ISSN (ONLINE): 2454-9762
ISSN (PRINT): 2454-9762
Available online at www.ijarmate.com

International Journal of Advanced Research in Management, Architecture, Technology and Engineering (IJARMATE) Vol. 3, Special Issue 2, April 2017

Face Identification And Recognition For Surveillance System Using ANN

A.Anithadevi
Department of CSE
Indra Ganesan College of Engineering
Trichy, India
aniashok18995@gmail.com

K.Bhuvaneswari
Department of CSE
Indra Ganesan College of Engineering
Trichy, India
bhanubhuvana110@gmail.com

M.Shanmuga Priya
Department of CSE
Indra Ganesan College of Engineering
Trichy, India
priyam004sa@gmail.com

Mrs. T. Sugashini
AP/CSE
Indra Ganesan College of Engineering
Trichy, India
sugashinitv@gmail.com

***Abstract*—**Human population is increasing at present. As the population is on an increase in other back side crimes and terrorism also increases which leads to security lack. Security is the most important factor to maintain peace among people. Manual monitoring is one of the security factor but it is not 100% effective in crowd and public surveillance area especially in Airport premises. In this project a system is proposed for human face detection and recognition in live videos. The system is segregated into five stages- Motion detection, Face acquisition, Face detection, Face recognition and Representation. The system is to detect the movement of the person and to capture the face from live video. If there is a face in the view, it is detected within a fraction of second. A multi-scale algorithm is used to search a face in low resolution. A face recognition system automatically identifies a human face from database images. ANN (Artificial Neural Network) must concentrate more on security in public surveillance area. The Artificial Neural Network (ANN) support to enhance the recognition process. After recognition process, an alert message is reached to the person who monitor the frame. The important aspects of this project is to identify the criminals easily with the help of Artificial Neural Network (ANN).Thus the highlight of this project is to reduce human intervention and processing time for detection and recognition process and also increase overall system efficiency.

Keywords—Motion Detection,Face Acquisition,Face Detection, Face Recognition,Representation

I. Introduction

Automatic human face detection is a challenging field of research with many useful real life applications. The use of computer vision in security applications and to minimize intervention of human beings has led the research in field of face biometrics. The fig 1 shows the system architecture. Face is vital part of human being that represents important information like expression, attention, identity etc. of an individual. The goal of face detection is to locate the occurrence of face in the frame and recognition system retrieves the identity of person for authorization. The main application of face recognition is "access control" that grants certain permissions to person detected. Numerous techniques are developed and available to detect and identify faces in images. Application of these methods to videos and real-time applications has many hindrances like processing speed, system complexity, and factors like illumination and pose variations, etc. In any image processing application first stage is pre-processing that removes unwanted artefacts from image or frame captured.

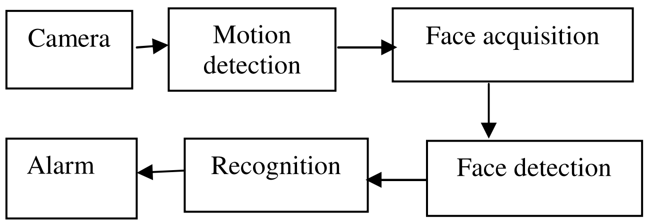

Fig 1. System Architecture

A face recognition system automatically identifies a human face from database images. The face recognition problem is challenging as it has to account all possible variation in face caused by change in facial features, illumination, occlusions, etc.., Recognition stage uses face to identify a person and claim identity. Face recognition is time consuming process as it has to undergo large number of

ISSN (ONLINE): 2454-9762
ISSN (PRINT): 2454-9762
Available online at www.ijarmate.com

International Journal of Advanced Research in Management, Architecture, Technology and Engineering
(IJARMATE) Vol. 3, Special Issue 2, April 2017

comparisons. The query image is compared with the images from database. Eigen face and principle component analysis technique is most popular method for recognition. The very important step in process is decision making which decides possible match. Most technologies available are resistant to modest changes in the features of face as prior stages bring all variations in features to ground level using additional processing

II. RELATED WORK

Their research in face detection and recognition has reached to far extent. This process has started from 1960 when the first semi-automated algorithm was developed. In 1988 principle component analysis was applied for face recognition which boosted the research eye recognition stage. In 1991, reliable real time automated recognition was reached by using residual error to detect faces where Eigen face technique was used. In recent decade Viola and Jones [1] gave a new boost to real time face detection with the use of wavelet based method. The very first stage in the system is motion sensing or detection. The rest stages rely highly on this stage.

A. Motion Detection

Background subtraction models have been widely used for detecting or localizing moving objects in video scenes. Any motion detection algorithm is based on the following principle: At the start of the algorithm frames are taken from camera feed and used to model background. Acquiring a background image which does not include any moving object is the most elementary way to model to the background. More details on background modelling can be found in survey article [2]. Background subtraction has many issues like videos with poor signal-to-noise ratio caused by a low quality camera, camera jitter, compression artifacts, noisy environment, illumination changes, natural movements of stationary objects like trees shaken by wind etc. that causes real challenge in preparing background model for any application [3]. Background subtraction is done by employing Gaussian Mixture Model (GMM) [4], [5]. To overcome changes due to small movements of stationary objects in background multimodal probability distribution functions (GMM) are used [7], [8].

B. Face Acquisition

When the system is attached to a video surveillance system, the recognition software searches the field of view of a video camera for faces. If there is a face in the view, it is detected within a fraction of a second. A multi-scale algorithm is used to search for faces in low resolution. An image or a video, if digitized is represented by a number of frames per unit of time, with each frame in turn represented by a number of components (three colors or more), each again represented by a set of pixel at a given precision (8 or more bits), scanning the 8 frame component on a raster, line by line. This is often referred to as first general representation, and was introduced taking into account practical issues such as camera and scan technologies, as well as simplicity of their representation. First generation image and video can be represented as one or more matrices whose elements correspond to a frame's component pixel. When compared to the first, second generation representation approach represents image and video as set of what is called attributes. A largely popular second general representation is that of object-based representation where to each object has been assigned some color, texture or motion attributes. The majority of image and video segmentation techniques try to take a first general image or video as an input and provide as output a second generation representation of them. Other image and video analysis tools extract other and provide what one generally calls a content-based representation in form of edges, features points, and others.

C. Face detection

The simplest method for face detection is based on skin color information [1]. Skin pixel values for r plane is in the range 0.36 to 0.456 and g value in 0.28 to 0.363. In [2] combined smoothed 2-D histogram and Gaussian model is used for skin pixel detection to overcome individual limitations to increase accuracy to 90%. In HSV model pixel is classified as skin pixel if it has values $0 \leq H \leq 50$ and $0.20 \leq S \leq 0.68$. In [5] combination of two models HSV and $Y\ CbCr$ is used with classification range of $Cb \leq 125$ or $Cb \geq 160$; $Cr \leq 100$ or $Cr \geq 135$; $26 < Hue < 220$. Skin color based method is sensitive to illumination variation and fails if background contains skin color like objects. Complex methods uses depth features to detect face based on geometry and texture pattern. Face is detected using edge maps and skin color thresholding in [6].Used $Y\ CbCr$ space for skin pixel detection and Viola-Jones [8] method to verify correct detection. LPA and LBP methods to extract textural features with LDA dimensionality reduction and SVM classification [9]. Used 3×3 block rank patterns of gradient magnitude images for face detection. Use logarithmic difference edge maps to overcome illumination variation with face verification ability of 95% true positive rate. Viola and Jones [8] proposed real-time AdaBoost algorithm for classification of rectangular facial features. Next stage after face detection is recognition where person's identity is verified by searching query face match in the database.

D. Recognition

There is a great diversity in the way facial appearance is interpreted for recognition by an automatic system. Currently a number of different systems are under development, and

ISSN (ONLINE): 2454-9762
ISSN (PRINT): 2454-9762
Available online at www.ijarmate.com

IJARMATE
Your ult-MATE Research Paper !!!

International Journal of Advanced Research in Management, Architecture, Technology and Engineering
(IJARMATE) Vol. 3, Special Issue 2, April 2017

which is most appropriate may depend on the application domain. A major difference in approaches is whether to represent the appearance of the face, or the geometry. Geometry is difficult to measure with any accuracy, particularly from a single still image, but provides more robustness against disguises and aging.

1) Transform two dimensional $N \times N$ image (x, y) into a vector of dimension $1 \times N2$ each face image is represented by the vector Γi.

2) Consider there are total M images. Average face image is calculated by

$$\Psi = 1/\mathrm{M}\ \Sigma\Gamma i\ Mi{=}1 \quad (1)$$

3) Mean centred image which is difference of each image from mean image is calculated as

$$\Phi = \Gamma i - \Psi i \quad (2)$$

Appearance information is readily obtained from a face image, but is more subject to superficial variation, particularly from pose and expression changes. In practice for most purposes, even appearance-based systems must estimate some geometrical parameters in order to derive a 'shape free' representation that is independent of expression and pose artifacts. This is achieved by finding facial landmarks and warping the face to a canonical neutral pose and expression. Facial features are also important for geometric approaches and for anchoring local representations. Face appearance representation schemes can be divided into local and global, depending on whether the face is represented as a whole, or as a series of small regions. Most global approaches are based on a principal components representation of the face image intensities. This representation scheme was devised first for face image compression purposes and subsequently used for recognition purposes. The latter coined the term Eigen faces for this type of representation. A face image is 10 represented as a vector of intensities and this vector is then approximated as a sum of basis vectors (Eigen faces) computed by principal component analysis from a database of face images. These principal components represent the typical variations seen between faces and provide a concise encapsulation of the appearance of a sample face image, and a basis for its comparison with other face images. This principal components representation is, like for example the Fourier transforms, a de-correlating transform to an alternative basis where good representations of the salient characteristics of an image can be created from only a few low-order coefficients despite discarding many of the higher-order terms.

E. Representation

The system translates the facial data into a unique code. This coding process allows for easier comparison of the newly acquired facial data to stored facial data. Having processed a face and extracted the features, these are stored or transmitted as a facial code (face template), which can be as small as 84 bytes. For each representation type, a distance or similarity measure is defined that allows 'similar' faces to be determined. Much of the art in biometrics is in the design of a model of the biometric data and, given a scheme for extracting the model parameters as a representation of the data, in creating a similarity measure that correctly discriminates between samples from the same person and samples from different people. As with any biometric system, some threshold on similarity must be chosen above which two face images are deemed to be of the same person.

III. METHODOLOGY

Background subtraction reduces the search area for further processing in system. This section describes each method in detail.

A. Motion Detection

Background subtraction models have been widely used for detecting or localizing moving objects in video scenes. Any motion detection algorithm is based on the following principle: At the start of the algorithm frames are taken from camera feed and used to model background. Acquiring a background image which does not include any moving object is the most elementary way to model to the background. Background subtraction has many issues like videos with poor signal-to-noise ratio caused by a low quality camera, camera jitter, compression artifacts, noisy environment, illumination changes, natural movements of stationary objects like trees shaken by wind etc. that causes real challenge in preparing background model for any application. Background subtraction is done by employing Gaussian Mixture Model (GMM). It describes advantages of multimodal mean real-time background modelling over other methods. To overcome changes due to small movements of stationary objects in background multi-modal probability distribution functions (GMM) are used. Another segmentation method temporal differencing uses pixel-wise differences between successive two to three frames to extract moving regions. Temporal differencing is dynamic background modelling method. The proposed system use dynamic background moving object detection method. This method is based on a neural-network model that deals with dynamic background and shadow elimination. Optical-flow-based motion segmentation uses timely directional vectors of moving objects to detect moving objects. Optical-flow-based methods are robust to camera motion. The below Fig 2 shows that moving objects in live video.

ISSN (ONLINE): 2454-9762
ISSN (PRINT): 2454-9762
Available online at www.ijarmate.com

International Journal of Advanced Research in Management, Architecture, Technology and Engineering
(IJARMATE) Vol. 3, Special Issue 2, April 2017

Fig 2. Motion Detection from live-video

Once the moving objects are identified face detection algorithms are employed in these regions to search and locate human face. Background model is established with absence of moving object. Adaptive background model is where a series of frames are used and an average is calculated from all of them over a period of time. This approach is useful if there are continuous changes in background scene. Non-adaptive background model is where a frame is taken and saved as the background. This approach is useful in static and indoor environments where there are very less variations in background over a time period. The background is then subtracted from current frame to segment moving objects. The decision of presence of moving object is based on the difference between the two images or frames.

B. Face Acquisition

The simplest method for face detection is based on skin color information. In combined smoothed 2-D Histogram and Gaussian model is used for skin pixel detection to overcome individual limitations to increase accuracy to 90%.

1) The Mean difference of skin color gives pixel detection.

(a) 2-D Histogram

(b) Gaussian model

2) In HSV model pixel is classified as skin pixel if it has values $0 \leq H \leq 50$ and $0.20 \leq S \leq 0.68$

C. Face Detection

Haar face detector is most popularly used face detection algorithm. Though background subtraction requires more computation time it greatly reduces the size of area in which face image is to be detected.

Integral images are also used for calculating the values of $(X2)$, μ and variance at any position in an image given in Eq. 3

$$\mu = 1/N\ (I1 + I4 - I2 - I3) \quad (3)$$

$$E(f\ (x, y)2) = 1/N\ (I21 + I24 - I22 - I23) \quad (4)$$

$$Var\ (f\ (x, y)) = E\ (f\ (x, y)2\ ) - \mu 2 \quad (5)$$

where N is number of elements within region D. The Haar classifier multiplies the weight of each rectangle by its area and adds to the results.

Fig 3. Face Detection

The fig 3 shows face detection which uses following detection methods. Face is detected using edge maps and skin color threshold value also used $YCbCr$space for skin pixel detection and Viola-Jones method to verify correct detection. Using 3× 3 block rank pattern of gradient magnitude images for face detection.

D. Recognition

The proposed system presents Eigen and fisher face as image based method and statistical shape and active shape models as feature based methods for face recognition. Feature based methods achieves robustness to illumination and pose variations with use of non-linear feature spaces.

The Eq. 6 calculate a covariance matrix is constructed as:

$$C = AAT \quad (6)$$

where $A = [\Phi 1, \Phi 2, \cdots, \Phi M]$. Size of matrix C is $N2 \times N2$

Calculate the eigenvectors and eigenvalues of the covariance matrix. The eigenvectors of this covariance matrix resembles facial images which look ghostly and are called Eigenfaces. They are the directions in which the images differ from the mean image. Eigen values are calculated in Eq.7

$$AATXi = \lambda iXi \quad (7)$$

where λ is eigenvalue and Xi are eigenvectors for i =1, 2, 3, ..,N. If the eigenvalues are all distinct, there are N associated linearly independent eigenvectors whose directions are unique.

E. Representation

When criminal activity or a threat is detected, security personnel and the proper authorities can be provided with real time information when assisting the situation. Various alerts can be set up, triggered by predefined operationally relevant

ISSN (ONLINE): 2454-9762
ISSN (PRINT): 2454-9762
Available online at www.ijarmate.com

International Journal of Advanced Research in Management, Architecture, Technology and Engineering
(IJARMATE) Vol. 3, Special Issue 2, April 2017

events. Information can be disseminated using text messaging, on-screen alerts, E-mail, geo-coded maps, pictures, and video. Accurate people detection can increase management efficiency in public transportation by marking areas with high congestion or signalling areas that need more attention. Estimation of crowds in underground transit system can be used to give passengers a good estimate of the waiting time in a queue. Multiple solutions to automate the crowd-counting process have been proposed, including solutions from a moving platform(such as a camera on the airport) that analyse the optic flow generated from moving objects as well as the moving platform.

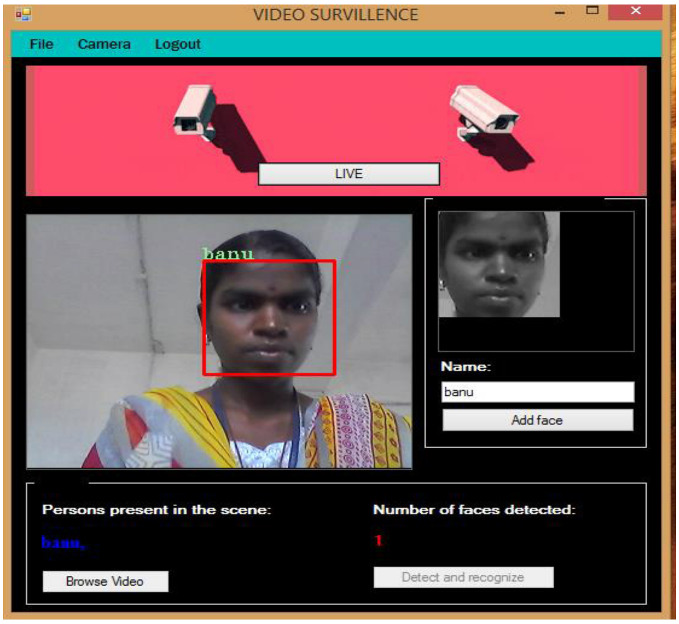

Fig 4. Representation process

The Fig 4 shows the representation process of this project. The person entered in the camera vision the system capture the face of that person. If the face is matched in the database saved faces then the face is highlighted with red color square box and the alarm is going on, otherwise the face is highlighted with green color square box.

IV. RESULTS

Table 1 shows the comparison carried out at each level and its recognition rate of existing system and the proposed system.

Levels	Coefficients	Recognition rate without normalized image	Recognition rate with normalized image
Level 1	150	85%	91%
Level 2	75	91%	91%
Level 3	20	93%	95%
Level 4	10	89%	93%
Level 5	5	87%	93%

Table 1. Experimental result

There are five levels are present. Level 1 represent motion detection process which is used to have a recognition rate up to 91% efficiency and accuracy that is possible with the normalized image. Recognition rate without normalized image exist the result with 85% only. Level 2 represent face acquisition process which is to capture the image from the video and the face can be acquired with 3x3 matrix pixel representation. Face acquisition process shrink the number of crowd counting and to decide whether it is face or not.

Level 3 represent Face detection process offers better performance in comparison to other levels. The face can be detected with the pose variation and it can produce the result as 95% when the recognition process carried out normalized image. Level 4 represent Face recognition process which is used to match the corresponding person in the database. Face recognition process have the higher performance with 93% compared to other process.

Level 5 include Representation process which is carried out through face recognition process. The representation process can run through alarm signals and it provides an efficient result with the rate of 93%.

V. CONCLUSIONS AND FUTURE WORK

The system here makes use of background subtraction in Haar face detector. Background subtraction requires more computation time affecting the overall system performance. The advantage of using this stage is reduced region of interest. In later stage face is detected only in the region provided by this stage. Face is important biometric feature of human anatomy and has many peculiar features. Popularly used haar detector is used for detecting faces in videos. Haar detector is improvised by using integral image to reduce computations. Eigen face method greatly reduces the computations by projection image from two-dimension to one-dimension. In this evaluation first frame of every video was modelled as background frame. In few videos moving object was present in background frame. There is vast scope for use of these methods in real-time applications like surveillance systems

ISSN (ONLINE): 2454-9762
ISSN (PRINT): 2454-9762
Available online at www.ijarmate.com

International Journal of Advanced Research in Management, Architecture, Technology and Engineering
(IJARMATE) Vol. 3, Special Issue 2, April 2017

where an alarm can be raised on detecting un-identified person in the premises.

VI. REFERENCES

1. Blanz.V and Vetter.T. (2003) Face recognition based on fitting a 3d morphable model, IEEE TPAMI.
2. Cristinacce.D and Cootes.T. (2006) Feature detection and tracking with constrained local models, In BMVC.
3. Everingham.M, Sivic.J, and Zisserman.A. (2006) – automatic naming of characters in TV video, In BMVC.
4. Han.J, Lao.W, P.H.N. (2003) Automatic video-based human motion analyzer for consumer surveillance system, IEEE Transactions on, vol. 55.
5. Hu.W, Tan.T, Wang.L, May bank.S. (2005) A survey on visual surveillance of object motion and behaviors, IEEE Transactions on Systems, Man and Cybernetics, vol. 34.
6. Jeeva.S, Sivabalakrishnan.M. (2001) Survey on background modeling and foreground detection for real time video surveillance. Procedia Computer Science, vol. 50 big Data.
7. Murphy-Chutorian.E and Trivedi.M. (2009) Head pose estimation in computer vision: A survey, IEEE TPAMI.
8. Narote.S.P,Narote.A.S,Tathe.S.V. (2016) Human Face Detection and recognition In videos, IEEE ICACCI.
9. Rowley.H, Baluja.S, and Kanade.T. (1998) Neural network-based face detection, IEEE TPAMI.
10. Viola.P, Jones.M. (2001) Rapid object detection using a boosted cascade of simple features, Proceedings of the 2001 IEEE Compute Society Conference on Computer Vision and Pattern Recognition, vol. 1.