An Intelligent Way of Crime Scene Detection on A Video Surveillance System

B.Naresh Kumar[1] Mr.I.Selvamani[2]
[1]M.E (Final Year) [2]M.E.,(Ph.D) Assistant Professor/ECE
[1,2] AVS Engineering College, Salem-636003.

Abstract---With the increasing demands of visual surveillance systems, vehicle and people identification at a distance have gained more attention for the researchers recently. Extraction of information from images and image sequences are very important for the analysis according to the application. We present an image processing software suite, based on the Mat lab environment, specifically designed to be used as a forensic tool by law enforcement laboratories in the analysis of crime scene videos and images. Here the normal and abnormal actions of a human in any of the under surveillance place is determined and its edges been traced out with various image segmentation techniques. From the preprocessed image database feature vectors are extracted using two methods namely receptive field method and features from Gabor Filter. Features extracted from the database is used as an input of the classifier, which is a feed forward neural network (shortly called FFNN) on a reduced subspace learned by an approach simpler than principal component analysis (shortly called PCA). Once the neural network has been trained, the database will be classified into two categories: normal faces and suspicious faces. Thus suspicious frames are extracted which in turn make ease the process of crime identification.
Index Term: Feed Forward Neural Network, Principal Component Analysis, Video Surveillance System.

I. INTRODUCTION

Video-surveillance systems are one of the main sources of information during investigations, thanks to their wide-spread and increasing presence in our country. However, the adopted closed-circuit devices are often affected by poor quality mainly because of economical and practical problems. Although this fact let us reflect if they can be considered more a deterrent for criminal actions rather than a valid identification system, in many cases also a low quality image can give useful information both during the first phase of the investigation and in courtrooms. As a consequence, the images and sequences coming from video-surveillance systems need to be digitalized in order to be processed by dedicated software to enhance features useful for crime analysis. Generally, this is done either to reduce the different kinds of corruptions that have been introduced in the acquisition, conversion, and storage processes of the data or to overcome the limits of the overall system.

The characteristic problems to deal with are:

– Low resolution of the images, which often implies the need to increase the size of the interesting details
– Lack of contrast
– Different types of noise or disturbances
– Blurring caused by motion or lack of focus
– Jitter or misalignment of lines due to the wear of video cassette recorder (VCR) heads

The human brain has its own limitations in the total number of persons it can accurately remember. A key potential advantage of a computer system is its capability to handle large datasets of face images. This is one of those applications where a crime scene or any abnormal activities is identified from a video database captured by a closed circuit television (shortly called as CCTV) camera for over a long time which went unnoticed. So this application eliminates tremendous work involving in searching for a crime scene manually. In Figure 1.1, the outline of a typical face recognition system is given. This outline carries the characteristics of a typical pattern recognition system.

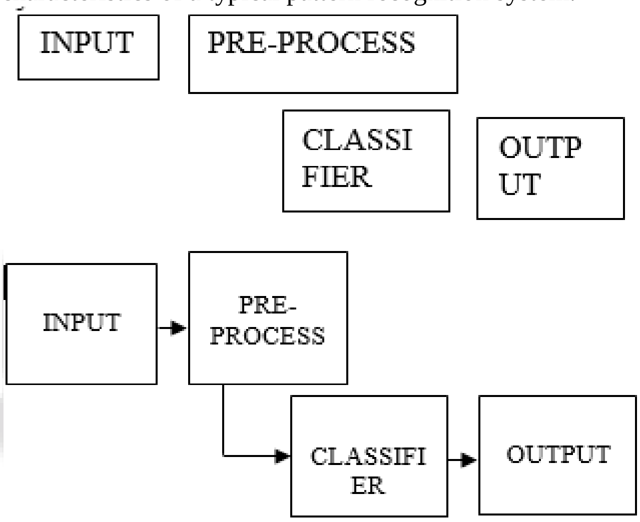

Fig. 1.1: Generic Representation of a Facial Recognition System

A. *Face Recognition*

Face recognition is a patter recognition task performed specifically on faces. It can be described as classifying a face either normal faces or suspicious faces, after comparing it with stored known individuals. It is desirable to have a system that has the ability of learning to recognize unknown faces.

Face recognition involves comparing an image with a database of stored faces in order to identify the individual in the input image. The related task of face detection has direct relevance to recognition because images must be analyzed and faces identified, before they can be recognized.

Face recognition, although a trivial task for the human brain has proved to be extremely difficult to imitate artificially, because although commonalities exist between faces, they can vary considerably in terms of age, skin color, orientation, facial expression and presence of facial furniture such as glasses or facial hair. The problem is further complicated by differing light conditions, image qualities and geometries, as well as the possibility of partial occlusion and disguise. Hence, a robust face recognition system should be capable of classifying a face image as known or unknown. Figure 1.1 outlines the typical block diagram of face recognition system. Normalized Face Image Face Image feature Vector Classified into normal and suspicious faces

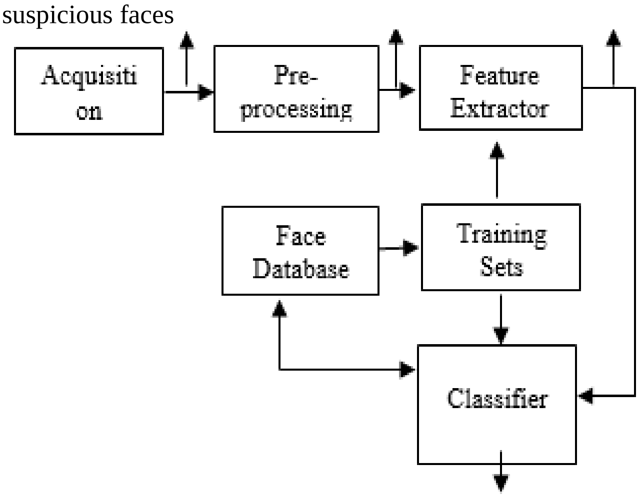

Fig. 1.2: Outline of a Typical Face

B. *Recognition System:*

Face recognition focuses on still images, which can be broadly grouped into image based and feature based approaches. Face recognition is commonly used in applications such as human-machine interfaces, automatic access control systems, which involve comparing an image with a database of stored faces in order to identify the subject in the image.

C. *Objective:*

Paper objectives are listed as follows:

- To design a model for an ideal facial recognition system for crime scene identification.
- To develop a program in MATLAB based on the designed model.
- To create a database set of face images.
- To validate and test the facial recognition system.
- To perform tests for program optimization and accuracy.

II. METHODOLOGY

- Collect a large video database using a CCTV surveillance camera**.**
- Convert it into frames by using software (free studio converter) into desired image format.
- Import those images into mat lab.
- Apply pre-processing techniques such as edge detection.
- Extract features using central point method.
- Extract features using Gabor filter.
- Input the features extracted by two methods individually to a neural network and train the database. Simulate it for different results.
- Testing and validation of the program and technique.
- Creating a user-friendly program in MATLAB from the source code.

Figure 1.3 illustrates the flowchart proposed for the design of the face recognition system, based on the methodology.

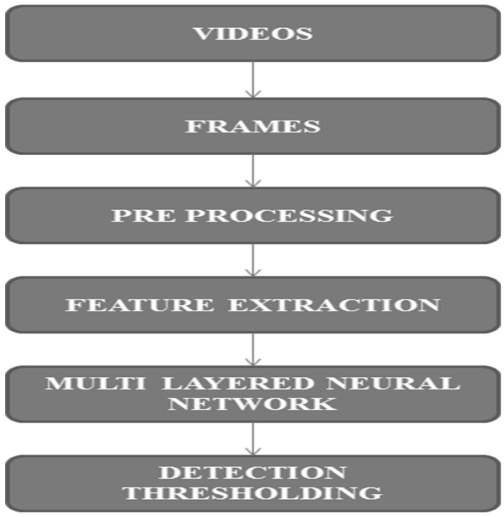

Fig. 1.3: Designed Face Recognition Algorithms

III. EXISTING STRUCTURE

In the existing method motion information, human skin color information, human shape information and variation of ambient lighting are combined. The moving objects in the video sequence images are extracted using the multi-frame differencing method with adaptive ambient illumination changes. The adaptive ambient illumination human skin feature extraction algorithm extracts human skin color in different lighting changes in order to tackle the problem that skin color is susceptible to illumination. Improve Hough transform is used to automatically determine the size of human head in different scenes. The below will discuss in detail about Hough transform.

The Hough transform is a feature extraction technique used in image analysis, computer vision, and digital image processing. The purpose of the technique is to find imperfect instances of objects within a certain class of shapes by a voting procedure. This voting procedure is carried out in a parameter space, from which object candidates are obtained as local maxima in a so-called accumulator space that is explicitly constructed by the algorithm for computing the Hough transform.

The classical Hough transform was concerned with the identification of lines in the image, but later the Hough transform has been extended to identifying positions of arbitrary shapes, most commonly circles or ellipses. The Hough transform as it is universally used today was invented by Richard Duda and Peter Hart in 1972, who called it a "generalized Hough transform" after the related 1962 patent of Paul Hough. The transform was popularized in the computer vision community by Dana H. Ballard through a 1981 journal article titled "Generalizing the Hough transform to detect arbitrary shapes".

In automated analysis of digital images, a sub problem often arises of detecting simple shapes, such as straight lines, circles or ellipses. In many cases an edge detector can be used as a pre-processing stage to obtain image points or image pixels that are on the desired curve in the image space. Due to imperfections in either the image data or the edge detector, however, there may be missing points or pixels on the desired curves as well as spatial deviations between the ideal line/circle/ellipse and the noisy edge points as they are obtained from the edge detector. For these reasons, it is often non-trivial to group the extracted edge features to an appropriate set of lines, circles or ellipses. The purpose of the Hough transform is to address this problem by making it possible to perform groupings of edge points into object candidates by performing an explicit voting procedure over a set of parameterized image objects (Shapiro and Stockman.

The simplest case of Hough transform is the linear transform for detecting straight lines. In the image space, the straight line can be described as $y = mx + b$ where the parameter m is the slope of the line, and b is the intercept (y-intercept). This is called the slope-intercept model of a straight line. In the Hough transform, a main idea is to consider the characteristics of the straight line not as discrete image points (x_1, y_1), (x_2, y_2), etc., but instead, in terms of its parameters according to the slope-intercept model, i.e., the slope parameter m and the intercept parameter b. In general, the straight line $y = mx + b$ can be represented as a point (b, m) in the parameter space. However, vertical lines pose a problem. They are more naturally described as $x = a$ and would give rise to unbounded values of the slope parameter m. Thus, for computational reasons, Duda and Hart proposed the use of a different pair of parameters, denoted r and θ (*theta*), for the lines in the Hough transform. These two values, taken in conjunction, define a polar coordinate.

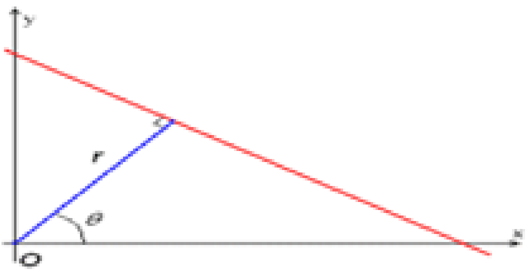

Fig. 3.1: Polar Coordinate

The parameter r represents the algebraic distance between the line and the origin, while θ is the angle of the vector from the origin to this closest point (see Coordinates). Using this parameterization, the equation of the line can be written as

$$y = \left(-\frac{\cos\theta}{\sin\theta}\right)x + \left(\frac{r}{\sin\theta}\right)$$

Which can be rearranged to $r = x\cos\theta + y\sin\theta$ .It is therefore possible to associate with each line of the image a pair (r,θ) which is unique if $\theta \in [0,\pi]$ and $r \in \mathbf{R}$, or if $\theta \in [0, 2\pi]$ and $r \geq 0$. The (r,θ) plane is sometimes referred to as *Hough space* for the set of straight lines in two dimensions. This representation makes the Hough transform conceptually very close to the two-dimensional Radon transform. For an arbitrary point on the image plane with coordinates, e.g., (x_0, y_0), the lines that go through it are the pairs (r,θ) with $r(\theta) = x_0\cos\theta + y_0\sin\theta$, where r (the distance between the line and the origin) is determined by θ.

This corresponds to a sinusoidal curve in the (r,θ) plane, which is unique to that point. If the curves corresponding to two points are superimposed, the location where they cross corresponds to a line that passes through both points. More generally, a set of points that form a straight line will produce sinusoids which cross at the parameters for that line. Thus, the problem of detecting collinear points can be converted to the problem of finding concurrent curves.

A. Implementation:

The linear Hough transform algorithm uses a two-dimensional array, called an accumulator, to detect the existence of a line described by $r = x\cos\theta + y\sin\theta$. The dimension of the accumulator equals the number of unknown parameters, i.e., two, considering quantized values of r and θ in the pair (r,θ). For each pixel at *(x,y)* and its neighborhood, the Hough transform algorithm determines if there is enough evidence of a straight line at that pixel. If so, it will calculate the parameters (r,θ) of that line, and then look for the accumulator's bin that the parameters fall into, and increment the value of that bin. By finding the bins with the highest values, typically by looking for local maxima in the accumulator space, the most likely lines can be extracted, and their (approximate) geometric definitions read off. (Shapiro and Stockman, 304) The simplest way of finding these *peaks* is by applying some form of threshold, but other techniques may yield better results in different circumstances - determining which lines are found as well as how many. Since the lines returned do not contain any length information, it is often necessary, in the next step, to find which parts of the image match up with which lines. Moreover, due to imperfection errors in the edge detection step, there will usually be errors in the accumulator space, which may make it non-trivial to find the appropriate peaks, and thus the appropriate lines.

The final result of the linear Hough transform is a two-dimensional array (matrix) similar to the accumulator -- one dimension of this matrix is the quantized angle θ and the other dimension is the quantized distance r. Each element of the matrix has a value equal to the number of points or pixels that are positioned on the line represented by quantized parameters (r, θ). So the element with the highest value indicates the straight line that is most represented in the input image.

– For each data point, a number of lines are plotted going through it, all at different angles. These are shown here as solid lines.
– For each solid line a line is plotted which is perpendicular to it and which intersects the origin. These are shown as dashed lines.
– The length and angle of each dashed line is measured. In the diagram above, the results are shown in tables.
– This is repeated for each data point.
– A graph of the line lengths for each angle, known as a Hough space graph, is then created.

The point where the curves intersect gives a distance and angle. This distance and angle indicate the line which intersects the points being tested. In the graph shown the lines intersect at the pink point; this corresponds to the solid pink line in the diagrams above, which passes through all three points.

IV. IMPLEMENTATION OF THE PROPOSED SYSTEM

A. Acquisition:

The video database is captured using a digital camera. To input a large video in mat lab is not feasible, so the video is converted to frames using software called studio converter. The frame rate is flexible with the software which has quite a few options. We sampled the video using the option 50 frames per video. The frames which are converted are of standard size 480X640 pixels.

V. PRE PROCESSING

A preliminary processing of data in order to prepare it for the primary processing or for further analysis is called preprocessing. The term can be applied to any first or preparatory processing stage when there are several steps required to prepare data for the user. For example, extracting data from a larger set, filtering it for various reasons and combining sets of data could be preprocessing steps.

A. Histogram Equalization:

A graphical representation, similar to a bar chart in structure, that organizes a group of data points into user-specified ranges. The histogram condenses a data series into an easily interpreted visual by taking many data points and grouping them into logical ranges or bins.

The histogram of an image shows us the distribution of grey levels in the image. That is a graph indicating the number of times each gray level occurs in the image. It is massively useful in image processing, especially in segmentation. Two different images can have same histogram. Figure 5.1 shows a model of a histogram. We can infer a great deal about the appearance of an image from its histogram. In a dark image, the gray levels would be clustered at the lower end. In a uniformly bright image, the gray levels would be clustered at the upper end. In a well contrasted image, the gray levels would be well spread out over much of the range.

B. Edge Detection:

Edge detection refers to the process of identifying and locating sharp discontinuities in an image. The discontinuities are abrupt changes in pixel intensity which characterize boundaries of objects in a scene. Classical methods of edge detection involve convolving the image with an operator (a 2-D filter), which is constructed to be sensitive to large gradients in the image while returning values of zero in uniform regions. There is extremely large number of edge detection operators available, each designed to be sensitive to certain types of edges.

For the purpose of edge detection, Gabor filter has been used. In image processing, a Gabor filter, named after Dennis Gabor, is a linear filter used for edge detection. Frequency and orientation representations of Gabor filters are similar to those of the human visual system, and they have been found to be particularly appropriate for texture representation and discrimination. In the spatial domain, a 2D Gabor filter is a Gaussian kernel function modulated by a sinusoidal plane wave. The Gabor filters are self-similar: all filters can be generated from one mother wavelet by dilation and rotation. A set of Gabor filters with different frequencies and orientations may be helpful for extracting useful features from an image. Figure 5.2.2.1 shows the edge detected image using Gabor filter.

i) $G(x,y;\theta,f)=\exp\{-\frac{1}{2}[(x_\theta^2/\sigma_x^2)+(y_\theta^2/\sigma_y^2)]\}\cos(2\pi f x_\theta)$,

ii) $x_\theta = x\cos\theta + y\sin\theta$,

iii) $y_\theta = -x\cos\theta + y\sin\theta$,

where θ represents the orientation of the normal to the parallel stripes of a Gabor function, f represents frequency of the sinusoidal factor, σ is the standard deviation of Gaussian envelope, x and y are the corresponding axes.

C. Cropping Face From An Image:

Two broad categories in image processing includes the modification in quality of picture like color, sharpness etc. and other is resizing like cropping. In first type of processing we need to modify the amplitude level and in the second category we select or reject some pixel values which are not desired. Cropping an image means creating a new image from the part of the original image. It is a real important task as it may be needed to extract some specific part of an image or to change aspect ratio is major concern in film making. It is generally used to remove the unwanted or irrelevant detail from the photo. Here it's need to crop only the faces from the frames.

VI. FEATURE EXTRACTION

In pattern recognition and in image processing, feature extraction is a special form of dimensionality reduction. When the input data to an algorithm is too large to be processed and it is suspected to be notoriously redundant (much data, but not much information) then the input data will be transformed into a reduced representation set of features (also named features vector). Transforming the input data into the set of features is called feature extraction. If the features extracted are carefully chosen it is expected that the features set will extract the relevant information from the input data in order to perform the desired task using this reduced representation instead of the full size input.

VII. CONCLUSION

In this paper, a general introduction on forensic image processing and its main issues has been exposed, and a new environment for the identification of suspicious images coming from video-surveillance devices has been proposed. Being developed in Mat lab, the system suffers from the drawbacks of this environment, i.e. mainly the speed of execution and possible problems with memory when working with large images. These can be considerably minimized by careful programming with efficient memory handling.

VIII. FUTURE WORK

This application is now limited with single face per frame. It can be further extended to work at any surveillance database say in a shopping mall or in road sides by using robust face detection techniques which can track multiple faces per frames.

REFERENCE

[1] Patel U., Thaker N. and Panchal P. "Shot Boundary detection and Key frame extraction from Uncompressed Video", International Conference of Signals, Systems

and Automation ICSSA 2011 at GCET, V.V.Nagar. ISBN No. 978-1-6123-3002-0.
[2] Oge Marques Florida Atlantic University, “Practical image and video processing using Matlab®”, book published by John wiley & sons, 2011. IEEE.
[3] N. Ikizler-Cinbis, S. Sclaroff, “Object, Scene and Actions: Combining Multiple Features for Human Action Recognition”, ECCV (1), 2010, pp. 494–507.
[4] In Su Kim, Hong Seok Choi, Kwang Moo Yi et al. Intelligent visual surveillance—A survey. International Journal of Control, Automation and Systems,2010, 8(5):926-939.
[5] Yong-feng QI, Yuan-lian HUO.Color-based method for face detection and location. Journal of Computer Applications, 2009, 29(3):785-78