Towards Expression invariant face recognition CS6350-Term Project Report, Group-2

Vinay Babu J - CS12B012 Abhishek yadav - CS12B032 Vamsi Krishna - CS10B055

***Abstract*— Face recognition and detection in various illumination conditions, poses and expressions is a challenging task. Information jointly contained in image space, scale and orientation domains can provide rich important clues not seen in either individual of these domains. The position, spatial frequency and orientation selectivity properties are believed to have an important role in visual perception. Our project attempts a novel face representation and recognition approach by exploring information jointly in image space, scale and orientation domains. Specifically, the face image is first decomposed into different scale and orientation responses by convolving multi-scale and multi-orientation Gabor filters. These multiple Gabor-filters help in alleviating variations of face expression and illumination. Second, local binary pattern analysis is used to describe the neighboring relationship not only in image space, but also in different scale and orientation responses. This way, information from different domains is explored to give a good face representation for recognition. Discriminant classification is then performed based upon LDA followed by k-Nearest Neighbour technique.**

I. INTRODUCTION

Face recognition has attracted much attention due to its potential value for applications and its theoretical challenges. In real world, the face images are usually affected by different expressions, poses, occlusions and illuminations, and the difference of face images from the same person could be larger than those from different ones. Therefore, how to extract robust and discriminant features which make the intraperson faces compact and enlarge the margin among different persons becomes a critical and difficult problem in face recognition. .

Till date, many face representation approaches have been introduced. Gabor Filters and local binary patterns (LBPs) are two representative features. **Gabor wavelets** capture the local structure corresponding to specific spatial frequency (scale), spatial locality, and selective orientation which are demonstrated to be discriminative and robust to illumination and expression changes.**LBP operator** which describes the neighboring changes around the central point, is a simple yet effective way to represent faces. It is invariant to any monotonic gray scale transformation and is, therefore, robust to illumination changes to some extent. Authors in [4] proposed LBP on Gabor magnitude and phase representation.This combination has improved face recognition performance significantly when compared to the individual representation.

II. OUR PROJECT

As a part of this project, we work on a novel face representation method that not only explores the information in the spatial domain, but also among different scales and orientations. The main procedure of the proposed joint information extraction is as follows:
Multi-scale and multi-orientation representations are derived by convolving the face image with a **Gabor filter bank** and formulated as a third-order volume. In order to reduce the computational complexity, we use an **effective GV-LBP** (E-GV-LBP) descriptor that models the neighboring changes around the central point in the joint domains simultaneously for face representation. Next, Discriminant classification is performed based on linear discriminant analysis (LDA) upon the histograms of E-GV-LBP patterns to lower the dimension and finally it is followed with **K-nearest neighbors(kNN)** technique for face recognition.

In this way, we encode the neighboring information not only in image space but also among different scales and orientations of Gabor faces.
The rest of the report is organized as follows. In Section III sub-section A describes the definitions of Gabor filters used,sub-section B reviews the LBP formulation, sub-section C describes the Effective GV-LBP used for obtaining histograms and sub-section D presents the LDA and k-NN techniques for face recognition. Section IV shows our results obtained on the popular FERET and ORL Face database. Section V and VI describes our conclusions and future work respectively.

III. E-GV-LBP BASED FACE REPRESENTATION

A. Gabor Faces

Gabor filters, which exhibit desirable characteristics of spatial locality and orientation selectively and are optimally localized in the space and frequency domains, have been extensively and successfully used in face recognition. The Gabor kernels we used are defined as follows:

$$\psi_{\mu,\nu} = \frac{k_{\mu,\nu}^2}{\sigma^2} exp\left(-\frac{k_{\mu,\nu}^2 z^2}{2\sigma^2}\right)\left[exp(ik_{\mu,\nu}z) - exp(-\frac{\sigma^2}{2})\right] \tag{1}$$

where μ and ν define the orientation and scale of the Gabor kernels, respectively,z= (x,y), and the wave vector $k_{\mu,\nu}$ is defined as:

$$k_{\mu,\nu} = k_\nu e^{i\phi_\mu} \tag{2}$$

where $k_{\mu,\nu} = k_{max}/f^v, k_{max} = \pi/2, f = \sqrt{2}, \phi_\mu = \frac{\pi\mu}{8}$.

The Gabor kernels in (1) are all self-similar since they can be generated from one filter, the mother wavelet, by scaling and rotating via the wave vector $k_{\mu,\nu}$. Hence, a band of Gabor filters is generated by a set of various scales and rotations. In this project, we use Gabor kernels at five scales $\mu \in \{0, 1, 2, 3, 4\}$ and eight orientations $\nu \in \{0, 1, 2, 3, 4, 5, 6, 7\}$ with the parameter $\sigma = 2\pi$ to derive the Gabor representation by convolving face images with corresponding Gabor kernels. For every image pixel we have totally 40 Gabor magnitude and phase coefficients, respectively, that is to say, we can obtain 40 Gabor magnitude and 40 Gabor phase faces from a single input face image.The Gabor filter bank looks like:

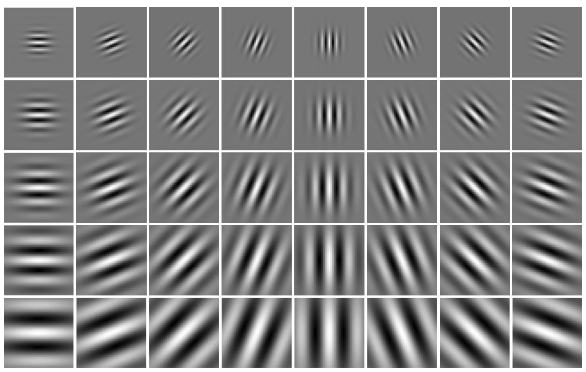


FIG. 1: Gabor Filter Bank

B. Gabor Volume Based LBP

LBP is introduced as a powerful local descriptor for micro-features of images. The basic LBP operator labels the pixels of an image by thresholding the 3X3-neighborhood of each pixel with the center value and considering the result as a binary number (or called LBP codes). An illustration of the basic LBP operator is shown in Fig. 2.

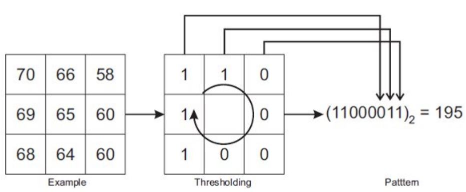


FIG. 2: Calculation of LBP code from 3X3 subwindow

Recently, the combination of Gabor and LBP has been demonstrated to be an effective way for face recognition. Our project focuses on exploring discriminative information by modeling the neighboring relationship not only in spatial domain, but also among different frequency and orientation properties.

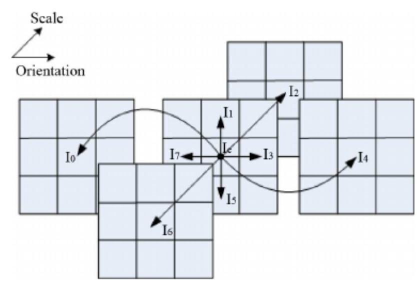
C. Effective GV-LBP

The input image is convolved with the Gabor-filters as described above and then LBP pattern is used to encode the 40 images obtained after convolution. The GV-LBP-TOP [4] is of high computational complexity. The length of the histogram feature vector and the computational cost are threefold compared to those of LGBPHS, so it is not very efficient in practical application. To address this problem, we used an effective formulation of GV-LBP **(E-GV-LBP)** which encodes the information in spatial, frequency and orientation domains simultaneously and reduces the computational cost. Fig.3 shows the definition of E-GV-LBP coding. For the central point I_c, I_0 and I_4 are the orientation neighboring pixels;I_2 and I_6 are the scale neighboring ones; $I_1, I_3, I_5 and I_7$ are the neighboring pixels in spatial domains. Like in LBP, all the values of these pixels surrounded are compared to the value of the central pixel, thresholded into 0 or 1 and transformed into a value between 0 and 255 to form the E-GV-LBP value

$$E - GV - LBP = \sum_{p=0}^{7} 2^p S(I_p - I_c) \tag{3}$$

where $S(I_p - I_c)$ is a threshold function defined as

$$S(I_p - I_c) = \begin{cases} 1, & \text{if } I_p - I_c \geq 0 \\ 0, & \text{if } I_p - I_c < 0. \end{cases}$$



The histogram features are computed based upon the E-GV-LBP codes to provide a more reliable description as

$$H(l) = \sum_{x,y} I(f(x, y) = l), l = 0, 1, ..., L - 1 \tag{4}$$

where I(.)∈0,1is an indication function of a boolean condition and $f(.)$ denotes the E-GV-LBP codes, L and is the number of the E-GV-LBP codes.

The histograms obtained for E-GV-LBP codes for a particular gabor filter($\mu = 2, \nu = 3$) convolved with input image Fig.3. are as shown in Fig.4.

D. LDA and k-NN

LDA is a representative subspace learning method which has achieved great success in face recognition. In this part, we conduct LDA on the selected features to learn the most discriminant subspace for classification. The essential idea of LDA is to disperse the samples from different classes and meanwhile gather the samples from the same class. Given the training samples Z = $Z_1, Z_2, ..., Z_n$ the between class scatter matrix S_b and within class scatter matrix S_w are defined as

$$S_b = \frac{1}{n} \sum_{i=1}^{L} n_i (m_i - m)(m_i - m)^T \tag{5}$$

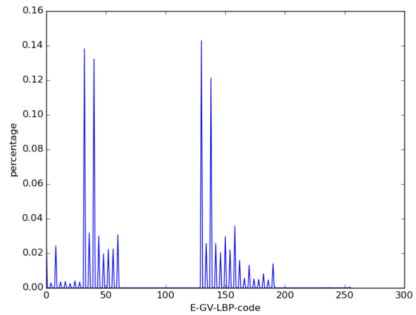


FIG. 3: Histogram of Magnitude distribution of E-GV-LBP code

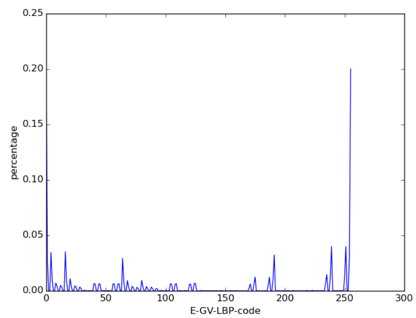


FIG. 4: Histogram of Phase distribution of E-GV-LBP code

$$S_w = \frac{1}{n}\sum_{i=1}^{L}\sum_{Z_j \in C_i} n(Z_j - m_i)(Z_j - m_i)^T \qquad (6)$$

where $m_i = \frac{1}{n_i}\sum_{Z_j \in C_i} Z_j$ is the mean vector in the class C_i, L is the class number and $m_i = \frac{1}{n_i}\sum_{i=1}^{L}\sum_{Z_j \in C_i} Z_j$ is the global mean vector.LDA aims to find the projective directions which maximize the ratio of between class scatter matrix to within class scatter matrix.

In k-NN method,

- The training examples are vectors in a multidimensional feature space, each with a class label. The training phase of the algorithm consists only of storing the feature vectors and class labels of the training samples.
- In the classification phase, k is a user-defined constant, and an unlabeled vector (a query or test point) is classified by assigning the label which is most frequent among the k training samples nearest to that query point
- The output is a class membership. An object is classified by a majority vote of its neighbors, with the object being assigned to the class most common among its k nearest neighbors (k is a positive integer, typically small). If k = 1, then the object is simply assigned to the class of that single nearest neighbor.

IV. EXPERIMENTS

A. Dataset: FERET Database

The FERET database is one of the largest publicly available databases.It contains colored images with different illuminations,poses,expressions,occulsions etc.The images in the database were cropped to a fixed size and are scaled and rotated to align along their eyes.

In this experiment, the training set contains 110 images of 22 persons, 5 each. Test dataset contains 66 images, 3 for each class/person. All the images in the training dataset are rotated, scaled and cropped to into 205*205 size according to the provided eye coordinates in the dataset.

The ROC and CMC curves obtained are shown in Fig 5 and Fig 6 respectively: The top 10 matches obtained on FERET database are as shown in Fig 7 with the input image on top and the top ten matches are below.

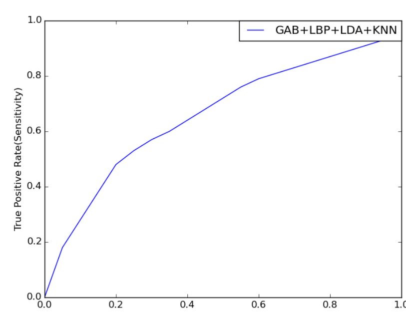


FIG. 5: ROC_ColorFeret

B. Dataset: ORL Face Database

The ORL Database of Faces, contains a set of face images taken between April 1992 and April 1994.There are ten different images of each of 40 distinct subjects. For some

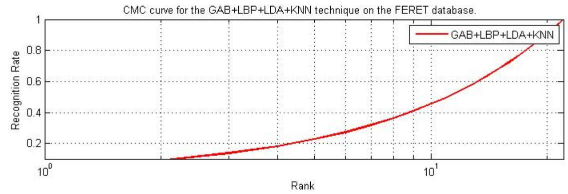


FIG. 6: CMS_ColorFeret

FIG. 7: Top 10 retrievals for FERET Database

subjects, the images were taken at different times, varying the lighting, facial expressions (open / closed eyes, smiling / not smiling) and facial details (glasses / no glasses). All the images were taken against a dark homogeneous background with the subjects in an upright, frontal position (with tolerance for some side movement). All the images in the training dataset are rotated, scaled and cropped to into 92*112 size according to the provided eye coordinates in the dataset.The top-10 matches for the ORL database are shown below with input sample image on top and corresponding top 10 matches shown in figure . The obtained ROC,CMS curves are as follows:

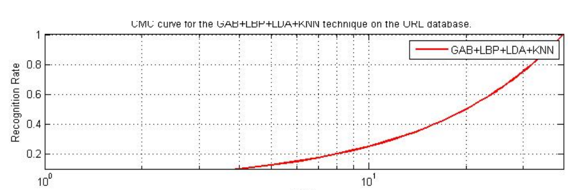


FIG. 8: CMS_ORL

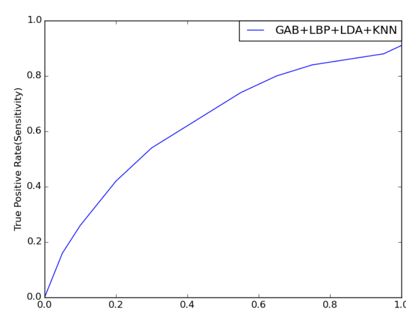


FIG. 9: ROC_ORL

V. CONCLUSION

From our experimental results we conclude that there are mainly three advantages for the proposed method.

- First,the Gabor feature is applied to the face images to alleviate the variations of facial expression and illumination.It is robust to any homogenous illuminations.
- Second,the LBP is utilized to model the neighboring

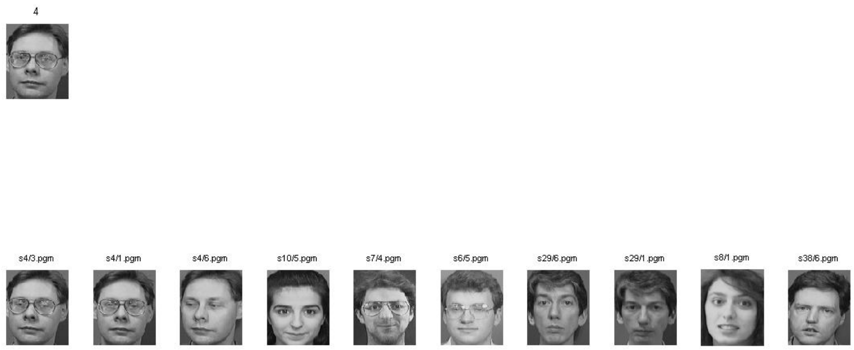


FIG. 10: Top 10 retrievals on ORL Database

relationship jointly in spatial, frequency and orientation domains. In this way, discriminant and robust information, as much as possible, could be explored.
- Third, LDA and k-NN analysis method is introduced to make the face representation compact and effective for face recognition.

VI. FUTURE WORK

- The dimension of LBP code obtained could be further reduced by performing statistical uniform pattern[4] to obtain better results
- we observed that different regions of face make different contributions for the performance of recognition e.g.,the areas nearby eyes and nose are more important than others. Therefore, it is sensible to assign different weights onto different blocks when measuring the dissimilarity of two images.
- Weighted histogram intersection with fisher criteria might be another possible approach for problem.

VII. ACKNOWLEDGEMENTS

”Portions of the research in this paper use the FERET database of facial images collected under the FERET program, sponsored by the DOD Counterdrug Technology Development Program Office”

REFERENCES

[1] C. Liu and H. Wechsler, Gabor feature based classification using the enhanced fisher linear discriminant model for face recognition, IEEE Trans. Image Process., vol. 11, no. 4, pp. 467476, Apr. 2002.
[2] Z. Lei, S. Z. Li, R. Chu, and X. Zhu, Face recognition with local gabor textons, in Proc. IAPR/IEEE Int. Conf. Biometr., 2007, pp. 4957.
[3] T. Ahonen, A. Hadid, and M. Pietikainen, Face description with local binary patterns: Application to face recognition, IEEE Trans. Pattern Anal. Mach. Intell., vol. 28, no. 12, pp. 20372041, Dec. 2006.
[4] Zhen Lei, Shengcai Liao, Matti Pietikinen, and Stan Z. Li ”Face recognition by exploring information jointly in space, scale and orientation.IEEE transactions on image processing”, vol.20,no.1, jan 2011.
[5] P.J. Phillips, H. Moon, S.A. Rizvi, P.J. Rauss, ”The FERET Evaluation Methodology for Face Recognition Algorithms,” IEEE Trans. Pattern Analysis and Machine Intelligence, Vol. 22, pp. 1090-1104, 2000