ISSN No: 2348-4845
International Journal & Magazine of Engineering, Technology, Management and Research
A Peer Reviewed Open Access International Journal

Efficient Heterogeneous Face Recognition using Scale Invariant Feature Transform



Shaik Imam Vali
M.Tech(DSCE),
St Johns College of Engineering & Technology, Yerrakota.

B.Venkatesh, M.Tech
Assistant Professor,
St Johns College of Engineering & Technology, Yerrakota.

M.Srinivasulu, M.Tech
Associate Professor,
St Johns College of Engineering & Technology, Yerrakota.

Abstract:

The Scale Invariant Feature Transform (SIFT) is an algorithm used to detect and describe scale-, translation- and rotation-invariant local features in images. The original SIFT algorithm has been successfully applied in general object detection and recognition tasks, panorama stitching and others. One of its more recent uses also includes face recognition, where it was shown to deliver encouraging results. SIFT based face recognition techniques found in the literature rely heavily on the so-called key point detector, which locates interest points in the given image that are ultimately used to compute the SIFT descriptors. While these descriptors are known to be among others (partially) invariant to illumination changes, the key point detector is not. Since varying illumination is one of the main issues affecting the performance of face recognition systems, the key point detector represents the main source of errors in face recognition systems relying on SIFT features. To overcome the presented shortcoming of SIFT-based methods, we present in this paper a novel face recognition technique that computes the SIFT descriptors at predefined (fixed) locations learned during the training stage. By doing so, it eliminates the need for key point detection on the test images and renders our approach more robust to illumination changes than related approaches from the literature. Experiments, performed on the Extended Yale B face database, show that the proposed technique compares favorably with several popular techniques from the literature in terms of performance.

Key words:

SIFT, key point detector, SIFT descriptor, face recognition.

Introduction:

Face recognition is extensively used in a wide range of commercial and law enforcement applications. Over the past years many algorithms have been proposed for facial recognition systems. These algorithms include two basic aspects: holistic, e.g. PCA (Principal Component Analysis) and LDA (Linear Discriminant Analysis), and feature-based, e.g., Gabor- and Scale Invariant Feature Transform-based (or SIFT-based) methods . Holistic approaches use the entire face region for the task of feature extraction and, therefore, avoid difficulties in the detection of specific facial landmarks. Feature-based approaches, on the other hand, extract local features from specific feature points of the face.Scale-Invariant Feature Transform (SIFT) is a known local feature extraction method which detects and describes local features in images. The features extracted by SIFT are invariant to image scale, orientation, change in illumination and substantial range of affine distortion. SIFT feature extraction method was initially developed for object recognition purposes. Lowe proposed to use SIFT features for face recognition in the same way as they were used for object recognition. Many authors used SIFT features in the field of face recognition. However, the capability of SIFT features in face recognition has not been systematically investigated.Though SIFT used for face recognition, issues like comparing similarities between tests and training image of two different persons, face authentication are not discussed. For efficient face recognition, we propose to extract SIFT features frommultiple face training image per person and set a threshold formaximum matching features. Since for a test image, there are many training images, thus image with the maximum matching features will be the recognized image.

ISSN No: 2348-4845
International Journal & Magazine of Engineering, Technology, Management and Research
A Peer Reviewed Open Access International Journal

For efficient heterogeneous face recognition, we evaluated SIFT algorithm on AT&T, YALE and IIT-KANPUR [9], [10], [11] database. Contour matching based face recognition is also experimentally studied. Even though contour matching provides computational simplicity, but gives better resultsonly with small databases.

A. Scale Invariant Feature Transform:

The presented model uses SIFT, developed by David Lowe to extract features. SIFT transforms image data into scale invariant coordinates relative to local features. The features are invariant to image scaling and rotation, and partially invariant to change in illumination. Large numbers of features can be extracted from typical images. SIFT features are extracted from a set of training images and stored in the database for face recognition. A test image is matched by individually comparing each feature from the test image to the existing database. Best match between the extracted features is based on Euclidean distance of their feature vectors. To generate set of image features, SIFT uses four major stages of computation. Four stages include Scale-Space extremaDetection, Key Point Localization, Orientation Assignment and Key Point Descriptor .

Scale Space Extrema Detection stage detects stable features across all possible scale spaces. Image is first convolved with Gaussian filters at different scales, which results into several Gaussian images. Koenderink and Lindeberg shown that under a variety of reasonable assumptions the only possible scalespace kernel is the Gaussian function. Scale space L(x,y,),is produced from the convolution of a variable-scale Gaussian G(x,y,),with an input image I(x,y)is defined by,

$$L(x,y,\sigma) = G(x,y,\sigma) * I(x,y)$$

Where

$$G(x,y,\sigma) = \frac{1}{2\pi\sigma^2} e^{-(x^2+y^2)/2\sigma^2}$$

is the Gaussian Kernelfunction. Then the differences of the adjacent Gaussian images are calculated to generate the Difference of Gaussian images. DoG images D(x,y,),can be computed from the difference of two nearby scales separated by constant multiplication Factor k:

$$D(x,y,\sigma) = \left(G(x,y,k\sigma) - G(x,y,\sigma)\right) * I(x,y)$$
$$D(x,y,\sigma) = L(x,y,k\sigma) - L(x,y,\sigma) \quad (2)$$

a factor of 2 (next octave), and then the process is repeated. After obtaining the DoG images in each octave, SIFT selects the interesting features from the DoG images. In order to detect the local maxima and minima of DoG, each sample point is compared with its eight neighbors in the current image and nine neighbors in the adjacent scales. Every sample point is selected only if it has larger value or smaller value than all its neighbors. Key point Localization is used to determine location and scale. Key points with low contrast and poorly localized features areto be rejected in order to get stable extreme points. Fitting function of the feature point is constructed according to Taylor expansion and is given by,

$$D(x) = D + \frac{\partial D^T}{\partial x} x + \frac{1}{2} x^T \frac{\partial^2}{\partial x^2} x \quad (3)$$

where D and its derivatives are evaluated at the sample point. The location of the extremumX^is determined by taking derivative and equating it to zero. And it can be given as

$$\hat{x} = -\frac{\partial^2 D^{-1}}{\partial x^2} \frac{\partial D}{\partial x} \quad (4)$$

The function value at the extremum helps to reject unstable extrema with low contrast and is given by

$$D(\hat{x}) = D + \frac{1}{2} \frac{\partial D^T}{\partial x} \hat{x} \quad (5)$$

Reject the extremas whose |D(x^)|<0.3 ,but it is not sufficient to reject key points with low contrast. The difference-of-Gaussian function will have a strong response along edges, even if the location along the edge is poorly determined. The principal curvatures of poorly defined peak in the difference-of-Gaussian function computed using Hessian matrix will have large value across the edge but a small value in the perpendicular direction.

Ratio of principal curvatures helps to eliminate poorly defined key points along the edge. Orientation Assignment determines the key point's direction which ensures the feature's rotation invariance. The direction of the key point is calculated by the image information of key point's neighborhood. For each sample L(x,y), the modulusvalue M(x,y)and the orientation (x,y)are computed using pixel differences and are given by:

$$m(x,y) = \sqrt{(L(x+1,y) - L(x-1,y))^2 + (L(x,y+1) - L(x,y-1))^2} \quad (6)$$

$$\theta(x,y) = \tan^{-1}((L(x,y+1) - L(x,y-1))/(L(x+1,y) - L(x-1,y))) \quad (7)$$

ISSN No: 2348-4845
International Journal & Magazine of Engineering, Technology, Management and Research
A Peer Reviewed Open Access International Journal

Once the modulus and angles are computed, the algorithm divides 00 to 3600 into 36 bins, where each contains 100. Then it statistics the modulus located in each bin and histogram of gradients is constructed. Each sample added to the histogram is weighted by its gradient magnitude and a Gaussianweighted circular window with an σ that is 1.5 times that of the scale of the key point. The peak of orientation histogram reflects the dominant direction of the area around the key point.

If there is any other statistic value which is greater than 80% of the highest peak, then it also used to create a key point with that orientation. Hence multiple key points can be created at the same location and scale but with different orientation. Key Point Descriptor is constructed for each key point, bycomputing the gradient magnitude and orientation in a region around the key point location. Key point descriptor is created at each image point of the 16x16 key point neighborhoods.

Each window is divided into four areas, where each area is a window of 4x4 sub regions. For each area, compute the histogram in eight directions using the gradient value. The feature description is calculated by considering the direction descriptions of all subfields. The length of feature descriptor will be 4 x 4 x 8=128 elements.

B. Contour Matching:

Contour matching is simple matching technique used for face recognition. Face recognition using Contour matching is processed into three steps like Image Processing, Contour Generation and Matching Algorithm [14].Image Processing improves the recognition performance. Preprocessing involves are histogram equalization, noiseelimination and normalization. Histogram equalization to enhance the contrast of images and is done by transforming the values in an intensity image.

Noise elimination is to remove noise from the image using Gaussian blurring. Normalization is done to compensate for illumination variations. Contour Generation is to generate contour lines from the image the whole face is treated as a contour map, with the areas of constant gray level brightness (i.e. the plains enclosed by the contour lines. Thus contour lines for a given face can be generated. Matching Algorithm is used for face recognition by matching the contours of two faces.

The contour of test image is matched with the contour of training image using template matching. Matching algorithm has two perform two steps likecalculating matching ratio and fragment removal threshold. Matching ratio H(f,g)is the maximum similarity between test and registered contour. Under template matching technique, input image is slid pixel by pixel across the registered image from top to bottom and from left to right. The pixels f(i,j) and g(k,l)are at position (I,j)and (k,l)of input and registered contour respectively, and gives the measure of horizontal and vertical displacement between thetwo pixels f(I,j),and g(k,l) during the sliding process.

To absorb the slight differences existing in the contour extractedfrom different pictures of the same person, a 5 x 5 window around a black pixel is used. If for a black pixel in the input contour there is another black pixel in the 5 x 5 neighborhoodof the corresponding position on the registered contour, then the pixels are said to be matched. Maximum similarityhi,j(^,^)between two images depends on the horizontal (^) and vertical displacement (^) respectively and is given by an equation

$$h_{i,j}(\hat{\alpha},\beta)^{\Delta} = max \sum_{i,j} h_{i,j}(\alpha,\beta) \qquad (8)$$

If the images are of the same person, then long segments of overlapping contour will be obtained. And for the images of the different persons, if greater matching ratio is obtained indicate greater number of short segments. Stable performance can be obtained only if short segments are eliminated. Final matching ratio may be obtained by

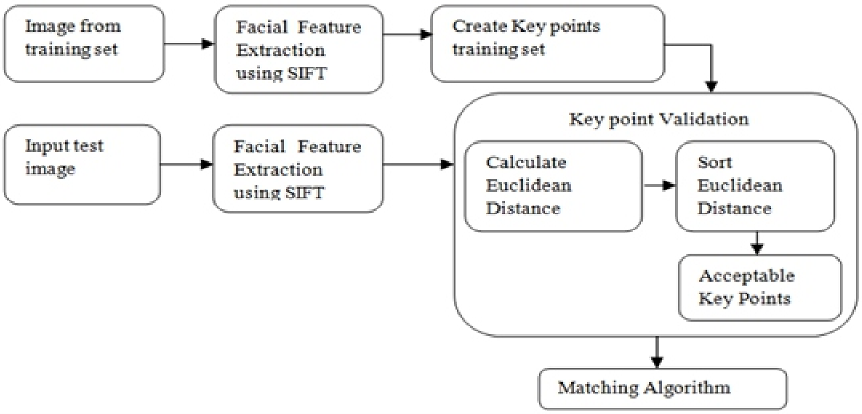
$$H(f.g) = \frac{2}{F+G} \sum h_{i,j}(\hat{\alpha},\hat{\beta}) \qquad (9)$$

Fand Gdenote the number of pixels in the isodensity lines ofthe input and registered pictures respectively.

III. PROPOSED MODEL:

The proposed model for face recognition shown in Fig.1 consists of three phases: Key Points Extraction, Key pointsValidation and Matching Algorithm

ISSN No: 2348-4845
International Journal & Magazine of Engineering, Technology, Management and Research
A Peer Reviewed Open Access International Journal



Key Points Extraction:

In this stage highly distinctive facial features are extracted from the images using SIFT. Key points for all the images in the training dataset as well as for the test image are extracted in this stage.

Extracted key points of all the training images are stored in the key point dataset. Steps involved in the key point extraction will be DoG Image Generation, Local Key Point Detection, Accurate Key Point Location and Eliminating Edge Response.

• Difference-of-Gaussian Images Generation: Location and scale of the key point need to be identified for key pointdetection. DoG proved to be the efficient function for detecting key point location. DoG function is computed from the difference of two nearby scales.

• Local Key Point Detection: Each key point is compared with its eight neighbors in the current image and nine-neighbors from adjacent scales that give local maxima or minima.

• Accurate Key Point Location: From the number of key points, those sensitive to noise means with low contrast will be rejected. This is done by setting up a threshold value for the key point which gives only stable key points.

• Eliminating Edge Response: Key points along the edges are poorly determined and unstable to noise, hence to be-rejected. Poorly defined peaks will be determined by principal of curvatures. Unwanted key points will berejected by setting additional threshold on principal of curvatures.

Key points Validation:

It is the stage which helps to find acceptable key points. Acceptable key points are the highly distinctive key points which are extracted from the image for processing face recognition. Detailed process involved in the Key Point Acceptance and Matching Algorithm is given the Fig.2. There are three steps involved in the process of validation: Calculating Euclidean Distances, Sorting Euclidean Distances and Acceptance of Key Points.

• Calculating Euclidean Distances: Once the key points for the test and training set are extracted, Euclidean distancecan be calculated by the formula given as:

$$ED = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2} \qquad (10)$$

x1,y,1 and x2,y2 are x and y coordinates of the test and training image respectively. Euclidean distances are calculated by comparing every test key point with entire set of key points of the training image and then form Euclidean distance array.

• Sorting Euclidean Distances: Once array of Euclidean distances generated by comparing all the test image keypoints with entire set of training image key points, it will be sorted to find first two closet neighbors.

• Acceptance of Key Points: From the Euclidean distances array, ratio of two closet neighbors is calculated. Thisratio helps to reject the false key points. If the ratio is greater less than 0.8 indicates those are not acceptable keypoints and can be rejected. But if the ratio is less than 0.8 then accept it as the valid key point and increment key point counter which keeps the count of matched key points.

ISSN No: 2348-4845
International Journal & Magazine of Engineering, Technology, Management and Research
A Peer Reviewed Open Access International Journal

Matching Algorithm based on Sum of Absolute Differences of key points: This is the heart of the face recognition model. Key point validation returns entire training set images with number of matched key points. First step of matching algorithm is to sort the training images with maximum count on the basis of the matching threshold Mt. The training images whose maximum count is greater than matching threshold Mt, indicates images which are closely matched. Second step is to determine the perfect match from the set of closely matched training images. To find exact match, Sum of Absolute Differences between key points of the closely matched training images and test image has been calculated. More formally, Itest and Itrainrepresents the test and training image, respectively. Distinctive Features for the test and training image will be represented as:

- $K_T^{I_{test}} = \{k_1^{I_{test}}, k_2^{I_{test}}, k_3^{I_{test}}, \dots \dots k_N^{I_{test}}\}$
- $K_{TN}^{I_{train}} = \{k_1^{I_{train}}, k_2^{I_{train}}, k_3^{I_{train}}, \dots \dots k_M^{I_{train}}\}$

Nand Mindicates number of key points in the test and training images respectively. Key points of the training images sorted on the basis of the MTwill be represented as:

- $K_{ST}^{I_{strainset}} = \{k_{1-M1}^{I_{strain}}, k_{1-M2}^{I_{strain}}, k_{1-M3}^{I_{strain}} \dots \dots k_{1-Mn}^{I_{strain}}\}$

Where M1, M2…. Mnare the key points of the sorted training images. Key points of one of the sorted training images will be represented as:

- $K_{ST}^{I_{strain}} = \{k_1^{I_{strain}}, k_2^{I_{strain}} k_3^{I_{strain}} \dots \dots k_M^{I_{strain}}\}$

Sum of Absolute Differences of key points, SADK between coordinates of the key points of the test image and sorted training images STof size Szis calculated using following expressions

$$D_{diff(x'y')} = \left(\sum_{i=0}^{N-1}\sum_{j=0}^{1}|(I_{test(i,j)})|\right) - \left(\sum_{x=0}^{M-1}\sum_{y=0}^{1}|(I_{strain(x,y)})|\right) \qquad (11)$$

$$Recognized\ Image = \min(SADK_{ST})$$

From array of Sum of Absolute Differences of key points sorted training images, the training image with the lowest difference value would be matched or recognized image.

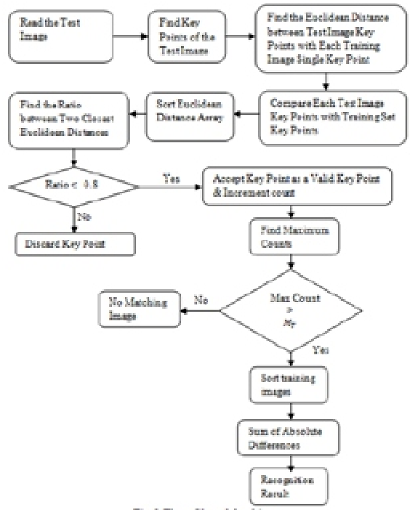


Fig.2 Flow Chart-Matching

EXPERIMENT RESULTS:

For systematic investigation of SIFT algorithm, experiments have been conducted on ATNT, YALE and IIT Kanpur databases, executed on an Intel Core 2 Duo CPU running on 2.26 GHz with 8 GB RAM with windows OS (64 Bit) and OpenCV 1.1(Visual Studio 2008) . With the mentioned hardware response time is 0.4secs for ATNT, YALE and IIT Kanpur database. In every database, there are 20 subjects and each subject had 5 different facial views representing different poses, varying lighting conditions and e xpressions. Each image is digitized and stored as 92 x 112 pixel array. The file is in JPEG format. For conducting the experiment on three different databases, three separate training and test set has been set. Each training set is set up by a ran dom selection of 3 samples for each person which includes 60 images totally and the testing set consist of the remaining.

In the experiment, highly distinctive key p oints are extracted for all the images in the training data s et. Corresponding coordinates of these key points are maintained in the array. Some sample images from different i mages and their corresponding extracted highly distinctiv e key points are shown in the Fig.3.For the test image also highly distinctive features are extracted using SIFT. Euclidean distance for every test key point is calculated using entire training set, to form Euclidean distance array. Euclidean distance array helps to find two closest neighbors and to reject false key points. Fi nal recognition is done by counting the number o f maximum matched key points of a test image with the entire set of training images.

ISSN No: 2348-4845
International Journal & Magazine of Engineering,
Technology, Management and Research
A Peer Reviewed Open Access International Journal

The training images with maximum count greater than threshold are sorted. Sum of Absolute Differences are calculated for the sorted training images. The training image with least SAD will be a recognized image.

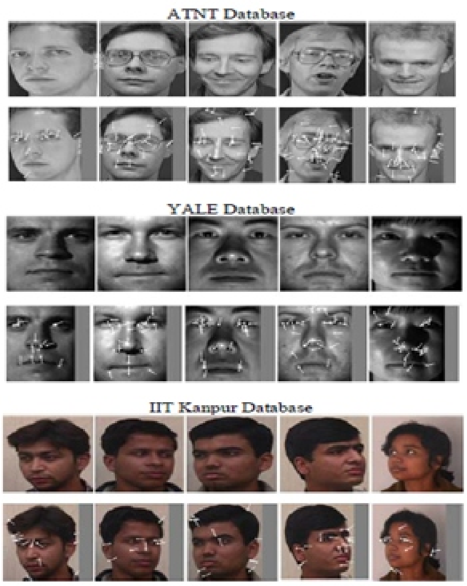


Fig. 3 Sample Images with keypoints

CONCLUSION:

Face recognition using SIFT algorithm and Contour matching has been studied. Face recognition based on Contour matching method is easy to implement due to its simplicity. But in contour matching shape of the contours get affected by titling or panning of a face. Similarly the varying lighting condition or two images with different expression of a same person also give the false contour lines, which in turns lower the recognition rate. Face recognition using Scale Invariant Feature Transform outperforms in many aspects. SIFT provides efficient face recognition results under varying lighting condition, scale, pose and expression. Our proposed efficient SADK matching algorithm of face recognition based on SIFT has given better results in terms of accuracy and efficiency. We experimented with three challenging data bases. Experimental results show that our SIFT based face recognition method is robust to illumination changes(YALE), changes in facial expressions(ATNT) and partial head and partial head movememnts including frontal, +45° and -45° (IIT Kanpur). In our experiment images with +90° and -90° and non ideal images taken by the camera are not included. Though the effect due to these are not examined,but left open for further study.

REFERENCES:

[1]G. Betta, D. Capriglione, C. Liguori, A. Paolillo, "Uncertainty evaluation in face recognition algorithms", IEEE, Pag.621-627 ISBN:9781424479344 ID:3025240 Contributo in Atti di convegno , 2011.

[2]JieZou, QiangJi and George Nagy, "A Comparative Study of Local Matching Approach for Face Recognition",IEEETrans.on Image processing, vol. 16, No. 10,pp.2617-2628,oct 2007

[3]S. M. Zakariya, r. Ali and m. A. Lone, "automatic face recognition Using multi-algorithmic approaches", s. Aluru et al. (eds.): ic3 2011, ccis 168, pp. 501–512, 2011.

[4]Yunfei Jiang and Ping Guo, "Comparative Studies of Feature Extraction Methods with Application to Face Recognition",IEEE,pp.3627-3632,2008 : IEEE International Conference on Systems, Man, and Cybernetics-SMC ,pp.3627-3632,2007DOI: 10.1109/ICSMC.2007.44137091

[5]Bin Jiang, Guo-Sheng Yang, Huan-Long Zhang, " Comparative Study Of Dimension Reduction And Recognition Algorithms Of DCT And 2DPCA" , Proceedings of the Seventh International Conference on Machine Learning and Cybernetics, Kunming, pp. 407-410, July 2008.

[6]SushmaNiketBorade and Dr. R.P.Adgaonkar, "Comparrative Analysis of PCA and LDA", ICBEIA,pp,203-206,2011

[7]Cong Geng and Xudong Jiang, "Face Recognition Using Sift Features", ICIP,pp.3313-3316,2009.

[8]Wang Yunyi, Huang Chunqing and QiuXiaobin, "Multiple Facial Instance for Face Recognition based on SIFT Features", IEEE conference on Mechatronics and Automation,pp.2442-2446,2009.

[9]http://www.cl.cam.ac.uk/Research/DTG/attarchive:pub/dataatt_faces.zip

[10]http://vision.ucsd.edu/datasets/yale_face_data/set_original/yalefaces.zip

[11]http://vis-www.cs.umass.edu/~vidit/IndianFaceDatabase

[12]D. G. Lowe, "distinctive image features from scale-invariant Keypoints", international journal of computer vision, vol. 2, no. 60,Pp. 91-110, 2004.

[13]Y ouliang Yang, Weili Liu and Lan Zhang, "Study on Improved Scale Invariant Feature Transform matching algorithm", PACCS,pp.398-401,2010

[14]S. T. Gandhe, K. T. Talele, and A.G.Keskar, "Face Recognition Using Contour Matching", IAENG International Journal of Computer Science,2008.