

Face Recognition Under varying Lighting Conditions and Noise Using Texture Based and SIFT Feature Sets

¹Vishnupriya. S, ²Dr. K. Lakshmi

¹Dept. of CSE, Roever Engineering College

²Dept. of CSE, Periyar Maniammai University

Abstract

Recognizing face images more reliable even under uncontrolled lighting conditions is one of the most important challenges for practical face recognition systems. We tackle this by combining the strengths of robust illumination normalization, local texture-based face representations, multiple feature fusion. Additionally we propose SIFT which is an approach for detecting and extracting local feature descriptors. By combining the results produced from both the above mentioned approaches, we increase the accuracy of recognizing the given image from group image even with difficult lighting conditions and noise.

Keywords

Face Recognition, Illumination Invariance, Image Preprocessing, Local Ternary Patterns, Gabor Features, Scale Invariant Feature Transform Features

1. Introduction

Face recognition by a computer system is a very popular and useful application of image analysis. It is a method to identify or verify the identity of a person. In an identification process, based on face recognition, the face image of an unknown identity is compared with face images of known individuals from a large database. Finally the system returns the determined identity.

In a verification process the face image of a person is compared with one face image from a database with the claimed identity. The system returns a value, which is a measure for the similarity between the two images. Of course this value needs to reach a certain threshold to acknowledge or reject the claimed identity. But most of the methods of face recognition were initially developed with face images collected under relatively well-controlled conditions including eigenfaces [6], fisherfaces [7], and self Quotient image [14], elastic bunch graph matching [9] and in practice they have difficulty in dealing with the range of appearance variations that commonly occur in natural images due to illumination, pose, facial expression, aging, partial occlusions, etc. Generally the changes induced by illumination are often larger than the differences between individuals, causing the face recognition system to misclassify input images.

This paper focuses mainly on the issue of robustness to lighting variations and positions. There are three traditional approaches for dealing with this issue as appearance-based, normalization-based, and feature-based methods [1].

In direct appearance-based approaches, training examples will be collected under different lighting conditions and directly (i.e., without undergoing any lighting preprocessing) used to learn a global model of the possible illumination variations. But since it requires large number of training images and an expressive feature set, it is essential to include a good preprocessor to reduce illumination variations.

Normalization based approaches seek to reduce the image to a more "canonical" form in which the illumination variations get suppressed. Histogram equalization is one simple example. Feature Based approaches extract illumination insensitive feature sets from

the given image. These features may be vary from geometrical features to image derivatives such as Edgmaps, Local Binary patterns and Local autocorrelation filters etc.

In this work, we propose an novel approach that combines the strengths of direct based, feature based and normalization based approaches of face recognition. We try to solve the problem of recognition under varying lighting conditions and noise.

The overall process of our proposed face recognition can be viewed as a pipeline consisting of:

1. An efficient preprocessing chain that eliminates most of the effects of changing illumination.
2. Extracting Local Ternary Patterns (LTP) which is an generalization of Local Binary Patterns (LBP) that captures small appearance details.
3. Extracting Gabor Wavelets which encode facial shape over a broad range
4. Subspace representation using Kernel Linear Discriminant Analysis (KLDA) incorporating features from fusing LTP and Gabor
5. Scale invariant feature transform which a method of detecting and extracting feature descriptors invariant is to Illumination, rotation and scaling
6. Final Recognition is done by classifying the feature vectors using SVM. Each stage increases resistance to illumination variations and makes the information needed for recognition more manifest.

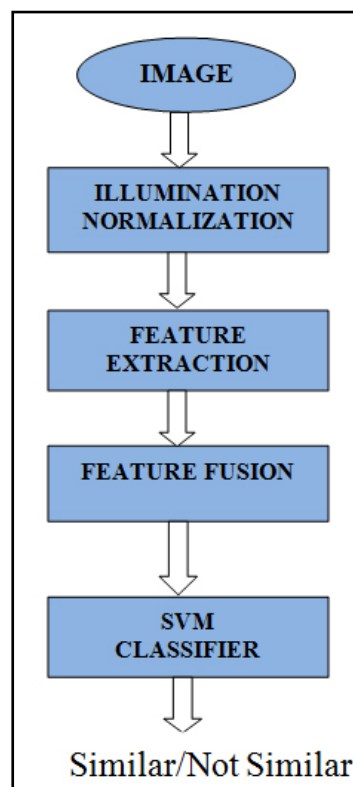


Fig. 1: Block Diagram of the System we used for Face Recognition

The main idea of our proposed system is to provide recognition of images under varying lighting conditions and noise. By extracting LTP and Gabor it is only possible to recognize the images under difficult lighting conditions. So we include the SIFT feature which provides the recognition for images with noise and different positions. Finally by combining all the three features we are able to recognize the images under difficult lighting, noise and in different positions and also increases the performance.

The rest of the paper is organized as follows. Section II presents our existing work, Section III briefly reviews our proposed system. Section IV reports experimental results, and Section V concludes.

II. Related Work

A. Local Binary Pattern

In recent years there has been much work proposed for face recognition. But these methods work well only under controlled conditions i.e. they do not produce satisfiable results under illumination variation, change in poses, occlusions etc. One such approach is extraction of local binary patterns [3]. The LBP Descriptor, which is based on the relative order of neighboring pixels has also shown promise for several applications. Binary Patterns are created for each pixel by comparing a pixel value with its neighboring intensities. The histogram of the such binary patterns computed over a region is used for texture description. But these features are sensitive to noise in near-uniform regions such as cheeks.

The disadvantage is that since LBP based representations typically subdivide the face into regular grid in order to compare histograms of LBPs within each region it seems to be somewhat arbitrary and it is likely to give rise to both aliasing and loss of spatial resolution.

B. Eigen Faces

Turk and Pentland [6] present a face recognition scheme in which face images are projected onto the principal components of the original set of training images. The resulting eigenfaces are classified by comparison with known individuals.

Turk and Pentland present results on a database of 16 subjects with various head orientation, scaling, and lighting. Their images appear identical otherwise with little variation in facial expression, facial details, pose, etc. Scale is renormalized to the eigenface size based on an estimate of the head size. The middle of the faces is accentuated, reducing any negative affect of changing hairstyle and backgrounds.

Even though eigenfaces is a fast, simple, and practical algorithm, it may be limited because optimal performance requires a high degree of correlation between the pixel intensities of the training and test images. This limitation has been overcome by using extensive preprocessing to normalize the images.

C. Template Matching

Template matching methods such as [10], operate by performing direct correlation of image segments. Template matching is only effective when the query images have the same scale, orientation, and illumination as the training images.

D. Graph Matching

Another approach to face recognition is the well known method of graph matching. In [9], Lades et al. present dynamic link architecture for distortion invariant object recognition which

employs elastic graph matching to find the closest stored graph. Objects are represented with sparse graphs whose vertices are labeled with a multiresolution description in terms of a local power spectrum, and whose edges are labeled with geometrical distances.

III. Proposed System

A. Preprocessing

This section includes a preprocessing [1] chain run before feature extraction that incorporates a series of stages that remove the effects of illumination variations, local shadowing, and highlights while preserving the essential elements of visual appearance.

Gamma Correction is a nonlinear gray-level transformation that replaces gray-level I with I_γ (for $\gamma > 0$) where $\gamma \in [0, 1]$ is a user-defined parameter. This method enhances the local dynamic range of the image in dark or shadowed regions while compressing it in bright regions and at highlights.

After gamma correction, DoG (Difference of Gaussian) filter is applied which suppresses the highest spatial frequencies potentially reduces both aliasing and noise without destroying too much of the underlying recognition signal.

The final stage of preprocessing chain is the equalization of variation which rescales the image intensities to standardize a robust measure of overall contrast or intensity variation.

B. Feature Extraction

This section includes LTP, Gabor and SIFT Feature extraction methods. Even though existing feature sets offer good performance under illumination variation there is still room for improvement.

The LTP [1] operator takes a local neighborhood around each pixel thresholds the pixel of the neighborhood at the central value in which gray-levels in a zone of width $\pm t$ around i_c are quantized to zero, ones above this are quantized to +1 and ones below it to -1, and the binary LBP code is replaced by a ternary LTP code.

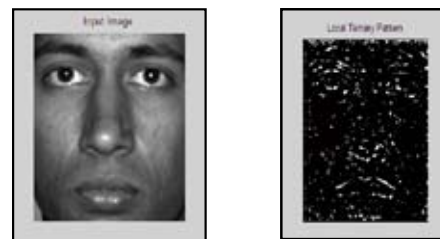


Fig. 2: From Left to Right: Image and its Binary Layer

$$D(X, Y) = \sum_{\text{pixels}(i,j) \text{ of } Y} w(d_x^{k_y(i,j)}(i, j)) \quad (4)$$

Gabor filter [2], works as a bandpass filter for the local spatial frequency distribution, achieving an optimal resolution in both spatial and frequency domains. Convolve the image with complex Gabor filters.

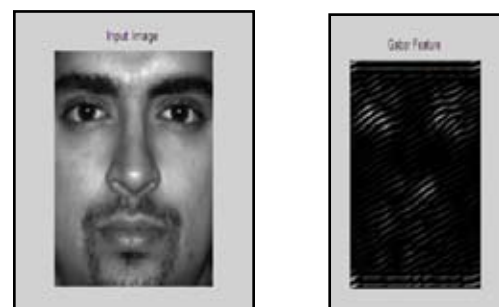


Fig. 3: From Left to Right: Image and its Gabor Feature

The optimal projection operator[1] is then $P = \bar{\phi} U \Lambda^{-1/2} V$ and test examples can be projected into the optimal discriminant space.

When a face image is presented to the system, its Gabor wavelet and LTP features are extracted, separately projected into their optimal discriminant spaces by KLDA method (4) and used to compute the corresponding distance scores. This similarity score is provided as input to the SVM classifier for final decision.

Scale Invariant Feature Transform (SIFT) is an approach for detecting and extracting local feature Descriptors [11] that are reasonably invariant to changes in illumination, image noise, rotation, scaling, and small changes in viewpoint.

The first step toward the detection of interest points is the convolution of the image with Gaussian filters at different scales, and the generation of difference-of-Gaussian images from the difference of adjacent blurred images.

Interest points (called keypoints in the SIFT framework) are identified as local maxima or minima of the DoG images across scales. Each pixel in the DoG images is compared to its 8 neighbors at the same scale, plus the 9 corresponding neighbors at neighboring scales. If the pixel is a local maximum or minimum, it is selected as a candidate keypoint.

To determine the keypoint orientation, a gradient orientation histogram is computed in the neighborhood of the keypoint (using the Gaussian image at the closest scale to the keypoint's scale). Peaks in the histogram correspond to dominant orientations. A separate keypoint is created for the direction corresponding to the histogram maximum, and any other direction within 80% of the maximum value.

Find the SIFT features from the input image. Match these features to the SIFT feature database and provide the resulting image which has the minimum Euclidean distance. Each keypoint specifies four parameters: 2D location, scale and orientation.

C. SVM Classification

The full method incorporates the aforementioned preprocessing chain and LBP or LTP features with distance transform based comparison and SIFT feature Extraction. However, since face recognition is a complex task for which it is necessary to include multiple types of features, we build a final classification stage that can handle residual variability and learn effective models from relatively few training samples.

We combine the similarity scores obtained from LTP, Gabor and SIFT features. Providing these scores as input to the SVM classifier we gain the final decision of whether the original image is recognized or not.

IV. Experiment Results

We illustrate the effectiveness of our proposed methods by presenting the experiments on Yale-B dataset with difficult lighting condition. The dataset contains 39 subjects under 9 poses and 20 illumination conditions. The size of each image is 128 X 128 or 200 X 200 pixels.

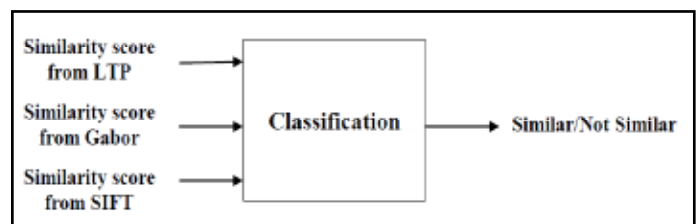
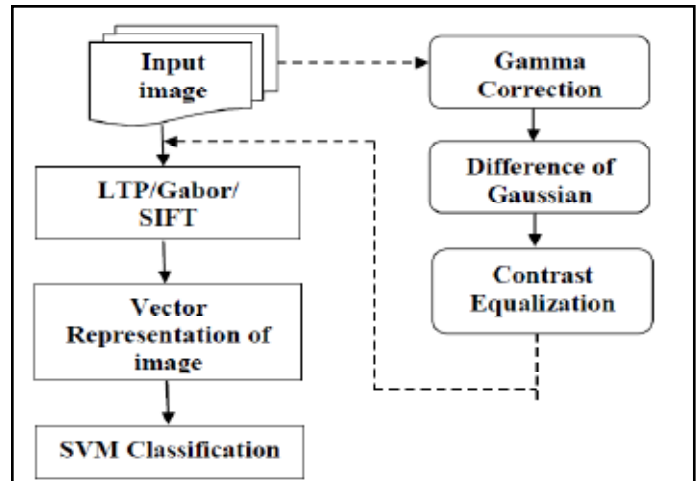


Fig. 4: Overall Architecture of Proposed Face Recognition (From Top to Bottom: Training and Testing the Classifier)

The results are divided into two sections, first focusing on KLDA based classifier with LTP and Gabor and the second on Scale invariant feature transform features.

A. Data Sets

For the experiments, images from the Yale Face Database B were used that contains images of 39 faces each under 64 different lighting conditions, and the test is performed on 39 images.

On Yale-B dataset initially all the images were converted into gray scale for best results. In order to suppress the overall illumination, preprocessing has been done. With γ values ranging from 0.2 to 0.5 in Gamma correction. These images were then applied difference of Gaussian filtering with σ_0 with the value of 4 and σ_1 with the value of 2 so that overall shading effects will be induced.

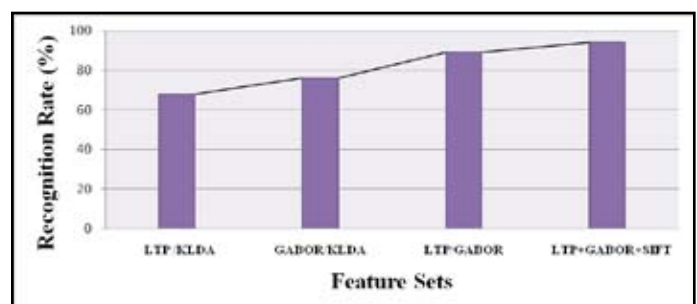


Fig. 5: Improvement Due to the Addition of SIFT Feature Along With LTP and Gabor Features on Extended Yale-B Dataset

Finally in order to equalize the overall contrast we provide the value of α to be 0.1 and τ be 10.

The preprocessed image is provided as input to the Feature extraction. In this approach LBP is replaced with LTP in order to avoid noise in the resulting feature. For the LTP features the image contains 8 X 8 LTP blocks with 59 uniform codes. Next Gabor filter is applied to each 200 X 200 preprocessed image with 5 scales and 5 directions in order to increase the spatial resolution.

The LTP and Gabor features are extracted from the preprocessed images and these features were fused. KLDA was used for dimensionality reduction. In order to test the recognition results alone on these two features

Further to increase the accuracy another feature descriptor called SIFT is included which divides the image into blocks of 16 X 16 pixels and extract a feature vector or keypoint which are invariant to noise, rotation, scaling and changes in view point from each block with the spacing of 8 pixels. SIFT descriptor is of 128 bit value and unique.

Finally the LTP, Gabor and SIFT features are extracted from the images.

Given a set of training examples, each marked as belonging to one of two categories, an SVM training algorithm builds a model that assigns new examples into one category or the other.

For classification we took 39 positive images and 39 negative images. We provide the similarity score obtained from the LTP, Gabor and SIFT feature as input to the classifier. Two classes are assigned to the classifier as similar and not similar.

Next the test images will be compared with the database images and when the similar face appears it displays that particular image as similar image otherwise it provide the results as not similar.

The final accuracy of the system is calculated by the following formulas.

$$\text{Accuracy} = \frac{TP + TN}{(P + N)}$$

where TP is true positive, FP is false positive, P is the total number of positives and N is the total number of negatives.

From the results obtained on testing with images under difficult lighting from Yale-B dataset, LTP and Gabor feature extraction method provides the accuracy as 89% along with preprocessing. By adding SIFT feature alone further improves the accuracy by 5% (from 89% to 94%) and also provides the recognition under noise and different positions.

Overall our proposed methods increases the performance for recognizing images under different lighting noise and and it is also very fast and easy to implement.

VIII. Conclusion

The new method for face recognition under uncontrolled lighting and noise has been proposed based on robust preprocessing and an extension of the LBP local texture descriptor and SIFT feature descriptor. The main contributions are as follows:

1. A simple, efficient image preprocessing chain for suppressing the overall illumination
2. A rich descriptor for local texture called LTP that generalizes LBP while fragmenting less under noise in uniform regions
3. A method for extracting distinctive invariant features called SIFT for detecting local feature descriptors that are invariant to noise, rotation, scaling and small changes in viewpoint
- 4) Final recognition is given by providing the similarity score obtained from each feature set as input to the SVM classifier.

By including the SIFT feature along with LTP and Gabor feature, the images which are invariant to lighting, noise and various position could also be recognized. From the experiments tested with the Yale-B dataset, the results shown that by including the SIFT feature along with LTP and Gabor the accuracy has been increased to 94%.

References

- [1] Xiaoyang Tan, Bill Triggs, "Enhanced Local Texture Feature Sets for Face Recognition Under Difficult Lighting Conditions", IEEE Transactions On Image Processing, Vol. 19, No. 6, June 2010.
- [2] X. Tan, B. Triggs, "Fusing Gabor and LBP feature sets for kernel based face recognition", in Proc. AMFG, 2007, pp. 235–249.
- [3] T. Ahonen, A. Hadid, M. Pietikainen, "Face recognition with local binary patterns", in Eur. Conf. Comput. Vis., Prague, Czech Republic, 2005, pp. 469–481.
- [4] T. Ahonen, A. Hadid, M. Pietikainen, "Face description with local binary patterns: Application to face recognition", IEEE Trans. Pattern Anal. Mach. Intell., Vol. 28, No. 12, pp. 2037–2041, Dec. 2006.
- [5] Y. Adini, Y. Moses, S. Ullman, "Face recognition: The problem of compensating for changes in illumination direction", IEEE Trans. Pattern Anal. Mach. Intell., Vol. 19, No. 7, pp. 721–732, Jul. 1997.
- [6] M. Turk, A. Pentland, "Eigenfaces for recognition", J. Cognitive Neurosci., Vol. 3, No. 1, pp. 71–86, 1991
- [7] P. Belhumeur, J. Hespanha, D. Kriegman, "Eigenfaces vs. Fisherfaces: Recognition using class specific linear projection", IEEE Trans. Pattern Anal. Mach. Intell., Vol. 19, No. 7, pp. 711–720, Jul. 1997.
- [8] G. Borgefors, "Distance transformations in digital images", Comput. Vis., Graphics Image Process., Vol. 34, No. 3, pp. 344–371, 1986.
- [9] L. Wiskott, J.-M. Fellous, N. Krüger, C. von der Malsburg, "Face recognition by elastic bunch graph matching", IEEE Trans. Pattern Anal. Mach. Intell., Vol. 19, No. 7, pp. 775–779, Jul. 1997.
- [10] R. Brunelli, T. Poggio, "Face recognition: Features versus Templates", IEEE Trans. Pattern Anal. Mach. Intell., Vol. 15, No. 10, pp. 1042–1052, Oct. 1993
- [11] Lowe, D. G., "Distinctive Image Features from Scale-Invariant Keypoints", International Journal of Computer Vision, 60, 2, pp. 91–110, 2004.
- [12] D. Wagner, G. Reitmayr, A. Mulloni, T. Drummond, D. Schmalstieg, "Pose tracking from natural features on mobile phones", Proceedings of the International Symposium on Mixed and Augmented Reality, 2008.
- [13] Y. Pang, Y. Yuan, X. Li, "Gabor-based region covariance matrices for face recognition", IEEE Trans. Circuits Syst. Video Technol., vol. 18, no. 7, pp. 989–993, Jul. 2008.
- [14] H. Wang, S. Li, Y. Wang, "Face recognition under varying lighting conditions using self quotient image", in Proc. IEEE Int. Conf. Autom. Face Gesture Recognition, 2004, pp. 819–824.
- [15] S. Yan, H. Wang, X. Tang, T. Huang, "Exploring feature descriptors for face recognition", in Proc. IEEE Int. Conf. Acoust., Speech, Signal Process. (ICASSP), 2007, pp. 629–632.



S.Vishnupriya received B.E degree in Computer Science and Engineering from M.Kumarasamy College of Engineering in 2010 and then completed M.E degree in Computer Science and Engineering at Periyar Maniammai University in 2012. Research interests includes Image Processing and Data Mining. At present working as Assistant Professor at Roever Engineering College, Anna University.



K.Lakshmi received B.E degree in Electronics and Instrumentation Engineering, in Annamalai University in 1994 and then received M.E degree in Computer Science and Engineering in Madurai Kamaraj University in 1999. Completed Ph.D in Computer Science and Engineering at Anna University in 2009. Currently working as Professor in Periyar Maniammai University. Research interests in Text

Mining, Machine Learning, Information Retrieval.