i belonging to an unique class, it makes much more sense that a data point belongs to a class c_i with a certain probability p_{c_i} . If the data set has C classes, then for each data point we can build a C -dimensional probability vector $\mathbf{p}(\mathbf{x}_i) = [p_1(\mathbf{x}_i), p_2(\mathbf{x}_i), \dots, p_c(\mathbf{x}_i)]^T$. he probability vector is computed based on class similarity between data points which combines the knowledge of class label as well as euclidean distance.

Let $\Delta(i, j)$ be the euclidean distance between data points \mathbf{x}_i and \mathbf{x}_j . Class similarity based distance denoted by $\Delta'(i, j)$ can be defined by

$$\Delta'(\mathbf{x}_i, \mathbf{x}_j) = \Delta(i, j) + \alpha \max(\Delta)(1 - \mathcal{S}(i, j)) \tag{1}$$

where, $\alpha \in [0, 1]$ is a tuning parameter. $\max(\Delta)$ indicates maximum pair-wise distance or data diameter. $\mathcal{S}(i, j)$ is class similarity between \mathbf{x}_i and \mathbf{x}_j , which is defined as,

$$\mathcal{S}(i, j) = \begin{cases} 1; & \mathbf{x}_i = \mathbf{x}_j \\ \mathbf{p}(\mathbf{x}_i)^T \mathbf{p}(\mathbf{x}_j); & \mathbf{x}_i \neq \mathbf{x}_j \end{cases} \tag{2}$$

We used Logistic Discrimination (LD) to find probability of each data point \mathbf{x}_i belonging to class c_i . Performing LD on high dimensional data causes huge computational burden, thus lower dimensional representation is sought using PCA. Let \mathbf{z}_i be a lower dimensional representation of \mathbf{x}_i , to find probability vector $\mathbf{p}(\mathbf{x}_i)$. The c^{th} element of $\mathbf{p}(\mathbf{x}_i)$ corresponding to class c can be computed by

$$p_c(\mathbf{x}_i) = \frac{\pi(\mathbf{z}_i; \alpha_c, \beta_c)}{\sum_{c=1}^C \pi(\mathbf{z}_i; \alpha_c, \beta_c)} \tag{3}$$

where the function, $\pi(\mathbf{z}_i; \alpha_c, \beta_c) = \frac{\exp(\alpha_c + \beta_c^T \mathbf{z}_i)}{1 + \exp(\alpha_c + \beta_c^T \mathbf{z}_i)}$

The neighbors for each data point \mathbf{x}_i will be chosen based on class similarity based distance given in Eq. (1).

Step 2: Calculating Reconstruction Weight: In this step, the neighborhood \mathcal{N}_{x_i} is expressed as a linear combination of neighbors with reconstruction weight w_{ij} s as $\sum_{j=1}^k w_{ij} \mathbf{x}_j$. The weight w_{ij} are computed by minimizing the reconstruction error i.e. error between \mathbf{x}_i and linear combination of $\mathbf{x}_j \in \mathcal{N}_{x_i}$.

$$\arg \min_{\mathbf{W}} \mathcal{E}(W) = \arg \min_{\mathbf{W}} \sum_{i=1}^N \left\| \mathbf{x}_i - \sum_{j=1}^k w_{ij} \mathbf{x}_j \right\|^2 \tag{4}$$

subject to $\sum_{j=1}^k w_{ij} = 1$.

For each data point \mathbf{x}_i , optimization problem given in (4) can be modeled as a least square problem $(\mathbf{X}_{N_i} - \mathbf{x}_i \mathbf{e}^T) \mathbf{w}_i = \mathbf{0}$ with a constraint $\mathbf{e}^T \mathbf{w}_i = 1$. Here, \mathbf{X}_{N_i} is a matrix having \mathbf{x}_j as its columns, where $\mathbf{x}_j \in N_{x_i}$. Note that \mathbf{X}_{N_i} includes \mathbf{x}_i as its own neighbor making it a matrix of dimension $l \times k + 1$. Solving the least square problem results in a closed form solution for \mathbf{w}_i given by Eq. (5).

$$\mathbf{w}_i = \frac{\mathbf{G}^{-1} \mathbf{e}}{\mathbf{e}^T \mathbf{G}^{-1} \mathbf{e}} \tag{5}$$

Here, \mathbf{e} is a vector of ones having dimension $k \times 1$ same as \mathbf{w}_i . $\mathbf{G} \in \mathcal{R}^{k \times k}$ is a Gramian matrix, each entry of \mathbf{G} is given by $\mathbf{g}_{\mathbf{p}\mathbf{l}} = (\mathbf{x}_i - \mathbf{x}_{\mathbf{p}})^T (\mathbf{x}_i - \mathbf{x}_{\mathbf{l}})$, for $\forall \mathbf{x}_{\mathbf{p}}, \mathbf{x}_{\mathbf{l}} \in N_{x_i}$.

Step 3: Finding Projection Matrix: Last step is dimensionality reduction or finding the projection matrix \mathbf{V} that explicitly maps l -dimensional data point \mathbf{x}_i to d -dimensional representation \mathbf{y}_i assuming that the neighborhood relationship among N_{x_i} with corresponding weights w_{ij} will be preserved in lower dimensional space, too.

The optimization problem to achieve such mapping can be formed as minimization of the sum of squares of reconstruction errors in lower dimensional space. The cost function is given by

$$\arg \min_{\mathbf{Y}} \mathcal{F}(\mathbf{Y}) = \arg \min_{\mathbf{Y}} \sum_{i=1}^N \left\| \mathbf{y}_i - \sum_{j=1}^k w_{ij} \mathbf{y}_j \right\|^2 \tag{6}$$

subject to orthogonality constraint, $\mathbf{V}^T \mathbf{V} = \mathbf{I}$.

Solving the optimization problem results in eigenvalue problem $\mathbf{X} \mathbf{M} \mathbf{X}^T \mathbf{V} = \lambda \mathbf{V}$. Here, columns of \mathbf{V} are eigen-vectors that corresponding to the smallest d eigen-values. The matrix $\mathbf{M} = (\mathbf{I} - \mathbf{W})(\mathbf{I} - \mathbf{W}^T)$. Note that $\mathbf{X} \mathbf{M} \mathbf{X}^T$ is symmetric and positive semi-definite. ONPP explicitly maps \mathbf{X} to \mathbf{Y} , which is of the form $\mathbf{Y} = \mathbf{V}^T \mathbf{X}$, i.e. each test sample \mathbf{x}_t can now be projected to lower dimension by just a matrix-vector product $\mathbf{y}_t = \mathbf{V}^T \mathbf{x}_t$.

Considering the under-sampled size issue where the number of samples N is less than dimension l . In such situation, the matrix $\mathbf{X} \mathbf{M} \mathbf{X}^T \in \mathbf{R}^{l \times l}$ will have maximum rank $N - c$, where c is number of classes. To ensure that the resulting matrix \mathbf{M} be non-singular, one may utilize an initial PCA projection that reduces the dimensionality of the data vectors to $N - c$. If \mathbf{V}_{PCA} is the projection matrix of PCA, then on performing the ONPP the resulting dimensionality reduction matrix is given by $\mathbf{V} = \mathbf{V}_{\text{PCA}} \mathbf{V}_{\text{ONPP}}$. Note that the PCA projection is most common pre-processing applied in many dimensionality reduction methods and in this article we are using it in our advantage to define new distance measure for local neighborhood.

Following Table 1 gives procedure to find Class-similarity based ONPP subspace and mechanism to recognize expression of the test image.

Table 1. Procedure for expression recognition using Class-Similarity based ONPP

Input: Dataset $\mathbf{X} \in R^{l \times N}$ representing N facial expression images in modular format and number of reduced dimension d
Output: Lower dimension representation $\mathbf{Y} \in R^{d \times N}$

- 1: Find low dimensional representation \mathbf{z}_i of data by projecting on d_{pca} dimensional space using PCA ($\mathbf{z}_i = V_{pca}^T \mathbf{x}_i$)
- 2: Use Logistic Regression on \mathbf{z}_i to find class probability vector \mathbf{p}_i
- 3: Calculate modified distance for all data point pairs $\Delta'(\mathbf{x}_i, \mathbf{x}_j)$ using equation (1)
- 2: Compute NN $\mathcal{N}_{\mathbf{x}_i}$ with modified distance $\Delta'(\mathbf{x}_i, \mathbf{x}_j)$
- 3: Compute the weight W for each neighbor data point $\mathbf{x}_j \in \mathcal{N}_{\mathbf{x}_i}$ as given in equation (5)
- 4: Compute Projection matrix $V \in R^{l \times d}$ whose column vectors are smallest d eigenvectors of matrix \mathbf{XMX}^T
- 5: Compute Embedding on lower dimension by $\mathbf{Y} = \mathbf{V}^T \mathbf{X}$
- 6: Project test facial expression image represented as modular vector \mathbf{x}_t on learned ONPP space to get low dimensional representation \mathbf{y}_t
- 7: Use 1-NN classifier to identify the class label for test image

4 Experiments

To approve the hypothetical conclusion of the proposed framework, experiments were performed on the four facial datasets. 1) JAFFE database [16] having 213 facial images of 10 Japanese female models of 7 facial expressions (6 basic facial expressions + 1 neutral). Out of 213 images, random 140 images were chosen for the training and the remaining 73 were used for testing. 2) The Video database [17] has videos of 11 persons. The single video contains four different facial expressions: Smiling, Angry, Open mouth and Normal. Out of 6668 images, randomly 70% images were chosen for training and remaining 30% images used as testing 3) In CK+ [18] there are 593 sequences across 123 persons giving 8 facial expressions. Table 2 reports average and best recognition result of 20 such iterations.

Comparison of performance of ONPP [14] and CS-ONPP on facial expression databases in the light of recognition score (in %) with corresponding subspace dimensions are reported in Table 3. Where as Table 4 reported the comparison between the MONPP [15] and CS-ONPP. CS-ONPP are reported along with tuning parameter Alpha and PCA dimension (d_pca).

Table 5 compares recognition results of the proposed technique with that of few State-of-the-art Local feature based facial expression recognition techniques. As can be seen from the comparison that for JAFFE database, similar performance is achieved with proposed method but with feature vector of only 110 dimensions. For CK+ data, best recognition is achieved at 510 dimensions,

Table 2. Best recognition Accuracy (%) achieved with proposed method of three benchmark databases along with related parameters: PCA subspace dimension (dpca), Number of Nearest Neighbors (k), tuning parameter α and ONPP subspace dimensions

Database	Recognition accuracy (%)	ONPP dimensions	dpca	alpha α	Number of nearest neighbors (k)
JAFFE	94.54	100	24	0.25	7
Video	94.76	110	20	0.50	13
CK+	86.76	510	22	0.25	5

Table 3. Comparison of performance of ONPP and CS-ONPP on facial expression databases in the light of recognition score (in %) with corresponding subspace dimensions. CS-ONPP are reported along with tuning parameter Alpha and PCA dimension (d_pca).

Databases	ONPP		CS-ONPP				
	RecAcc	Subspace dim	RecAcc	Subspace dim	d_pca	Alpha	NN_k
JAFFE	93.62	155	94.54	100	24	0.25	7
VIDEO	94.12	175	94.76	110	20	0.5	13
CK+	85.66	705	86.76	510	22	0.25	5

Table 4. Comparison of performance of MONPP and CS-ONPP on facial expression databases in the light of recognition score (in %) with corresponding subspace dimensions. CS-ONPP are reported along with tuning parameter Alpha and PCA dimension (d_pca).

Databases	MONPP		CS-ONPP				
	RecAcc	Subspace dim	RecAcc	Subspace dim	d_pca	Alpha	NN_k
JAFFE	93.76	145	95.35	85	20	0.25	9
VIDEO	94.4	170	95.66	105	20	0.5	14
CK+	87.14	675	87.88	485	18	0.5	9

which is not superior to other methods, but provides an advantage in terms of reduced feature vectors. Table 6 repeats the comparisons given in [23] for Local features based Facial Expression Recognition methods along with the dimensions of feature vectors for the given method.

Table 5. Comparison with Recognition Accuracy reported in some State-of-the-Art facial expression methods

DataBase	Existing methods	Classifiers	Accuracy
JAFFE	2D Geometry & Gabor Wavelet	NN with PROP	90.00%
	Gabor Filter [19]	ANN	88.75%
	HSOG [10]	1-NN	94.15%
	EPCA [20]	1-NN	89.01%
	DBN	Adaboost	91.8%
	WMDNN[21]	–	92.89%
	Proposed: ONPP using CS	1-NN	94.54%
CK+	DCV [22]	HMM	88.00%
	Gabor filter [19]	ANN	88.94%
	HSOG [10]	1-NN	92.54%
	CNN+loss layer	–	90.66%
	CNN, RBM	SVM	92.05%
	EPCA [20]	1-NN	91.72%
	Proposed: ONPP using CS	1-NN	86.76%

Table 6. Comparison of proposed method with Local feature based methods in the light of Feature vector length for JAFFE dataset for 256×256

Holistic approaches		
Method	Feature length	Recognition accuracy (%)
Local Binary Pattern (LBP)	65536	89.42
Local Gradient Code (LGC)	65536	90.38
Histogram of Gradients (HOG)	20736	85.71
Local Directional Pattern (LDP)	14337	85.20
Modular approaches		
Method	Feature length	Recognition accuracy (%)
Histogram of 2^{nd} Order Gradient (HSOG)	38582	94.15
Proposed ONPP with CS	11541	94.56

5 Conclusion

Though, Local Features based methods have been successfully applied to Facial Expression Recognition problems, the resulting feature vector lengths usually are of order 10^5 which slow down classification process. The article proposes a Dimensionality Reduction based method which can be employed in FER.

Basically, state-of-the-art DR method ONPP is used and a novel approach of neighborhood selection based on class similarity is proposed to suit FER application. Proposed method is tested on three benchmark databases and proved to be gaining huge margin in terms of feature vector length while maintaining same recognition accuracy.

References

1. Sandbach, G., Zafeiriou, S., Pantic, M., Yin, L.: Static and dynamic 3D facial expression recognition: a comprehensive survey. *Image Vis. Comput.* **30**(10), 683–697 (2012)
2. Pantic, M., Rothkrantz, L.J.: Automatic analysis of facial expressions: the state of the art. *IEEE Trans. Pattern Anal. Mach. Intell.* **22**(12), 1424–1445 (2000)
3. Lowe, D.G.: Distinctive image features from scale-invariant keypoints. *Int. J. Comput. Vis.* **60**(2), 91–110 (2004)
4. Bay, H., Tuytelaars, T., Van Gool, L.: SURF: speeded up robust features. In: Leonardis, A., Bischof, H., Pinz, A. (eds.) *ECCV 2006*. LNCS, vol. 3951, pp. 404–417. Springer, Heidelberg (2006). https://doi.org/10.1007/11744023_32
5. Dalal, N., Triggs, B.: Histograms of oriented gradients for human detection. In: 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, CVPR 2005, vol. 1, pp. 886–893. IEEE (2005)
6. Zhang, B., Shan, S., Chen, X., Gao, W.: Histogram of gabor phase patterns (HGPP): a novel object representation approach for face recognition. *IEEE Trans. Image Process.* **16**(1), 57–68 (2007)
7. Ojala, T., Pietikainen, M., Maenpaa, T.: Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. *IEEE Trans. Pattern Anal. Mach. Intell.* **24**(7), 971–987 (2002)
8. Ahonen, T., Hadid, A., Pietikainen, M.: Face recognition with local binary patterns. In: Pajdla, T., Matas, J. (eds.) *ECCV 2004*. LNCS, vol. 3021, pp. 469–481. Springer, Heidelberg (2004). https://doi.org/10.1007/978-3-540-24670-1_36
9. Zhu, C., Bichot, C.-E., Chen, L.: Multi-scale color local binary patterns for visual object classes recognition. In: 2010 20th International Conference on Pattern Recognition (ICPR), pp. 3065–3068. IEEE (2010)
10. Sujata, Mitra, S.K.: A modular approach for facial expression recognition using HSOG. In: *Proceeding of 9th International Conference on Advances in Pattern Recognition (ICAPR 2017)*. IEEE (2017)
11. Roweis, S.T., Saul, L.K.: Nonlinear dimensionality reduction by locally linear embedding. *Science* **290**(5500), 2323–2326 (2000)
12. He, X., Niyogi, P.: Locality preserving projections. In *Advances in Neural Information Processing Systems*, pp. 153–160 (2004)
13. Cai, D., He, X., Han, J., Zhang, H.-J.: Orthogonal Laplacian faces for face recognition. *IEEE Trans. Image Process.* **15**(11), 3608–3614 (2006)
14. Kokiopoulou, E., Saad, Y.: Orthogonal neighborhood preserving projections: a projection-based dimensionality reduction technique. *IEEE Trans. Pattern Anal. Mach. Intell.* **29**(12), 2143–2156 (2007)
15. Koringa, P., Shikkenawis, G., Mitra, S.K., Parulkar, S.K.: Modified orthogonal neighborhood preserving projection for face recognition. In: Kryszkiewicz, M., Bandyopadhyay, S., Rybinski, H., Pal, S.K. (eds.) *PREMI 2015*. LNCS, vol. 9124, pp. 225–235. Springer, Cham (2015). https://doi.org/10.1007/978-3-319-19941-2_22

16. Lyons, M., Akamatsu, S., Kamachi, M., Gyoba, J.: Coding facial expressions with Gabor wavelets. In: 1998 Proceedings of Third IEEE International Conference on Automatic Face and Gesture Recognition, pp. 200–205. IEEE (1998)
17. Shikkenawis, G., Mitra, S.K.: On some variants of locality preserving projection. *Neurocomputing* **173**, 196–211 (2016)
18. Kanade, T., Cohn F.J., and Tian Y.: Comprehensive database for facial expression analysis. In *Automatic Face and Gesture Recognition, 2000. Proceedings. Fourth IEEE International Conference on*, pages 46–53. IEEE, 2000
19. Saabni, R.: Facial expression recognition using multi radial bases function networks and 2-D Gabor filters. In: 2015 Fifth International Conference on Digital Information Processing and Communications (ICDIPC), pp. 225–230. IEEE (2015)
20. Trivedi, M., Mitra, S.K., et al.: A modular approach for facial expression recognition using Euler principal component analysis (E-PCA). In: 2018 IEEE Applied Signal Processing Conference (ASPCON), pp. 204–208. IEEE (2018)
21. Yang, B., Cao, J., Ni, R., Zhang, Y.: Facial expression recognition using weighted mixture deep neural network based on double-channel facial images. *IEEE Access* **6**, 4630–4640 (2018)
22. Wang, Y.-K., Huang, C.-H.: Facial expression recognition with discriminative common vector. In: 2007 Third International Conference on Intelligent Information Hiding and Multimedia Signal Processing, IHHMSP 2007, vol. 2, pp. 431–434. IEEE (2007)
23. Kumari, J., Rajesh, R., Pooja, K.M.: Facial expression recognition: a survey. *Procedia Comput. Sci.* **58**, 486–491 (2015)