ISSN: 2277-9655
(I2OR), Publication Impact Factor: 3.785

IJESRT

INTERNATIONAL JOURNAL OF ENGINEERING SCIENCES & RESEARCH TECHNOLOGY

VIOLA JONES AND SIFT FRAMEWORK FOR FACE RECOGNITION

Priyanka Raosaheb Borude*, Dr.S.T.Gandhe
PG Student ,Electronics& Telecomm.Engg,SITRC, Nasik,India
Principal, Electronics& Telecomm.Engg,SITRC, Nasik,India

ABSTRACT

Faces are highly deformable objects which can easily change their appearance over time. All face areas are not subjected to the same variability. Hence decoupling the information from independent areas of the face is of much importance to improve the robustness and efficiency of any face recognition technique.So in this paper the SIFT (Scale Invariant Feature Transform) and Viola Jones Framework is proposed for efficient face recognition using Nearest neighbor classifier. The proposed framework works efficiently on ORL face dataset even under the challenges of varying expression, varying pose, occlusion with the recognition rate of 90%.

KEYWORDS:
Face recognition, Occlusion, SIFT, Nearest neighbor classifier

INTRODUCTION

Face recognition has attracted much attention in last decade because of its wide applications. However, face recognition is yet an unresolved problem as human face is not rigid object and it can be transformed easily under different situations. So, how to represent the intrinsic attributes of a human face effectivelyis more important to increase the accuracy of face recognition systems.[2]

Human faces can be characterized on the basis of global as well as local features. Although local features on the face can be highly discriminative, but they suffer for local changes in the appearance of face or face occlusion. On the other hand, global features are easier to capture they are basically less discriminative than that of localized features, but are less sensitive to deformations in the face due to the partial deformability of the facial structure. The optimal face representation must then allow the matching of global facial features, and also determination of local similarity measurement for the face. So there is a need to provide a combined framework that extracts both global and local features of a face so as to give better representation of face for recognition and verification purposes.

This paper represents a framework using SIFT (Scale Invariant Feature Transform) and Viola Jones algorithm for face recognition. First Scale Invariant Feature transform extracts stable keypoints of given query face while the Viola Jones algorithm extracts facial features i.e, mouth,eyes,nose,of face. Then Nearest neighbor classifier provides a correct match for given query image by matching extracted features of face.

STATE OF ART

A general face recognition process starts with face detection, feature extraction, and then face identification and verification. Feature extraction stage in face recognition forms a firm basis for face recognition task. Hence locating facial features in face image is a crucial step for number of applications.[1] Dimension reduction is important step in pattern recognition and machine learning. It is difficult to directly use raw data not only because significant parts of data have not been extracted, but also because of extremely high dimensionality of raw data. This section well describes the Facial Feature Extraction state of art methods and face recognition techniques used traditionally
Some image processing techniques extract feature points such as nose, eyes, mouth and then used as input data to application. For some application it has been the central step. Various approaches have been proposed to extract these facial points from video sequences and images of faces. The four basic techniques in this direction are as follow:

A. Geometry-based Techniques

In these methods features are extracted by using sizes and relative positions of the important components of face. This group of methods concentrates in two directions. Firstly it detects edges, directions of crucial components or region of images contain ing important components, it then builds feature vectors from these edges and directions. Then it includes use of filters like Canny filter, methods based on the grayscales difference of important components and unimportant components, the gradient analysis method, Gabor Wavelets Transform(4),(3) ,LBP(5) Adaboost method(6).

B. Color Segmentation Based Techniques

This approach makes use of skin color for isolating face region. Then any non-skin color region in the face is viewed as a candidate for eyes or mouth. By using Color models such as HSV, RGB ,or YCbCr with certain range of color pixels, skin region is detected. After getting the skin region, facial features viz. Eyes and Mouth are extracted. The image is then obtained after application of skin color statistics and then it is subjected to binarization where it is transformed to gray-scale image and then to a binary image by applying suitable threshold. This is done to eliminate the saturation and hue values and it then considers only the luminance part. Then this luminance part is changed to binary image with some threshold because the features for face are darker than the background colors. After thresholding, closing and opening operations are performed for removing noise. These are a sort of the morphological operations, that are used to remove holes.

C. Appearance Based Techniques

This type of methods use the statistical methods and linear transformation to find the basic vectors to represent the face. Several methods have been proposed in the literature to satisfy this aim, like ICA and PCA [7].Basically PCA method is used to reduce the number of dimensions of feature space , keeping principle features so as to minimize loss of information. PCA method considers second-order statistic in the data. But, PCA method has some demerits because high order dependencies exist in PCA analysis. While ICA (independent component analysis) is the analysis technique that uses second-order statistic as well as high order statistics in the data.

D. Template Based Techniques

This group of methods extracts facial features such as mouth, eyes etc. based on template function and appropriate energy function. For ex. if eye needs to be extracted then first an eye template is used to find the eye from face image. The correlation of eye template with various overlapping regions of the face image is found out. Then the region with maximum correlation with the template refers to eye region.

To extend the feature extraction stage in face recognition for the data dimensionality, different appearance-based techniques have been successfully used, such as the Principal Component Analysis (PCA) [8], Fisher Discriminant Analysis (FDA) [8],,Linear Discriminant Analysis (LDA) [8], Independent Component Analysis (ICA) [8], and. Other methods have been also studied for the extraction of salient facial features by application of cascaded scale-space filtering [9,10,11,12]. For most of the times, one missing part is the bridge between the geometry of the face itself and features extracted from the face images.

.

PROPOSED METHOD

A.Viola Jones Algorithm

It is a "paradigmatic" method for real-time object detection .In this method the training is slow, but detection is very fast.It includes following Key steps:

I..Obtaining Image Features

II. Integral images computation for fast feature evaluation

III. Boosting for feature selection

IV..Attentional cascade for rejection of non-face windows

I.Obtaining Image Features: Features are obtained by applying rectangle filters .Feature value is evaluated as

Value = $\sum$ (pixels in white area) – $\sum$ (pixels in black area)

II.Integral images computation for fast feature evaluation : The integral image evaluates a value at each pixel (x,y) which is equal to the sum of the pixel values above and to the left of (x,y) pixel inclusive.It can be rapidlycomputed in one pass through the image. For a 24x24 detection region, the possible number of rectangle features is nearly 180,000! In this case, it is impractical to calculate the entire feature set .So here we can create a good classifier using just a small subset of all possible features which is selected by Boosting scheme.

III. Boosting for feature selection : Boosting is a classification scheme that works by combination of weak learners into a more accurate ensemble classifier . Weak learners are then defined on the basis of rectangle features.For a set of weak classifiers

$$h_j(\mathbf{x}) = \begin{cases} -s_j & \text{if } f_j < \theta_j \\ s_j & \text{otherwise} \end{cases}$$

Then repetitively classifiers are combined to form a linear combination as:

$$h(\mathbf{x}) = \text{sign}\left(\sum_{j=1}^{M} \alpha_j h_j(\mathbf{x})\right)$$

Best weak learner is found for current weighted training set then increase the weights of training examples that are misclassified by current weak learner .Final classifier is then computed as linear combination of all weak learners where weight of each learner is related to its accuracy. For each round of boosting: we evaluate each rectangle filter on each example then select best threshold for each filter and then select best filter/threshold combination .

IV.Attentional cascade for rejection of non-face windows: This step start with simple classifiers which reject many of the negative sub-windows while detecting almost all positive sub-windows. Here positive results the first classifier triggers the evaluation of a second (more complex) classifier, and so on. Then a negative outcome at any point leads to the immediate rejection of the sub-window. The chain classifiers are more complex and have lower false positive rates. For training the cascade we adjust weak learner threshold to minimize false negatives (as opposed to total classification error) .Each classifier is then trained on false positives of previous stages [14]

B.Scale Invariant Feature Transform

Scale Invariant Feature Transform (SIFT) , a novel technique was introduced by David Lowe to extract features from images. These features are invariant to scale, rotation, and 3D projective transform. They provide an efficient matching across a wide range of change in affine distortion, 3D viewpoint,addition of noise and varying illumination.Thus SIFT provide a set of stable features of an object that are not affected by occlusion, clutter, unwanted noise present in the image. As SIFT features are highly distinctive in nature they have been used for correct matching on various pair of feature points with higher probability in between a test sample and a large database .Following four importan steps are utilized to generate the set of stable feature based on SIFT.

I. *Scale-Space Extrema Detection*
II. *Keypoints Localization in Laplacian Space*
III. *Assignment of Orientation*
IV. *Keypoint Descriptor*

***I.Scale-Space Extrema Detection*:** This is a filtering approach used to identify image locations and scales that are identifiable from different views. Difference of Gaussian (DoG) which is image enhancement technique and scale space functions are used to find stable keypoints. Difference of Gaussian is used for detection of key-points in scale-space and scale space extrema is located by taking difference between two images, out of which one is scaled by some constant time with respect to other. Now to find the local minima and maxima, each stable feature point is then compared with its 8 neighbors at the same scale and also with its 9 neighbors up and down by one scale. If this value is the maximum or minimum of all these stable feature points then this point is considered as an extrema.

II.Keypoints Localization in Laplacian Space: For locating the keypoints, some points after detection of stable keypoint locations which are having a low contrast or are poorly localized on an edge are eliminated. This is done by evaluating the Laplacian space. After detection of the location of extremum value, if the difference of Gaussian(DoG) pyramids is less than a given threshold value, then the point is excluded. For a large principle curvature across the edge and a small curvature in the perpendicular direction in the difference of Gaussian function, then poor extrema is located and eliminated.

III.Orientation Assignment: This step assigns a consistent orientation to the key-points on the basis of local image characteristics. From the gradient orientations of sample points, the orientation histogram is created in a region around the key-point. Orientation assignment is then followed by key-point descriptor which can be presented according to this orientation. A window of size 16x16 is then chosen to create the histogram. The orientation histogram consist of 36 bins covering almost 360 degree range of orientations. Pixel differences are then

used for evaluation of orientation and gradient magnitude .Every sample is then weighted by a Gaussian-weighted circular window and its gradient magnitude.

IV.Keypoint Descriptor: Here, the feature descriptors which shows local shape distortions and varying illumination are computed. A detailed fitting is then used to the nearby data for the location, peak magnitude and edge response and then candidate locations are found out.A consistent orientation is assigned to each of the feature point on the basis of local image properties to achieve invariance to image rotation,. The histogram of orientations is then produced from the gradient orientation at all sample points in a circular window of a feature point. The peaks in this histogram shows the dominant directions for each of the feature point. 8 orientation planes are then defined for illumination invariance. Then finally, by application of a Gaussian filter the gradient magnitude as well as the orientation are smoothened and are then sampled over a 4 x 4 grid with 8 orientation planes. [13]

RESULTS AND DISCUSSION

For experimentation total 6 different faces are considered with different specificities in the form of occlusion, varying pose,and varying expression as in fig 1.

In the implementation SIFT and Viola Jones framework is proposed for face recognition by extracting static (eyes, nose) and dynamic (mouth) and stable features on the face that are robust to pose variance occlusion and varying expression. For face recognition first static and dynamic features are extracted by Viola Jones algorithm and then stable SIFT features (keypoints) are extracted from all faces in ORL database. Then given a query image, SIFT and facial features are extracted from that face whish are then compared against features from each face in the database, by using nearest neighbor classification algorithm based on Euclidean Distance Matching. The face in the ORL database with the maximum number of matching points is then considered as nearest face and is then used for identification and classification of face. After doing analysis of results obtained on ORL dataset it is found that the proposed framework is able to overcome the challenges of varying expression ,varying pose. Here first query image Fig2. is chosen from database ,then query iamge is preprocessed with Viola Jones Feature detector detects facial features as in fig 3. and then SIFT (Scale Invariant Feature Transform) so that it obtains the stable key points on face that are invariant to varying pose, varying expression as in fig 4.

Fig 1. Six different subjects from the ORL dataset

Fig 2. Query Image from dataset

Fig 3 Facial Features of query image.

Fig 4. Stable Features of query image.

After this step the nearest neighbor classifier is used that is based on Euclidean Distance matching between the probable features and then finally all recognized faces of query image are obtained from database as in fig 5.Face recognition performance is evaluated by two performance measures i.e., recognition rate and error rate.

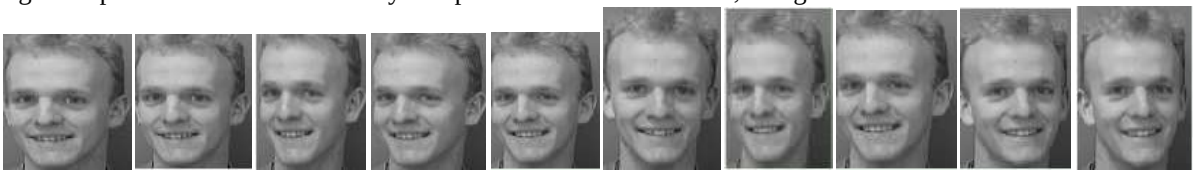

Fig 5.Recognized faces of query face

Table 1. Performance measures for face recognition

Performance measure	Definition	Value (%)
Recognition Rate	No. of Correctly detected faces/Total no.of faces	90
Error Rate	No. of Correctly detected faces/Total no.of faces	10

CONCLUSION

In this paper, an efficient face recognition system is presented using Scale Invariant Feature Transform(SIFT) and Viola Jones Algorithm respectively. The proposed systems shows efficiency towards recognizing the subjects under different specificities like varying pose, occlusion, varying expression. The capability of the system is shown by the results to cope for varying expression occurring in the the query image. Moreover the system well detects facial features for varying expressions. In the face recognition method, local facial features obtained by Viola Jones method and stable keypoints obtained from SIFT algorithm on face are extracted for further processing. The optimal face representation using nearest neighbor classifier allows matching the localized facial features and keypoints efficiently by searching the correspondence of keypoints .

REFERENCES

1. Jongju Shin and Daijin Kim Hybrid Approach for Facial Feature Detection and Tracking under Occlusion IEEE SIGNAL PROCESSING LETTERS, VOL. 21, NO. 12, DECEMBER 2014
2. Jun Luo1, Yong Ma2, Erina Takikawa2, Shihong Lao2, Masato Kawade2, Bao-Liang Lu1 PERSON-SPECIFIC SIFT FEATURES FOR FACE RECOGNITION 1424407281/ 07/$20.00 ©2007 IEEE , ICASSP 2007
3. J.Lim, Y. Kim J. Paik" Comparative Analysis of Wavelet-Based Scale-Invariant Feature Extraction Using Different Wavelet Bases" International Journal of Signal Processing, Image Processing and Pattern Recognition Vol. 2, No. 4, December, 2009
4. T. S. Lee, "Image representation using 2D Gabor wavelets,"PAMI, IEEE Trans. on, vol. 18, pp. 959-971, 1996.

5. Shu Liao, Wei Fan, Albert C. S. Chung and Dit-Yan Yeung, "Facial Expression Recognition Using Advanced Local Binary Patterns, Tsallis Entropies And Global Appearance Features", IEEE International Conference on Image Processing, pp. 665-668, 2006
6. M. Jones and P. Viola, "Face Recognition Using Boosted Local Features", IEEE International Conference on Computer Vision,2003.
7. Bruce A. Draper, Kyungim Baek, Marian Stewart Bartlett, J. Ross BeveRidge, "Recognizing Face with PCA and ICA", Computer Vision and Image Understanding 91, pp. 115–137, 2003
8. G. Shakhnarovich and B. Moghaddam. Facerecognition in subspaces. In S. Li and A. Jain, editors, Handbook of Face Recognition, pages 141–168. Springer Verlag, 2004.
9. L. Wiskott, J. Fellous, N. Kruger, and C. Malsburg. Face recognition by elastic bunch graph matching. IEEE Transactions on Pattern Analysis and Machine Intelligence, 19:775–779, 1997.
10. J. Bigun. Retinal vision applied to facial features detection and face authentication. Pattern Recognition Letters, 23(4):463–475, 1997.
11. G. Zhang, X. Huang, S. Li, Y. Wang, and X. Wu. Boosting local binary pattern (lbp)-based face recognition. In L. 3338, editor, SINOBIOMETRICS 2004, pages 179–186. Springer Verlag, 2004.
12. G. Heusch, Y. Rodriguez, and S. Marcel. Local binary patterns as an image preprocessing for face authentication. IDIAP-RR 76, IDIAP, 2005
13. Dakshina Ranjan Kisku1, Phalguni Gupta2, Jamuna Kanta Sing3 Face Recognition using SIFT Descriptor under Multiple Paradigms of Graph Similarity Constraints1 International Journal of Multimedia and Ubiquitous Engineering Vol. 5, No. 4, October, 2010
14. Y. Freund and R. Schapire, A short introduction to boosting, Journal of Japanese Society for Artificial Intelligence, 14(5):771-780, September, 1999.