

# Facial Recognition using CCTV images based on Rough Set Theory and Scale Invariant Feature Transform

Geetha Mary A., Akhilesh Saklecha  
VIT University Vellore India

## ABSTRACT

Facial recognition is a highly sophisticated tool utilized by current governmental agencies to identify criminals notorious for fleeing the crime scene. The use of close circuit monitoring for automatic identification burgeons a new field of study. This paper presents some techniques that can be employed to recognize individuals on the basis of CCTV footage. Scale Invariant Feature Transform is applied on the existing database of criminal's images. Finally, an efficient matching is done through Rough Set Theory. If a match is found the concerned government authorities can be alerted using an android application.

**Key words:** Facial Recognition, SIFT, Rough Set, CCTV images, reduct, core.

## INTRODUCTION

The ratio of egregious criminals to intelligent law agencies is constantly increasing. The law enforcing agencies are often found desperate for information about anti-social elements, and by the time any substantial clue comes up, the damage is already done. This brings us to the immediate need for the use of technological advancements. One of the inventions that can greatly assist the governmental bodies is close circuit television. Using close circuit television cameras the facial images of people crossing the camera can be captured and then it can be identified if it matches any of the images in the database.

For robust facial identification, features need to be generated from the images that are consistent enough for varying orientation and scale. Scale Invariant Feature Transform [1] is a highly descriptive technique that produces feature vectors for an image which are scale and illumination independent. Earlier techniques used for facial recognition can be classified broadly into two categories: Holistic and feature based. Holistic techniques e.g.: PCA (Principal Component Analysis[2]) and LDA(Linear Discriminant Analysis[2]) use the entire facial region for feature extraction and hence face no difficulties in detecting special landmarks, but are highly time consuming. Feature based (Gabor and SIFT based [3] [4] [5]), select local features from specific regions of face.

For matching of the features, a widely employed technique is least square distance calculation. The conjoined problem with least square distance calculation is that it is too time consuming and is very inefficient if the dataset is too large. This can be rectified by using Rough Set Theory [6] to reduce the number of attributes to compare and using a rough set based comparison technique.

The paper presents the use of CCTV images to recognize the faces of criminals, SIFT will be used to generate feature vectors for the criminal dataset that will be initially acquired from police records. Rough set theory will be applied on this dataset to gain knowledge about the core and reduct of the dataset. The image that is grabbed by the CCTV cameras will be

transformed by SIFT to generate feature vectors. Matching of the test image features with the ones stored in dataset will be done by using core and reduct information, if a match is found, the corresponding government agency can be alerted.

The paper will be organized in three sections. The next section will preview terminologies of rough set theory needed for the paper, following which will be SIFT's brief introduction. The last section will provide the algorithm for comparing the feature vectors of test image and the stored image.

## OVERVIEW OF THE CONCEPTS USED

### ROUGH SET THEORY

An information system can be defined as  $\mathbf{A} = (U, A)$

Where  $U$ - non-empty, finite set called universe  $A$ - non-empty finite set of attributes i.e  $a: U \rightarrow V_a$  for  $a \in A$  where  $V_a$  is value set of  $a$ . If the information system is of the form  $\mathbf{A} = (U, A \cup \{d\})$ , where 'd' is the decision attribute, elements of  $A$  are called conditions. Any subset  $B$  of  $A$  determines a binary relation  $I(B)$  on  $U$ , *indiscernibility relation*, and is defined as follows:

$$xI(B)y \text{ if and only if } a(x) = a(y) \text{ for every } a \in A,$$

where  $a(x)$  denotes the value of attribute  $a$  for element  $x$ .

Let  $B$  be a subset of  $A$  and let ' $a$ ' belong to  $B$ .

- We say that  $a$  is *dispensable* in  $B$  if  $I(B) = I(B - \{a\})$ ; otherwise  $a$  is *indispensable* in  $B$ .
- Set  $B$  is *independent* if all its attributes are indispensable.
- Subset  $B'$  of  $B$  is a *reduct* of  $B$  if  $B'$  is independent and  $I(B') = I(B)$ .

Let  $B$  be a subset of  $A$ . The *core* of  $B$  is the set off all indispensable attributes of  $B$ .

$$Core(B) = \cap Reduct(B) \dots eq(1)$$

### SCALE INVARIANT FEATURE TRANSFORM

Scale Invariant Feature Transform was introduced by David Lowe. The transform generates features that are scale and orientation invariant. There are six steps in finding the features. The first step is the construction of a scale space in which stepwise blurred images are generated for a level and then the image is halved and again blurring is done. This is done generally till four octaves. For the next step, two consecutive images from a single octave are chosen and subtracted, this produces a difference of Gaussian images (DoG). The third step is to find the maxima and minima within the DoG images, using the location of maxima minima, subpixel maxima and minima are generated for clear identification of keypoints. Keypoints generated in the previous step contains some redundant and superficial keypoints. To remove these, two concepts are applied. The keypoints that correspond to low contrast features, are removed using a contrast filter. For further improving the keypoints, the keypoints that lie on a flat region or an edge are discarded, since corner region produces good keypoints. After finalizing the keypoints, an orientation is assigned to each keypoint. Gradient direction around each keypoint is calculated, and using a histogram, most prominent direction is chosen as the orientation of keypoint. The final step is calculating the SIFT

features. A 16X16 window is taken around the keypoint and is broken into sixteen 4X4 windows. Gradient magnitude and orientation are calculated for each window Using a Gaussian weighing function, according to the distance of pixels from the keypoint, their magnitude and orientation is taken and constructed into a 128 dimensional vector (16 pixels into 8 bins histogram for finding the magnitude, for every 4X4 window:  $4 \times 4 \times 8 = 128$ ).

## PROPOSED METHOD

### I. FLOW DIAGRAM

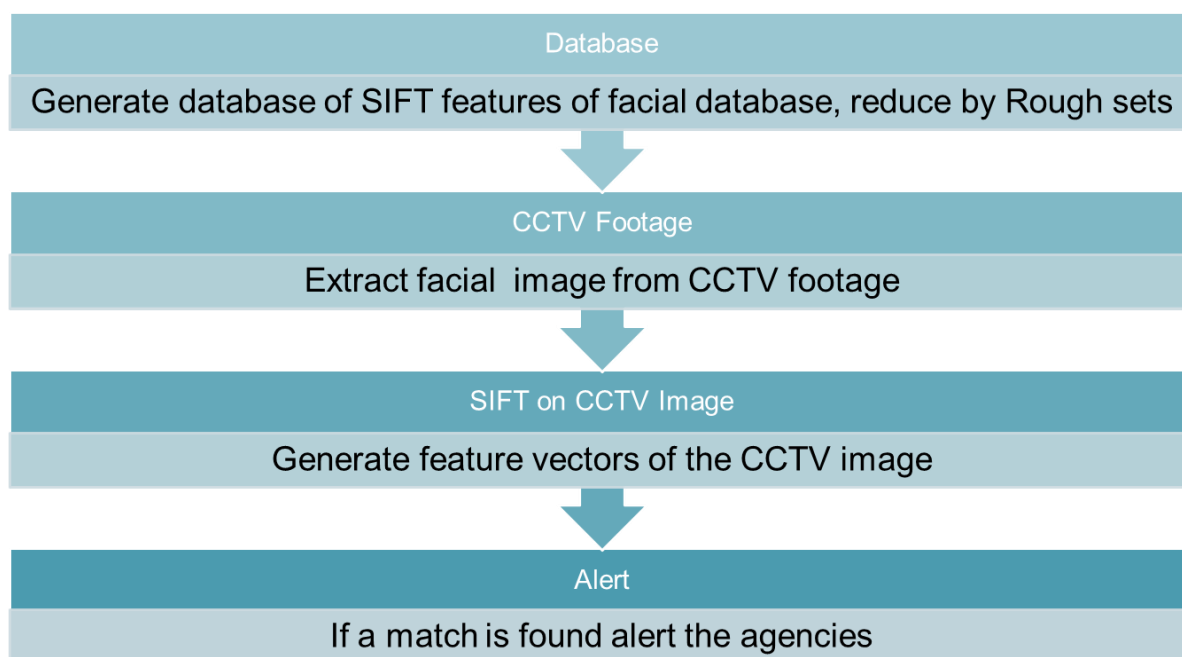


Figure 1 Flow Diagram

### II. CALCULATION OF SIFT FEATURES OF IMAGES IN DATASET

On a database of criminal images (which can be obtained from the government records) scale invariant feature transform will be applied in the first step of the process, and a number of distinct and image exclusive features will be generated. These features will serve as base for matching the facial image that will be obtained later by CCTV cameras. This paper utilizes Indian Face Database [8] for getting facial images.

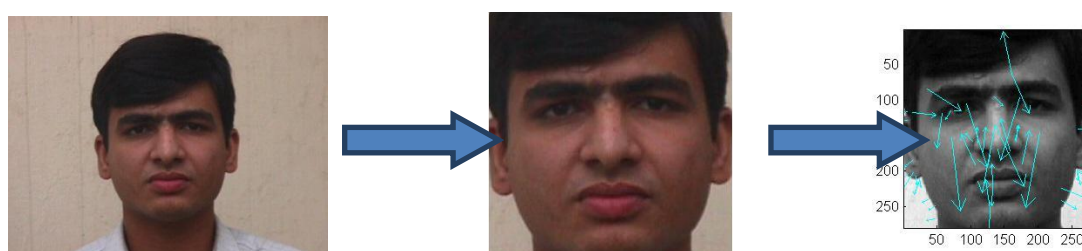


Figure 2 Calculation Of SIFT Features on image obtained from IIT Kharagpur facial database

### III. REDUCT CALCULATION ON THE FEATURES GENERATED

The features generated by SIFT may contain some extraneous and superficial attributes that will slow the comparison. To eliminate these, and use only the most important features, reducts and core will be calculated for the feature vectors generated. Reducts will be calculated using the concept of discernibility matrix introduced by A. Skowron [9]. It is constructed by following rules:

Let the discernibility matrix be  $M(B)$  where  $B \subseteq A$ , an element of discernibility matrix can be defined as

$$c_{ij} = \{ a \in B : a(x_i) \neq a(x_j) \}$$

Reduct can be obtained from the discernibility matrix:  $B' \subseteq B$  is a reduct if  $B'$  is minimal subset of  $B$  such that :  $B \cap c \neq \emptyset$  for any non-empty entry  $c$  in  $M(B)$ .

Core can be obtained by  $eq(1)$ .

The reduct and core information thus obtained is stored for comparison

### IV. PROCURING IMAGE FROM CCTV

For the identification of the criminal, the image from CCTV cameras is used. The video that is being recorded by CCTV cameras will be analysed frame by frame, by the system and if a facial structure is found in the frame, that facial structure will be cropped out and then will be analysed. The facial structure is identified through MATLAB's toolbox that uses Viola Jones algorithm for identifying facial presence in an image.

The image thus obtained from CCTV camera, will mostly be of a low quality since the cameras that record the images are of minimal quality. Hence it is necessary to enhance the image before using it for feature detection. The enhancement is done by 'adaptive histogram equalization' that sharpens some of the features of facial region, helpful in comparing with the dataset.

When the face is extracted from the frame, Scale Invariant Feature transform is applied on the image to produce feature vectors required for comparison.

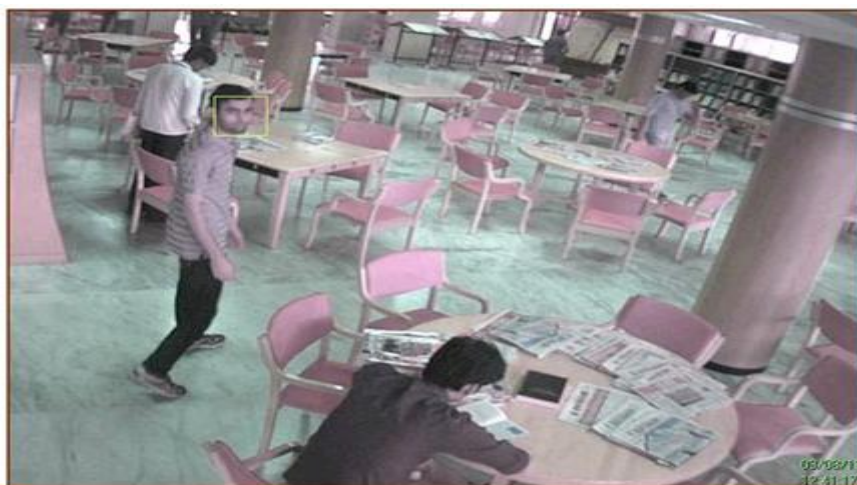


Figure 3 Facial Identification using Viola Jones Algorithm implemented in MATLAB toolbox



**Figure 4 Enhancement using Adaptive Histogram Equalization**

## V. COMPARING THE TEST IMAGE FEATURES WITH FEATURE DATASET

For an efficient application, the matching of test image with dataset image should be done as swiftly with as much accuracy as possible. The necessity of the application is matching should be done quickly since the criminals are known to be fast moving. The least square distance calculation requires  $2n$  operations for comparing two vectors containing 'n' attributes. If the dataset for comparison is too large then the overall time complexity for least square distance calculation will be monumental. In this paper, we followed a new matching technique derived from rough set theory.

The reducts and core information that is stored in the II step will be used to reduce the number of calculations required to find a match. Initially the value of the core of each feature vectors will be compared against the test feature vectors. If the core values are matching, then the reduct information will be used to dismiss the irrelevant feature vectors. If after comparing the reducts, a match is found, it will be alerted to the concerned governmental agency to take further steps upon. The proposed method uses a tree like comparison strategy in which the root is core and down the tree, different reducts are placed such that the overall comparisons required to find a match are reduced. The tree is such arranged through the reduct information obtained from the indiscernibility matrix calculated in second step.

## CONCLUSION

The paper presents an efficient and dynamically applicable method to get hold of criminals specializing in vanishing after committing a crime and staying hidden for days after. If the method comes into application, the criminals' facial data can prove to be an effective tool to find their location. The paper utilizes SIFT for generating feature vectors, Viola Jones

algorithm for facial detection and Rough set theory for matching the feature vectors of the set of images found through close circuit monitoring.

The advantages of the proposed method are : criminals located in any part of the world can be identified using a centralized dataset of criminal images. The manual scrutinizing of the close circuit footage for identification of a criminal will no longer be required. Automatic analysis of the footage eliminates human error that comes along with manual analysis.

The limitations of the proposed method are: an image must be existing in the database for any criminal to be identified, If the criminal's face is not detected by Viola Jones algorithm, a matching will not be done. CCTV footage may sometime be impossible to enhance and hence get a matching.

## REFERENCES

- [1]. Turk M., Pentland A.: Face Recognition Using Eigenfaces, Proc. IEEE Conference CVPR. pp. 586-591, 1991
- [2]. Etemad K., Chellappa R.: Discriminant Analysis for Recognition of Human Face Images, J. of the Opt. Society of America A, Vol. 14, No. 8, pp. 1724–1733, 1997.
- [3]. Wiskott L., Fellous J.M., Kruger N., Malsburg C. von der: Face Recognition by Elastic Bunch Graph Matching, IEEE TPAMI, Vol. 19, No. 7, pp. 775–779, 1997.
- [4]. David G. Lowe. Distinctive image features from scale-invariant key points, International journal of computer vision60, 2 pp.91-110November 2004
- [5]. Janez Križaj, Vitomir Štruc, and Nikola Pavešić. Adaptation of SIFT features for robust face recognition, Proceedings of the 7th international conference on Image Analysis and Recognition - Volume Part I(ICIAR'10) 2010
- [6]. Z. Pawlak, "Rough Sets: Theoretical Aspects of Reasoning about Data." Kluwer Academic Publishers, Netherlands.
- [7]. Ming Hu, Qiang Zhang and Zhiping Wang: Application of Rough Sets to Image Pre-processing for Face Detection. Proceedings of the 2008 IEEE International Conference on Information and Automation June 20 -23, 2008.
- [8]. Vidit Jain, Amitabha Mukherjee. The Indian Face Database. <http://vis-www.cs.umass.edu/~vidit/IndianFaceDatabase/> , 2002.
- [9]. Aly M.: Face Recognition using SIFT Features, CNS/Bi/EE report 186, 2006
- [10]. Ying-Ho Liu, Wen-Hsiung Lin, Fu Chang, and Chin-Chin Lin. 2005. Fast Vector Matching Methods and Their Applications to Handwriting Recognition. In *Proceedings of the Eighth International Conference on Document Analysis and Recognition (ICDAR '05)*. IEEE Computer Society, Washington, DC, USA

- [11]. Jun Luo, Yong Ma, Erina Takikawa, Shihong Lao, Masato Kawade, Bao-Liang Lu  
Person-Specific SIFT Features for Face Recognition, ICASSP 2007
- [12]. Paul Viola and Michael J. Jones. 2004. Robust Real-Time Face Detection. *Int. J. Comput. Vision* 57, 2 (May 2004).