International Journal of Advance Research in Science and Engineering
Vol. No.5, Issue No. 02, February 2016
www.ijarse.com
IJARSE
ISSN 2319 - 8354

PHOTOGRAPH RETRIEVAL BASED ON FACE SKETCH USING SIFT WITH PCA

Tayyaba Hashmi

ME Information Technology, Shah & Anchor Kutchhi Engineering College

University of Mumbai, (India)

ABSTRACT

The problem of matching a sketch to a gallery of mug shot images is addressed. Previous research in sketch matching only offered solutions to matching highly accurate sketches that were drawn while looking at the subject (viewed sketches). To identify sketches much efficient algorithm is presented here. Both sketches and photos are considered for extracting feature descriptors using SIFT (Scale Invariant Feature Transform). Proposed method can be used to match a few sketches against a mug shot gallery containing several images. Its matching performance can be improved using race and gender information to reduce the target gallery size. The proposed project will lead to state-of-the-art accuracy when matching viewed sketches. The proposed project represent both sketches and photos using SIFT feature descriptors and multiscale local binary patterns(MLBP).

***Keywords:* Photo Retrieval, Graphical Method, SIFT Feature Descriptors, MLBP, PCA.**

I INTRODUCTION

Image retrieval in sketch matching is a challenging task. Sketch matching is one of the most tedious task when carried out manually. Performing manual examination of facial images or videos for a match with huge database of mug shots is very difficult. To identify sketches, we present a framework called local feature-based discriminant analysis (LFDA)[1]. For example, when a crime is committed, always the major clue leading to sketch of the criminal is verbal description of the criminals by witnesses. Police departments then release drawings from the descriptions in hopes of identifying the criminals. Identifying criminal normally consists of witnesses looking through books of true mugshot photographs. However, some departments, utilize a computer to assist in the search. The identification officer inputs the height, weight, and race of the suspect into a program, which searches the database and outputs a long list of possible candidates. The officer then calls up the image of each person on this list and compares it with the sketch obtained from witness descriptions. The images are obtained from databases of real photographs maintained by police and other governmental agencies. An automated comparison process would save many person-hours by eliminating the officer's manual search, and it might reduce the dependence on the public. The problem we address is amongst the most difficult instances of the face identification problem. It takes as input a police artist sketch of a suspect and matches it to a database of photographs. The sketch is first transformed to an actual digitized photograph. The transformed sketch is then compared to the database images to obtain a set of candidates. This paper presents a method for extracting distinctive invariant features from images that can be used to perform reliable matching between different views of an object or scene[2]. This approach transforms an image into a small collection of local feature vectors. The first stage identifies key locations in scale space by looking for locations that are maxima or minima of a difference-of-Gaussian function. Each point is used to generate a feature vector that describes the local image region. The resulting feature vectors are called Modified SIFT keys.

II REVIEW OF LITERATURE

2.1. The common steps utilized by most face recognition algorithms

The challenges in designing automated face sketch recognition algorithms are numerous. Finding the probe sketch match in the mugshot database, such challenges manifest in the following stages performed by most face sketch recognition algorithms:

(i) Preprocessing,

(ii) Feature description,

(iii) Feature extraction, and

(iv)Feature based matching.

This section provides an overview of each of the above mentioned stages in automated face sketch recognition, the main focus will be on the face representation and feature extraction stages. This is because to improve the search time, face sketch recognition has generally relied on improvements in these two stages to increase recognition speed with accuracies.

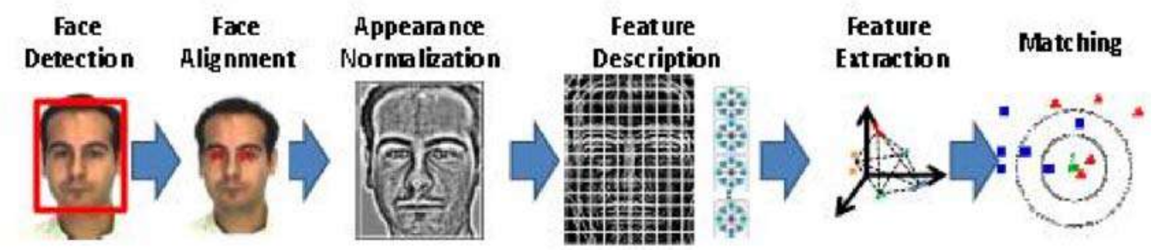

Fig 1: The common steps utilized by most face recognition algorithms [3].

2.2. Preprocessing

In this step often face from sketch and photo is identified and aligned. Digital images in the gallery may also be noisy and of suboptimal quality because of printing and scanning of images. Some algorithm enhances the forensic sketch and photograph for better recognition like MCWLD [4], Multiscale MRF Model.

2.3. Feature description

The feature description is to find a specific representation of the image that can highlight relevant information. This representation can be found by maximizing a criterion or can be a pre-defined representation. Usually, a face image is represented by a high dimensional vector containing pixel values (holistic representation) or a set of vectors where each vector summarizes the underlying content of a local region by using a high level transformation (local representation)[5].

2.4. Global Feature Extraction

Feature extraction is the most widely used feature description technique in appearance based face recognition methods. Despite its poor generalization abilities in unconstrained scenarios, it is being used for the main reason that any local extraction technique is a form of information reduction in that it typically finds a transformation that describes a large data by few numbers. Since from a strict general object recognition standpoint, face is one class of objects, and thus discriminating within this class puts very high demands in finding subtle details of an image that discriminates among different faces.

2.5. Local Feature Extraction

The goal of local feature extraction thus becomes to represent these local regions effectively and comprehensively. Here we review the most commonly used local feature extraction techniques in face recognition namely the Gabor wavelet transform based features , discrete cosine transform DCT-based features and more recently proposed Local binary pattern LBP features.

III REPORT ON PRESENT INVESTIGATION

3.1. Problem

The problem is amongst the most difficult instances of the face identification problem. It takes as input a police artist sketch of a suspect and matches it to a database of photographs. The sketch is first transformed to an actual digitized photograph. The transformed sketch is then compared to the mugshot database images to obtain a set of candidates. Also if we Consider every important places is having a camera at the entrance door, which captures the image of everyone entering the place. If some crime happens at the place, and someone sees the criminal, he and the photos initially captured can act as eye witness and mug shot photos respectively, can be used to catch the criminal using forensic sketch matching. The Proposed system will accurately matches forensic sketches with their corresponding photo images using feature-based approach. User will feed forensic sketch to the system to find a suitable match for the respective image found in the database. The feature descriptors of both the sketch and photo images are computed and matched for equality, once a match is found the user is alerted and the corresponding suspect's information is extracted from the database and displayed to the user. Large numbers of features can be extracted from typical images using scale invariant feature transform (SIFT) algorithm [5]. These features are highly distinctive, which allows a single feature to be correctly matched with high probability against a large database of features.

International Journal of Advance Research in Science and Engineering
Vol. No.5, Issue No. 02, February 2016
www.ijarse.com
IJARSE
ISSN 2319 - 8354

3.2 Proposed Block Diagram

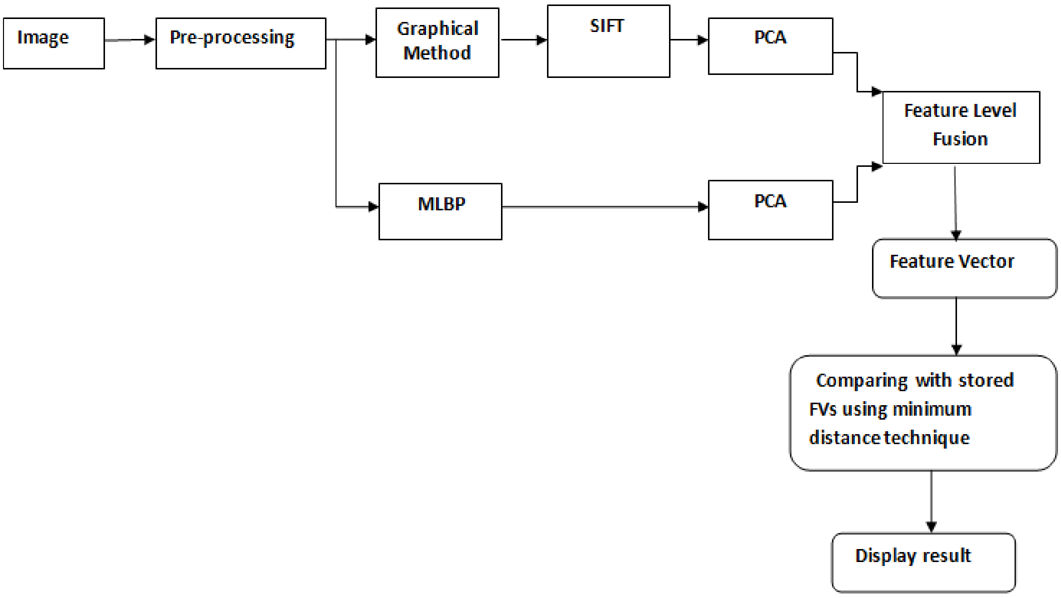

Fig 6.System Block Diagram

3.2.1. Pre-processing

Pre-processing is an important and diverse set of image preparation programs that act to offset problems with the band data and recalculate DN values that minimize these problems. Among the programs that optimize these values are atmospheric correction sun illumination geometry; surface-induced geometric distortions; spacecraft velocity and attitude variations (roll, pitch, and yaw); effects of Earth rotation, elevation, curvature (including skew effects), abnormalities of instrument performance (irregularities of detector response and scan mode such as variations in mirror oscillations); loss of specific scan lines (requires destriping), andothers. Once performed on the raw data, these adjustments require appropriate radiometric and geometric corrections. Resampling is one approach commonly used to produce better estimates of the DN values forindividual pixels. After the various geometric corrections and translations have been applied,the net effect is that the resulting redistribution of pixels involves their spatial displacementsto new, more accurate relative positions. However, the radiometric values of the displacedpixels no longer represent the real world values that would be obtained if this new pixel arraycould be resensed by the scanner.The particular mixture of surface objects or materials in the original pixel has changedsomewhat (depending on pixel size, number of classes and their proportions falling within thepixel, extent of continuation of these features in neighbouring pixels [a pond may fall withinone or just a few pixels; a forest can spread over many contiguous pixels]). In simple words,the corrections have led to a pixel that at the time of sampling covered ground A beingshifted to a position that have A values but should if properly located represent ground B.

3.2.2. Graphical Method

In this method we select pixels on which we apply SIFT. A pixel,or picture element is a physical point in a raster image, or the smallest addressable element in an all points addressable display device; so it is the smallest controllable element of a picture represented on the screen. The address of a pixel corresponds to its physical coordinates. LCD pixels are manufactured in a two-dimensional grid, and are often represented using dots or squares, but CRT pixels correspond to their timing mechanisms and sweep rates. Each pixel is a sample of an original image; more samples typically provide more accurate representations of the original. The intensity of each pixel is variable. In color image systems, a color is typically represented by three or four component intensities such as red, green, and blue, or cyan, magenta, yellow, and black. In some contexts, the term pixel is used to refer to a single scalar element of a multicomponent representation, while in others the term may refer to the entire set of such component intensities for a spatial position. In color systems that use chroma subsampling, the multi-component concept of a pixel can become difficult to apply, since the intensity measures for the different color components correspond to different spatial areas in such a representation.

3.2.3. SIFT

SIFT is a mathematical algorithm for extracting interest point features from images that can be used to perform reliable matching between different views of objects[6].For any object in an image, interesting points on the object can be extracted to provide a "feature description" of the object. This description, extracted from a training image, can then be used to identify the object when attempting to locate the object in a test image containing many other objects. To perform reliable recognition, it is important that the features extracted from the training image be detectable even under changes in image scale, noise and illumination. Such points usually lie on high-contrast regions of the image, such as object edges. Another important characteristic of these features is that the relative positions between them in the original scene shouldn't change from one image to another. For example, if only the four corners of a door were used as features, they would work regardless of the door's position; but if points in the frame were also used, the recognition would fail if the door is opened or closed. Similarly, features located in articulated or flexible objects would typically not work if any change in their internal geometry happens between two images in the set being processed. However, in practice SIFT detects and uses a much larger number of features from the images, which reduces the contribution of the errors caused by these local variations in the average error of all feature matching errors. SIFT can robustly identify objects even among clutter and under partial occlusion, because the SIFT feature descriptor is invariant to uniform scaling, orientation, and partially invariant to affine distortion and illumination changes. This section summarizes Lowe's object recognition method and mentions a few competing techniques available for object recognition under clutter and partial occlusion.

International Journal of Advance Research in Science and Engineering
Vol. No.5, Issue No. 02, February 2016
www.ijarse.com
IJARSE
ISSN 2319 - 8354

3.2.4. MLBP

Multiscale Local binary patterns (MLBP) is a type of feature used for classification in computer vision. It has since been found to be a powerful feature for texture classification; it has further been determined that when LBP is combined with theHistogram of orientedgradients (HOG) descriptor, it improves the detection performance considerably on some datasets. In the LBP histogram, an image pair is first split into sub-regions. The similarity score ofeach local LBP histogram pair is measured using the similarity function. The similarity scoresare then concatenated to form an input feature vector for feature selection process.The LBP operator was originally designed for texture description[7]. Overcompletefeatures can be provided by shifting and scaling the local regions. In general, thetotal sample size of inter-person pairs is larger than that of intra-person pairs. This will giverise to a bias for feature selection.

3.2.5. PCA

The PCA method has been extensively applied for the task of face recognition[8].Principal Components Analysis (PCA) is a dimensionality reduction technique used extensively in Remote Sensing studies (e.g. in change detection studies, image enhancement tasks and more). The PCA involves a mathematical formula that transforms a number of correlated variables into a number of uncorrelated variables called principal components[9]. PCA is in fact a linear transformation applied on (usually) highly correlated multidimensional (e.g. multispectral) data. The input dimensions are transformed in a new coordinate system in which the produced dimensions (called principal components) contain, in decreasing order, the greatest variance related with unchanged landscape features. PCA is a standard decorrelation technique which projects the input signal into a space where features have no correlation with each other. It is a common technique forsignal representation or signal compression because PCA can reduce the dimensionality bykeeping the space which encapsulates the maximum amount of signal variation and throwingout dimensions with small variation which are regarded as noise.

3.2.6. Feature level Fusion

Fusion at the feature level involves the integration of feature sets corresponding to multiple modalities. Since the feature set contains richer information about the raw biometric data than the match score or the final decision, integration at this level is expected to provide better recognition results. Some of the identifying biometric feature now under investigation for potential use include hand geometry, blood vessel patterns in the retina or hand, fingerprints, voice-prints, and handwritten signature dynamics[10]. However, fusion at this level is difficult to achieve in practice because of the following reasons: (i) the feature sets of multiple modalities may be incompatible (e.g., minutiae set of fingerprints and eigen-coefficients of face); (ii) the relationship between the feature spaces of different biometric systems may not be known; and (iii) concatenating two feature vectors may result in a feature vector with very large dimensionality leading to the 'curse of dimensionality' problem. We describe a technique that utilizes the fused feature vectors of face and hand geometry in order to improve the performance of a multimodal biometric system.

3.2.7. Feature Vector

In pattern recognition and machine learning, a feature vector is an n-dimensional vector of numerical features that represent some object. Many algorithms in machine learning require a numerical representation of objects, since such representations facilitate processing and statistical analysis. When representing images, the feature values might correspond to the pixels of an image, when representing texts perhaps term occurrence frequencies. Feature vectors are equivalent to the vectors of explanatory variables used in statistical procedures such as linear regression. Feature vectors are often combined with weights using a dot product in order to construct a linear predictor function that is used to determine a score for making a prediction.The vector space associated with these vectors is often called the feature space. In order to reduce the dimensionality of the feature space, a number of dimensionality reduction techniques can be employed.

IV RESULTS AND DISCUSSIONS

In order to find photo based on sketch through this Modified- SIFT Algorithm, sketches and photographs pair is taken from the CUHK data set. An example of the sketch are shown in the Fig. 7

Fig 7. Sketch to Photo

Following Graph and table shows the accuracy of proposed method and comparative accuracy using MLBP and graphical method with LFDA and without LFDA.

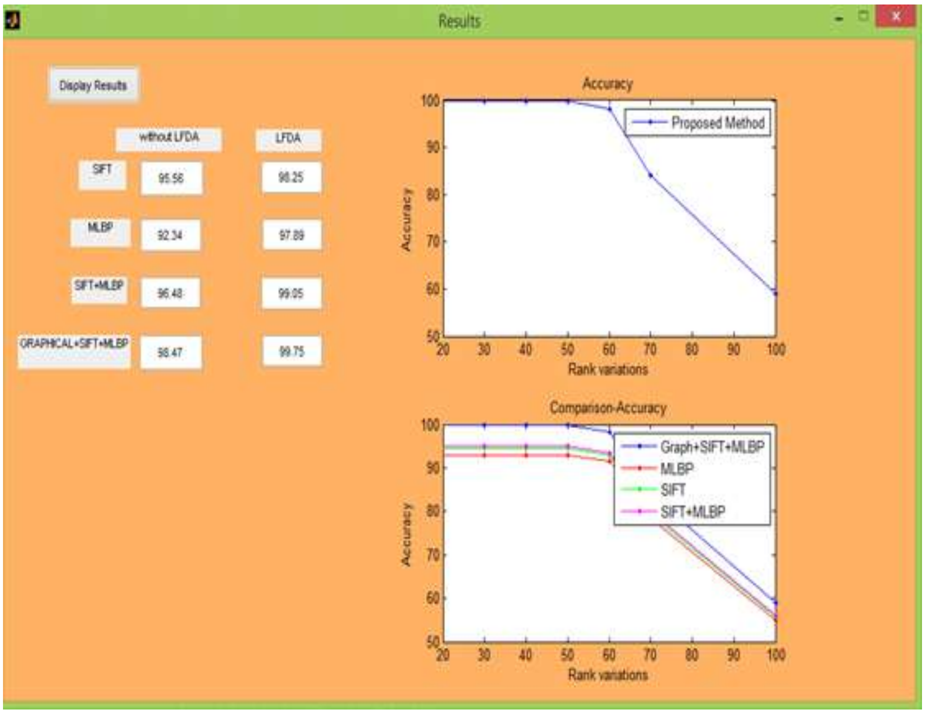

Fig 8. Accuracy of the Proposed method

Following Graph shows the accuracy of proposed method using rank method. It is shown that the accuracies differ depending upon the rank.

Rank K -An accuracy measurement where,

k =the number of returned matches before achieving 100% recognition rate

for example: rank-50 i.e. whether or not the true subject is present within the top-50 images that were or top-50 retrieved images.

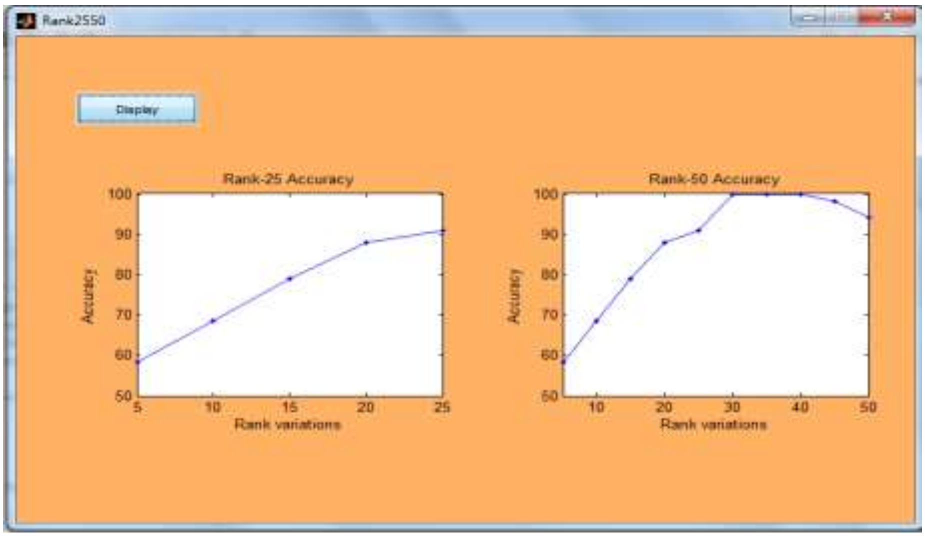

Fig 9. Accuracy based on rank

V FUTURE SCOPE

The studies presented in this thesis offers into some new research challenges for using both SIFT and MLBP. In this approach some better method can be implemented for short-listing. Also the application of PCA in this offers the most avenues for future research. The release of the datasetusing online editing software will hopefully prompt other researchers to explore orthogonal ideas on how to solve the problem.

VI CONCLUSIONS

The proposed Modified-SIFT method is extremely quick in analyzing the Database. Upon close observation it has been found that it is almost 30 times faster than the contemporary software such as SIFT. The percentage of accuracy has been traded off in lieu of speed. To achieve the speed, scrutinizing method of relevant feature has been simplified. Our proposed Modified-SIFT can be used exclusively or in complement with other method such as PCA. If used in complementary form speed and accuracy can be delicately optimized. Attendant advantages of Modified-SIFT include use of PCA. Principal component analysis (PCA) is one of the statistical techniques frequently used in signal processing to the data dimension reduction or to the data decorrelation. PCA takes advantage of eigenvectors properties for determination of selected object orientation. This method provides a powerful tool for data analysis and pattern recognition which is often used in signal and image processing as a technique for data compression, data dimension reduction or their decorrelation as well. There are various algorithms based on multivariate analysis or neural networks that can perform PCA on a given data set. We introduce PCA as a possible tool in image enhancement and analysis.

REFERENCES

[1] Brendan F. Klare, Zhifeng Li, Anil K. Jain-"Matching Forensic Sketches to Mug Shot Photos", IEEE Trans. Pattern Analysis and Machine Intelligence, VOL. 33, NO. 3, MARCH 2011.

[2] D. Lowe, "Distinctive Image Features from Scale-Invariant Keypoints," Int'l J. Computer Vision, vol. 60, no. 2, pp. 91-110, 2004.

[3] Heterogeneous Face Recognition By Brendan F. Klare A Dissertation Submitted to Michigan State Universityin partial fulfillment of the requirements for the degree of Doctor of Philosophy Computer Science and Engineering.

[4] Himanshu S. Bhatt, Student Member, IEEE,Memetically Optimized MCWLD for Matching Sketches with Digital Face Images SUBMITTED TO IEEE TRANSACTIONS ON IFS.

[5] M. SaquibSarfraz 2Olaf HellwichFeature Extraction and Representation for Face Recognition 2Computer Vision and Remote Sensing, Berlin University of Technology Sekr. FR 3-1, Franklin str. 28/29, 10587, Berlin German.

[6] Morteza Zahedi, Sahar Yousefi. GENDER RECOGNITION BASED ON SIFT FEATURES, School of Information Technology and Computer Engineering, Shahrood University of Technology, Shahrood, Iran ,International Journal of Artificial Intelligence & Applications (IJAIA), Vol.2, No.3, July 2011.

[7] T. Ahonen, A. Hadid, and M. Pietikainen. Face description with local binary patterns: Application to face recognition. IEEE Trans. Pattern Analysis & Machine Intelligence, 28(12):2037{2041, Dec. 2006}.

[8] Gottumukkal, [R. &Asari, V. K. (2004). An improved face recognition technique based on modular PCA approach, Pattern Recognition Letter, vol. 25, no. 4, pp. 429– 436.

[9] D. Padmini, Different Levels of Image Fusion Techniques in Remote Sensing Applications and Image Classification, Assistant Professor, Department of Computer Science, Sri Sankara Arts and Science College, Kanchipuram, Enathur, India, International Journal of Advanced Research in Computer Science and Software Engineering, Volume 4, Issue 8, August 2014.

[10] Daugman, J. (1993). High confidence visual recognition of persons by a test of statistical independence. IEEE Transactions on Pattern Analysis and Machine Intelligence, 151148–1161.