

HOG & Sparse Based Face Recognition System

Shahina¹, Arun P.S²

¹Dept of Electronics & Communication Engineering Name of organization - MBCET
Trivandrum, Kerala, India

²Dept of Electronics & Communication Engineering Name of organization - MBCET
Trivandrum, Kerala, India

Abstract: Face detection and recognition is one of the most challenging problems in the field of image processing. One of the recent techniques in face recognition is by the sparse representation. In this paper, combination of Sparse and Histogram of Oriented Gradients (HOG) are used. In the first stage dimensionality reduction is done to reduce the amount of random variables and the features are extracted using Principal Component Analysis (PCA). The extracted features are classified using Sparse Representation via l_1 -minimization. For better invariance to illumination and expression changes, Histogram of Oriented Gradients (HOG) feature is also extracted to represent face images. This method offers better recognition performance when compared to the existing methods. The proposed face recognition method is evaluated on the standard ORL data base and yields very impressive results.

Index Terms: Face Recognition, Sparse Representation, Histogram of Oriented Gradient, Principal Component Analysis

I. Introduction

Face recognition remains one of the most important and challenging research topic in computer vision and many other practical application domains [1]. Many applications based on face recognition are access control, video surveillance, human computer interfaces etc. A variety of techniques have been adopted for Automatic face recognition, but still it is a difficult problem. This difficulty is mainly due to artifacts that obscure the image features, like the variations in illumination, facial expression, and head pose. Techniques that may offer efficient feature extraction with a high discrimination power and low computational complexity are crucial [3].

There are different approaches available for face Recognition like Holistic approach, Feature-based approach and Hybrid approach. In Holistic approach, the entire facial images are taken for the face recognition. In Feature-based approach local features like nose, eyes are segmented and it can be used as input data for the face recognition. The Hybrid approach is the combination of holistic and feature based approach. In this approach both the local and whole face is used for the face recognition. Some of the holistic based approaches are Principal Component Analysis (PCA), Linear Discriminant Analysis (LDA), Independent Component Analysis (ICA) etc... and some of the Feature based approaches are Linear Binary Pattern (LBP), Scale Invariant Feature Transform (SIFT), Histogram of Oriented Gradient (HOG) etc...

Standard sparsity concepts mainly focus on the sparsity of coefficient vector while other important characteristics are less considered. Wright et al. [6] proposed a standard sparse representation based classification (SRC) method for face recognition recently. In their work, each test sample is casted as a sparse linear combination of all training samples. This method outperforms many existing methods as it can handle disguises and variations to some extent. Sparse representation based classification (SRC) has been demonstrated to be superior to nearest neighbor (NN) and nearest subspace (NS) based classifiers in various subspaces (e.g. PCA or LDA). When applied to face recognition, it can also be efficiently customized to handle errors due to occlusion and corruption. Following Wright et al.'s work, in the past year, several extensions of SRC based face recognition were proposed. In [8], Zhou et al. applied a Markov Random Field model to SRC based face recognition for improving performances under severe contiguous occlusion. Yang and Zhang [9] used image Gabor-features for SRC in order to reduce the cost in coding occluded faces meanwhile improving accuracy. In [10], Yang et al. reviewed five representative l_1 -minimization methods in the context of SRC based face recognition.

However, there are still some defects. For one thing, if face image is contaminated by sparse occlusions, an occlusion dictionary is introduced to code the occlusion components. However, the occlusion dictionary used in SRC method is a high dimensional orthogonal matrix which makes the standard sparse coding computationally complex. In this paper, in order to increase the accuracy and to

maintain a better invariance to illuminant changes HOG features are also extracted and is combined with the sparse coefficient vectors to represent the facial images .

II. Methodology

In this approach, first the database is created by extracting image features and storing them in a data base. For any input facial images its features will be found to match them with the database. Input will be a matrix of training sample and the output will be identified image which is present in the data base.

2.1 HOG Feature Based Face Representation

The algorithm for extracting the HOGs counts occurrences of the edge orientations in a local neighborhoods of an image. Here, the image is first divided into small connected regions, called cells, and for each cell a histogram of edge orientations are computed. The histogram channels are evenly spread in the regions 0-180° or 0-360°, depending on whether the gradients obtained is 'unsigned' or 'signed'. The histogram counts are then normalized to compensate for the illumination. This is done by accumulating a measure of the local histogram energy over the somewhat larger connected regions and using those results to normalize all the cells in the block. The combination of these histograms will represents the final HOG descriptor. Invariance to the scale and rotation can also be achieved by the extracting descriptors from only salient points (key-points) in the scale space of the image following a rotation normalization[11]. The steps involves are:

- (1) Scale-space extrema detection.
- (2) Orientation assignment.
- (3) Descriptor extraction.

The first step is intended to achieve the scale invariance. The second step is used to find the dominant gradient orientation. All the orientations counts are then made relative to the dominant directions. Fig. 1 shows an example of the patch with their corresponding HOGs. In order to use the HOG features for the face recognition five regions are extracted from facial images which are eyes, nose, lips, and chin regions. And also, the size of the patch used to extract the HOG features is important[11].

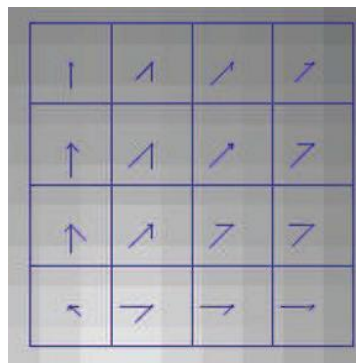


Fig 1: Example HOG descriptors, patch size=8 x 8. Each cell of the patch shows the orientation of the gradients[17] .

2.2 Sparse representation based face classification

Supposing a training set A consists of the facial images from k classes, where $A = \{A_1, A_2, \dots, A_k\}$. Ideally, giving sufficient training samples of class i , where $A_i = \{v_{i,1}, v_{i,2}, \dots, v_{i,ni}\} \in R^{m \times ni}$, a test facial image $y \in R^m$ belongs to the same class could be well approximated by a linear combination of the training samples from A_i , which can be written as:

$$y = \sum_{j=1}^{ni} a_{i,j} v_{i,j} \tag{1}$$

Since A is the dictionary which contains all the training samples, where $A = \{v_{1,1}, v_{1,2}, \dots, v_{k,ni}\}$. Then equation (1) can be written in the form as below :

$$y = Ax_0 \in R^m \tag{2}$$

where $x_0 = \{0, \dots, 0, a_{i,1}, a_{i,2}, \dots, a_{i,ni}, 0, \dots, 0\}^T$ is the coefficient vector in which most coefficients are zero except the ones associated with class i

Due to the fact that a valid test sample y can be sufficiently represented only using the training samples from the same class, and this representation is the sparsest among all others, to find the identity of y then equals to find the sparsest solution of (2). This is the same as solving the following optimization problem

(ℓ_0 -minimization) However, solving the ℓ_0 -minimization of an underdetermined system of linear equations is NP-hard. In the case for large number of training samples, it equals to find the minimal ℓ_1 -norm solution [12]. Therefore, the SRC procedure presented in [7] is shown as below.

Algorithm 1. The SRC Algorithm

1. Normalize the columns of A to have unit ℓ_2 norm
2. Solve the ℓ_1 -minimization problem.

$$x_1^{\wedge} = \arg \min \|x\|_1 \text{ subject to } \|Ax - y\|_2 \leq \epsilon \quad (3)$$

3. Compute the residuals by:

$$r_1(y) = \|y - A\Gamma_i(x_1^{\wedge})\|_2 \quad (4)$$

For $i=1,..,k$, where Γ_i is the characteristic function which selects the coefficients associated with the i -th class.

4. Output the identity by:

$$\text{identify}(y) = \arg \min_1 r_1(y) \quad (5)$$

2.3 Architecture

Figure 2 shows the architecture of the proposed method. In this method the HOG features are combined with SRC for face recognition. First the input image is divided into several blocks (2 x 2, 4 x 4, 8 x 8) and their facial features like left eye, right eye, nose are extracted. Then the magnitude and direction of orientations are computed from the following equation (6) and (7) and its histogram are plotted. Calculation of gradient over each pixel (x,y) is done by using 1-D filter, $h_1 = [-1, 0, 1]$ and $h_2 = h_1^T$

$$g(x, y) = \sqrt{g_x^2 + g_y^2} \quad (6)$$

and the gradient direction is given by,

$$\theta(x, y) = \tan^{-1}\left(\frac{g_y}{g_x}\right) \quad (7)$$

Normalisation is usually done after calculating the histogram vectors, for the better invariance to illumination and expression. Four different normalization schemes are L2-norm, L2-Hys, L1-sqrt, and L1-norm. Here the L2-norm scheme due to its better performance [11]. Dimension reduction is carried out by PCA. HOG features are taken after dimension reduction as a sample to represent each face image in this paper. The original Sparse Representation Classification (SRC) will take the entire feature space extracted from the whole image as the dictionary component. In this approach HOG features are also included into the dictionary component from one sub-block. In this way, the SCVs are computed for all sub-blocks of the input image using ℓ_1 -minimization (equation 4). The SCV's within all the division level are combined to get an enhanced SCV (E-SCV) by elementary wise summation. Then the E-SCVs are normalized to have the unit ℓ_2 -norm. At last all the normalized E-SCVs from all the division levels are being summed to form the augmented SCV (A-SCV) for the classification stage.

2.4 Dimensionality Reduction

The training set has a common resolution of $r \times c$, and every image is represented in a column vector of $r \times c$ by 1, and all the images in the training set represent one column of array \mathbf{T} . Now the dimensionality reduction is done using Principal Component Analysis (PCA). The eigen vectors are computed from the covariance matrix of \mathbf{T} and the eigen faces are computed from the covariance matrix \mathbf{T} for computing the Eigen faces. The main advantage of feature extraction is the reduction in data dimension and computational cost. Along with these features the HOG features are also extracted. Extracted features are then ℓ_1 -minimized and sparse representation will be created.

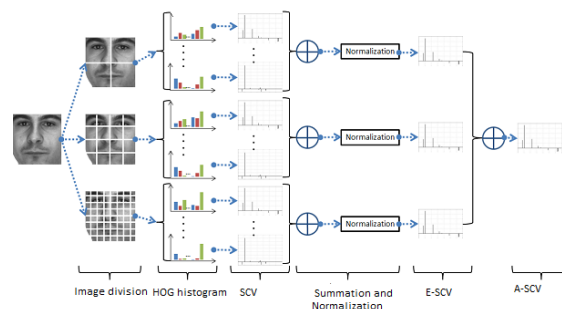


Fig 2: Architecture of the proposed method

For the input image first it is normalized for the basis and then features are extracted and the extracted features are compared with the values in the data base and the match is found appropriately. Along with the PCA features another important features like left eye ,right eye, nose etc. are also used.

III. Experimental Results

The simulation of this method was done on MATLAB environment for all the steps like normalization, feature extraction, training and testing. The data base used in this method is the standard ORL data base. The ORL face database contain ten different images of each of the 40 persons and a total of 400 images are present. The images used in this data base are in BMP format. The size of each image is 112 x 92 pixels with 256 grey levels per pixel. The data set is initially divided into two parts-one for training and other for testing, In this paper, three image per person of different lightning conditions are used for training stage and the remaining images are used for testing stage. The figure 3 shows the sample images in the data base.

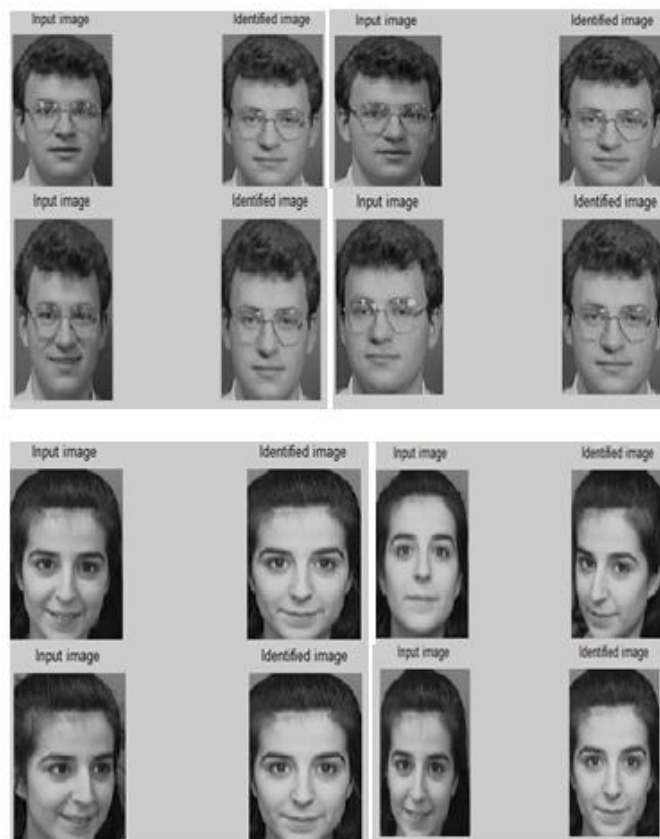


Fig 3: sample image in the data base

The data base used is complex since it consists of different facial expression and illumination conditions. The increased number of identities as well as the reduced number of training samples increases the complexity in face recognition. Comparing with other existing face recognition techniques, HOG based sparse representation has significant improvement in the recognition rates.

For comparison, we also compared the performance of using holistic feature based sparse algorithm. Sparse HOG, can handle 50% occlusion using only one gallery image. While PCA based results work very poorly even though there is no occlusion exists. This is may be due to the session reasons, when the test image and the gallery image are from different sessions, some variations such as hairstyle, pose, illumination, can affect the accuracy of alignment which is directly affect the holistic feature representation. This again states that our proposed Sparse-HOG with PCA is robust with different scenarios.

Some of the test image with different expressions and its identified image from the training set in the ORL data base is shown in figure 4



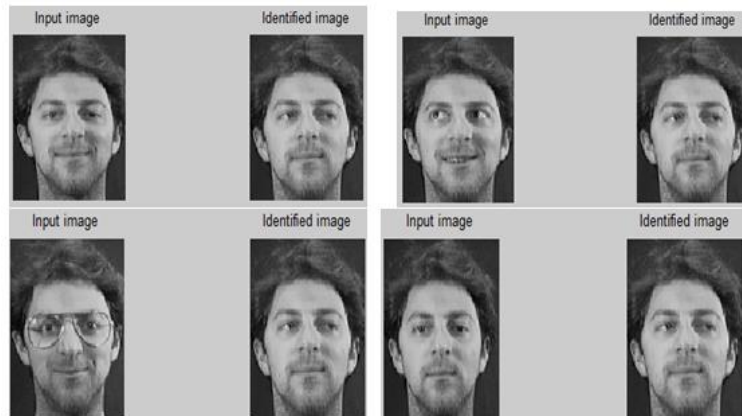


Fig5: Some of the Test image and its identified image by the proposed method

Table 1 shows the Recognition rates of the holistic feature based sparse representation and HOG feature based sparse representation

TABLE 1: Recognition Results on ORL data base

RECOGNITION RATES	
Algorithm	ORL Data Base
PCA+Sparse	70.2%
HOG+PCA+Sparse	82.5%

IV. Conclusion

The HOG features are widely used in applications like pedestrian detection and tracking [11], but has rarely been used in face recognition. In this paper, we extended the holistic based sparse representation by incorporating the HOG features. The proposed methods performs superior to the original Sparse, HOG, based methods and also several other existing face recognition algorithms in terms of the recognition accuracy.

References

- [1]. W.Zhao,R. Chellappa, P. J.Phillips, and A. Rosenfeld,“Face Recognition: A Literature Survey,” ACM Computing Surveys, 35(4), pp. 399-458, 2003.
- [2]. Jayavel J and Shahnazeer C K “Sparse Representation for Face Recognition” IJCSNS International Journal of Computer Science and Network Security, VOL.14 No.7, July 2014
- [3]. A. Yang, J. Wright, Y. Ma, and S. Sastry, “Feature selection in face recognition: A sparse representation perspective,” UC Berkeley Tech Report UCB/EECS-2007-99, 2007
- [4]. E. J. Candes, X. D. Li, Y. Ma, and J. Wright, “Robust principal component analysis?” J. ACM, vol. 58, no. 1, pp. 1–37, 2011.
- [5]. A. Yang, A. Ganesh, Z. Zhou, S. Sastry, and Y. Ma. Fast l_1 -minimization algorithms and an application in robust facerecognition: a review. Technical Report UCB/EECS-2010-13,UC Berkeley, 2010.
- [6]. J. Chen, S.Shan, C. He, G. Zhao, X.Chen, and W. Gao, “WLD: A Robust Local Image Descriptor”, IEEE Transactions on Pattern
- [7]. J. Wright, A. Ganesh, A. Yang, and Y. Ma, “Robust face recognition via sparse representation,” IEEE Transactions on Pattern Analysis and Machine Intelligence, vol.31, no.2, pp. 210-227, 2009.
- [8]. Z. Zhou, A. Wagner, H. Mobahi, J. Wright, and Y. Ma, “Face Recognition with Contiguous Occlusion using Markov Random Fields,” IEEE International Conference on Computer Vision (ICCV), 2009
- [9]. M. Yang and L. Zhang, “Gabor Feature based Sparse Representation for Face Recognition with Gabor Occlusion Dictionary,” 11th European Conference on Computer Vision (ECCV 2010), Sept. 2010.
- [10]. A. Yang, A. Ganesh, S. Sastry and Y. Ma, “Fast l_1 - Minimization Algorithms and an Application in Robust Face Recognition: A Review,” In Proc. of IEEE 17th International Conference on Image Processing (ICIP 2010), Sept. 2010.
- [11]. N. Dalal and B. Triggs, “Histograms of oriented gradients for human detection,” in CVPR, 886-893, 2005.
- [12]. D.L. Donoho, “For most large underdetermined systems of linear equations the minimal l_1 -norm solution is also the sparsest solution,” Comm. Pure and Applied Math., vol. 59, no. 6, pp. 797–829, 2006
- [13]. G. Tzimiropoulos, S. Zafeiriou, and M. Pantic, “Sparse representations of image gradient orientations for visual recognition and tracking,” in Proc. IEEE Comput. Soc. Conf. CVPRW, 2011, pp. 26–33.
- [14]. A.M. Martinez and R. Benavente, “The ORL database,” Technical report, CVC Technical report, no. 24, 1998.
- [15]. [15]P. Belhumeur, J.Hespanha, and D. Kriegman,“Eigenfaces vs. Fisher-faces: Recognition using class specific linear projection,” IEEE Trans.Pattern Anal. Mach. Intell., vol. 19, no. 7, pp. 711–720, Jul. 1997
- [16]. E. J. Candes, X. D. Li, Y. Ma, and J. Wright, “Robust principal component analysis?” J. ACM, vol. 58, no. 1, pp. 1–37, 2011.
- [17]. D. Lowe, “Distinctive image features from scale invariant keypoints”, IJCV, 60(2), 2004, pp. 91–110.