IJCCN International Journal of Computer Communications and Networks,
VOLUME 4, ISSUE 1, FEBRUARY 2014, ISSN 2289-3369
WWW.IARTC.NET

DESIGN AND DEVELOPMENT OF REAL TIME TRACKING OF HUMAN FACES SECURITY IDENTIFICATION SYSTEM

V.K. Narendira Kumar [1] and B. Srinivasan [2]

[1]Assistant Professor, Department of Information Technology,
[2]Associate Professor, PG & Research Department of Computer Science,
Gobi Arts & Science College (Autonomous), Gobichettipalayam – 638 453,
Erode District, Tamil Nadu, India.
Email ID: [1] kumarmcagobi@yahoo.com, [2] srinivasan_gasc@yahoo.com

ABSTRACT

Robust tracking of persons in real-world environments and in real-time is a common goal in many video applications. In this paper a computational system for the real-time tracking of multiple persons in natural environments is presented. Face detection has diverse applications especially as an identification solution which can meet the crying needs in security areas. The region extractor is based on the integration of skin-color, motion and silhouette features, while the face detector uses a simple, rule-based face detection algorithm and SVM. Exemplary results of the integrated system working in real-world video sequences. New intelligent processing methods, as well as security requirements make multiple-person tracking a hot area. This application is robust tracking in real-world environments and in real-time.

KEYWORDS

Face, Skin, Color, Shadow, Tracking, and Noise Filtering.

1. INTRODUCTION

In the last years we have seen an increasing interest on video processing applications. In this context, the aim of this paper is to present a computational system for real-time tracking of multiple persons in real-world environments, which integrates state-of-the-art methodologies for the analysis of movement and color, as well as for the analysis of faces.

The movement analysis subsystem is based in the system proposed in [1, 2], which uses background subtraction and selectively to exclude from the background model moving visual objects and their shadows, while retaining ghosts. The color analysis subsystem uses a standard skin detector algorithm [3], which increases the performance of the whole system by reducing the searching region for the faces. Finally, the face analysis subsystem is based mainly on the face detection system proposed in [5]. This system uses simple, rectangular feature face detectors (a kind of Haar wavelets), the integral image for fast computation of these feature detectors, asymmetrical Adaboost as a boosting strategy for the training of the classifiers, and a cascade structure for combining successively more complex classifiers.

Automatic recognition of human faces is one of the challenging problems in pattern recognition. A comprehensive survey of still and video-based face recognition techniques can be found in [1]. Various methods have been proposed in the literature such as Eigen face, elastic graph matching, line edge map and support vector machines [7]. The method proposed in this paper satisfies the following requirements for a face authentication technique:

- Invariant to size and tilt of the face

- Invariant to variations in natural lighting conditions
- Able to authenticate a subject within a reasonable time
- New subject can be added to the system without using the features of other subjects.

1.1. Video Analysis

Video analysis and video surveillance are active areas of research. The key areas are video-based detection and tracking, video-based person identification, and large-scale surveillance systems. A significant percentage of basic technologies for video-based detection and tracking program called Video Surveillance and Monitoring (VSAM). This program looked at several fundamental issues in detection, tracking, auto calibration, and multi-camera systems. System issues detection and tracking but also issues of event detection and automatic system [7].

1.2. Related Work

A large literature exists concerning movement analysis in video streams. As an example, last year was held the PETS 2002 event, in which several state-of-the-art tracking and surveillance systems were presented and tested [4]. Different approaches have been proposed for moving object segmentation; including frame difference, double frame difference, and background suppression. In the absence of any a priori knowledge about target and environment the most widely adopted approach is background suppression. Among other approaches, the work proposed in [1] has shown good performance in the analysis of real-world video sequences. It includes selective background update [2], a verification step for including ghosts into the background model and the use of the HSV color space for dealing with shadows.

Regarding face detection, different approaches have been proposed to solve this task. Main approaches can be classified as feature-based (low-level analysis, feature analysis and active shape models) and image-based (linear subspace methods, neural networks and statistical approaches). Image-based approaches have shown a better performance than feature based. Among them, systems like the ones developed by “Sung & Poggio”, “Rowley”, “Schneider man” and SNoW have shown very good results.

However, the system proposed by Viola and Jones outperforms previous systems in terms of processing speed. This system uses simple, rectangular features (a kind of Haar wavelets), a cascade of filters that discard non-face images, the integral image for fast computation of these filters and asymmetrical Ad boost as a boosting strategy for the training of the detectors.

Furthermore, a module for color analysis, based on a standard skin detector algorithm increases the system performance when color images are used. We believe that the proposed system outperformed similar systems in the task of robust, real-world, real-time tracking of multiple-persons. For instance, a related multiple-person tracking was proposed in [3]. This system is composed by an *Interesting Region Extractor* module and a *Face Detector* Module. The region extractor is based on the integration of skin-color, motion and silhouette features, while the face detector uses a simple, rule-based face detection algorithm and SVM. Although this is a real-time system (20 frames/second using 320x240 pixels on a Pentium III, 500 MHz), it seems that it is not robust enough to work on real-world environments.

2. Background of the Study

Several approaches have been proposed for moving object segmentation; including frame difference, double frame difference, and background suppression. In the absence of any a priori knowledge about target and environment the most widely adopted approach is background suppression. Among other approaches, the work proposed in has shown good performance in the analysis of real-world video sequences. It includes selective background update, a verification step for including ghosts into the background model and the use of the HSV color space for dealing with shadows. Face detection approaches can be classified as:

1. feature-based (low-level analysis, feature analysis and active shape models)
2. Image-based (linear subspace methods, neural networks and statistical approaches). Image-based approaches have shown a better performance than feature based.

2.1. Face Identification

Facial images are the most common biometric characteristic used by humans to make a personal recognition, hence the idea to use this biometric in technology. This is a nonintrusive method and is suitable for covert recognition applications. The applications of facial recognition range from static ("mug shots") to dynamic, uncontrolled face identification in a cluttered background (subway, airport). Face verification involves extracting a feature set from a two-dimensional image of the user's face and matching it with the template stored in a database.

The most popular approaches to face recognition are based on either: 1) The location and shape of facial attributes such as eyes, eyebrows, nose, lips and chin, and their spatial relationships, or 2) The overall (global) analysis of the face image that represents a face as a weighted combination of a number of canonical faces [2]. It is questionable if a face itself is a sufficient basis for recognizing a person from a large number of identities with an extremely high level of confidence [6]. Facial recognition system should be able to automatically detect a face in an image, extract its features and then recognize it from a general viewpoint (i.e., from any pose) which is a rather difficult task. Another problem is the fact that the face is a changeable social organ displaying a variety of expressions.

3. SYSTEM METHODOLOGY

The system proposed uses background subtraction and selectively to exclude from the background model moving visual objects and their shadows, while retaining ghosts. The color analysis subsystem uses a standard skin detector algorithm which increases the performance of the whole system by reducing the searching region for the faces. Finally, the face analysis subsystem is based mainly on the face detection system.

The system uses simple, rectangular feature face detectors (a kind of Haar wavelets), the integral image for fast computation of these feature detectors, asymmetrical Adaboost as a boosting strategy for the training of the classifiers, and a cascade structure for combining successively more complex classifiers. Face detection is complemented by a face tracking module based on heuristics.

The influence of the neural network is more in pattern recognition, for this reason the neural network based classifiers incorporates both statistical and structural information and achieve better performance. Multilayered networks (mlns) implementing Back propagation (bp) Algorithms. RBF neural network have been applied in many engineering and scientific applications including face recognition. RBF nn’s includes:
1. Universal approximates
2. Simple topological structure.
3. Implements fast learning algorithms based on locally tuned neurons.

Based on the merits of the RBF nn’s and the efficient feature extraction methods, high speed RBF nn’s classifiers for estimating near optical parameters according to the property of feature space instead of using the gradient descent training algorithm, the system proposed can achieve high training and recognition speed facilitating in real time application of the proposed face recognition system.

The proposed system outperformed similar systems in the task of robust, real-world, real-time tracking of multiple-persons. A related multiple-person tracking proposed is composed by an Interesting Region Extractor module and a Face Detector Module [3]. The region extractor is based on the integration of skin-color, motion and silhouette features, while the face detector uses a simple, rule-based face detection algorithm and SVM. Although this is a real-time system (20 frames/second using 320x240 pixels on a Pentium III, 500 MHz), it seems that it is not robust enough to work on real-world environments.

3.1. Method Description

The proposed method for tracking multiple human faces skin-colored objects operates as follows. At each time instance, the camera acquires an image on which skin-colored blobs (i.e. connected sets of skin-colored pixels) are detected. The method also maintains a set of object hypotheses that have been tracked up to this instance in time [7]. The detected blobs, together with the object hypotheses are then associated in time. The goal of this association is (a) to assign a new, unique label to each new object that enters the camera's field of view for the first time, and (b) to propagate in time the labels of already detected objects.

3.2. Requirements

- Invariant to size and tilt of the face.
- Invariant to variations in natural lighting conditions.
- Able to authenticate a subject within a reasonable time.
- New subject can be added to the system without using the features of other subjects.

4. System Study

In figure 1 is shown a block diagram of the proposed system. The system is composed by 3 main subsystems: *Movement Analysis*, *Color Analysis* and *Face Analysis* [3].

4.1. Movement Analysis Subsystem

The movement analysis subsystem includes the following sub-systems as follows:

Shadow Detection: The shadow detection is categorized in to 2 types. A darkened shadow and a lighten shadow in the background which are detected from the empirical and by the illumination values.

Blob Identification: With the help of 8-connectivity, the system detect all the blobs of connected candidate moving points. Blobs with small area are discarded as noise while the rest are validated as actual MVOs (Moving Visual Object).

Blob Analysis and Tracking: With every MVO we compute its average speed by means of frame-difference. By using a threshold on *AS* we separate the MVO as a moving MVO and stopped MVO.

Background Update: The background model is computed as a statistical combination of a sequence of previous frames and the previously computed background (adaptability). The statistical function used is the median. In order to improve the background update, the system use selectivity, so the background is updated [4].

4.2. Color Analysis Subsystem

Skin color detection involves (a) estimation of the probability of a pixel being skin-colored, (b) hysteresis thresholding on the derived probabilities map, (c) connected components labeling to yield skin-colored blobs and, (d) computation of statistical information for each blob. The color analysis subsystems are categorized as follows:

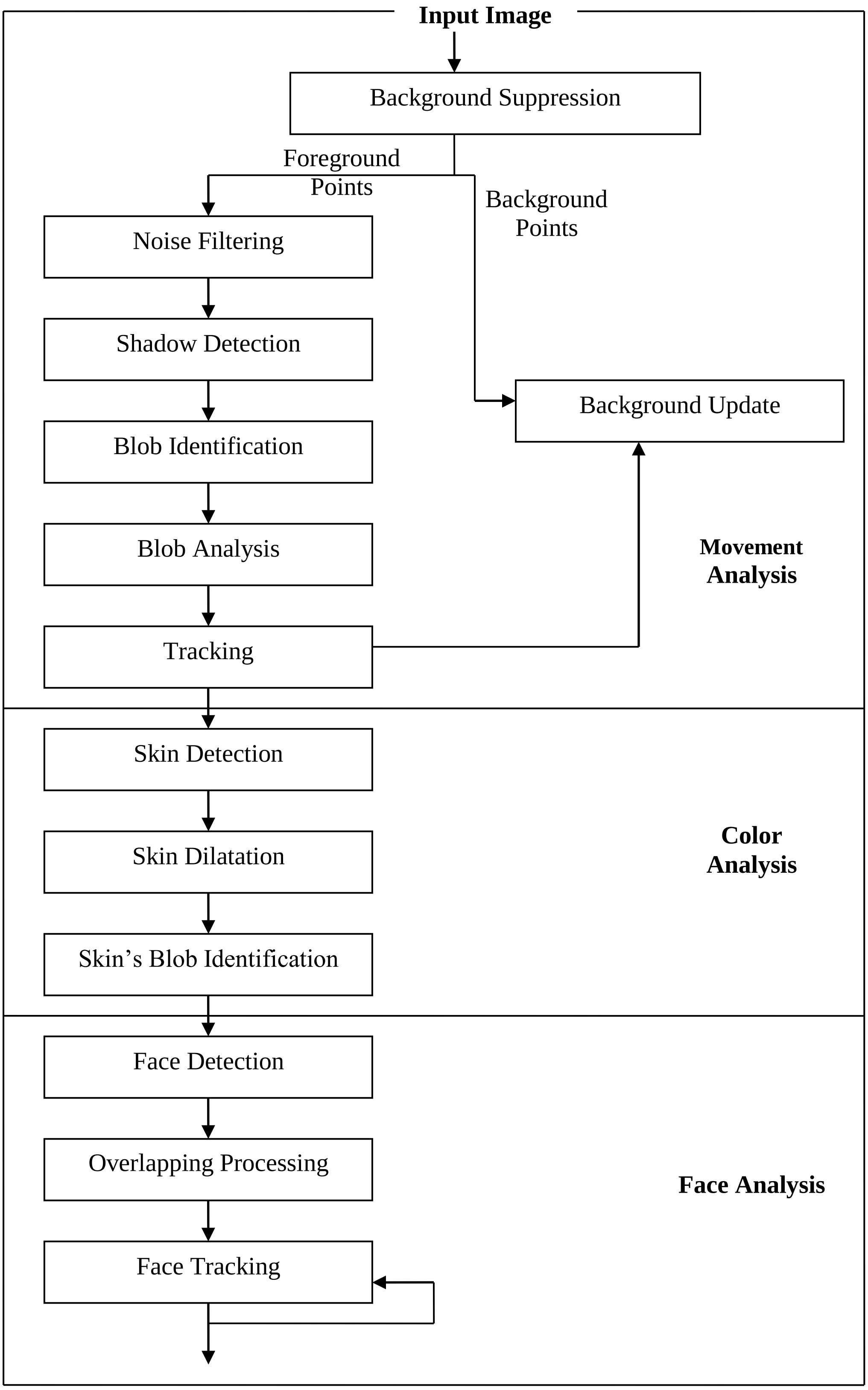


Figure 1: Block Diagram of the Real Time Tracking of Multiple Human Faces Identification

Skin Detection: To reduce the search area for face detection (increasing the processing speed and decreasing the false detection rate), the system used simple rules to verify if a point belonging to MVO has a skin color or not, using the pixels' normalized RG color space information.

Skin Dilation: Among the selected points, some are discarded as noise applying a 5x5 morphological opening, and afterward a dilatation using a 3x3 square structuring element is performed.

Skin's Blob Identification: A region-based labeling is preformed to compute the connected skin's blobs of skin pixels Blobs with small area are discarded as noise.

4.3. Face Analysis Subsystem

The face analysis subsystem is area in which the radial basis function neural network is implemented.

Face Detection: The implemented detection subsystem detects frontal faces with small in-plane rotations and it is based mainly on radial basis function neural network. This face detector corresponds to a cascade of filters that discard non-faces and let faces to pass to the next stage of the cascade.

Overlapping Detections Processing: Face windows obtained in the face detection module are processed and fused for determining the size and position of the final detected faces. Overlapping detections were processed for filtering false detections and for merging correct ones.

Face Tracking: This module was used to filter false detections. This filtering corresponds to an inter-frame operation, while the filtering applied in the Overlapping Detections Processing module to an intra-frame operation. Face detections belonging to consecutive frames were considered to be the same face, by applying the same heuristic used to process overlapping detections [4].

5. SYSTEM DESIGN

System design is a transition from a user-oriented document to a document oriented to programmers or database personnel. It goes through logical and physical design with emphasis on the following:

- Preparing input/output specifications.
- Preparing security and control specifications.
- Specifying the implementation plan.
- Preparing a logical design walkthrough before implementation.

5.1. Study and Analysis

The complete analysis for the requirements of the development of the system is determined in this phase. Learn the technical system and the logical aspects for the requirements of the system. The methods adopted and algorithms followed in the system provide a better solution for the current problems. The limitations of the scope lie purely on the wealth that is provided.

Step 1: Object Detection. This step detects objects of interest as they move about the scene. The object detection process is independently applied to all the static cameras present in the scene.

Step 2: 2-D Object Tracking. The objects detected in Step 1 are tracked within each camera field of view based on object appearance models [6].

Step 3: 3-D Object Tracking. The 2-D object tracks are combined to locate and track objects in a 3-D world coordinate system. This step uses the 3-D wide-baseline stereo tracking discussed previously. The result of the 3-D tracking is an association between the same object as seen in two overlapping camera views of the scene.

Step 4: 3-D Head Detection. To locate the position of the head in 3-D, we use the head detection technique described earlier. Given a 3-D track, the head is first detected in the corresponding 2-D views. The centric of the head in the two views is used to triangulate the 3-D position of the head.

Step 5: Active Camera Assignment. This step determines which of the available active cameras will be used for which task. Let us consider the example of a scene with three objects and a face cataloger system with two available active cameras. This step will employ an algorithm that uses an application dependent policy to decide the camera assignment.

Step 6: 3-D Position-Based Camera Control. Given the 3-D position of the head and a PTZ camera that has been assigned to the object, the system automatically steers the selected active camera to foveae in on the measured location of the head. There are several ways of controlling the pan-tilt and zoom parameters. For example, the zoom could be proportional to the distance of the object from the camera and inversely proportional to the speed at which the object is moving.

Step 7: Face Detection-Based Camera Control. Once the (frontal) face image is detected, the camera is centered on the face and the zoom is increased. The pan and tilt of the camera are controlled based on the relative displacement of the center of the face with respect to the center of the image. Given the intrinsic calibration parameters of the camera and the current zoom level (i.e., focal length), the relative image coordinate displacements are translated into desired (relative) pan/tilt angles. To avoid any potential instability in the feedback control strategy, we use a damping factor in the process.

5.2. Designing

The skeleton of the entire process is prepared in this phase. The scheduling and interactivity of the system for its completed in designing. The work structure including look and feel is generated for the system.

5.3. Coding

The algorithm implementation and the system functionality are finalized in this module. The system is complete integration of all the working models Developed.

6. IMPLEMENTATION OF THE SYSTEM

In order to implement this real-time tracking of multiple face efficiently, ASP.NET program is used. This program could speed up the development of this system because it has facilities to draw forms and to add library easily.

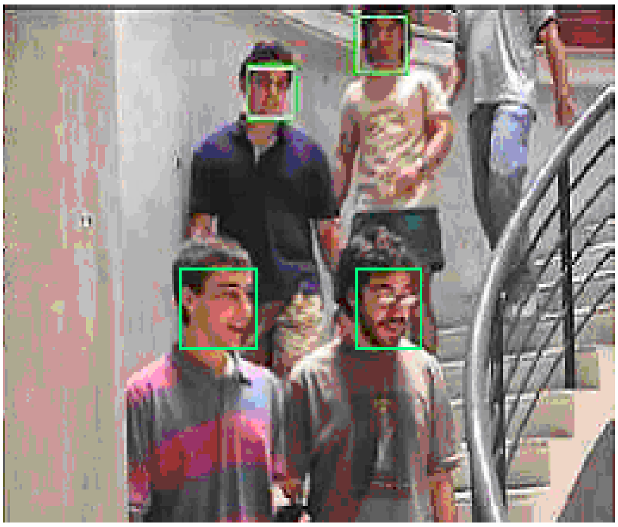


Figure 2: The results of the proposed system on the video. Face detection windows are shown. All faces were detected at least once.

The proposed multiple person tracking system using real-world video sequences. Video sequences were 320x240 pixels size. The videos were captured with a digital video camera with a frame rate of 29.74 frames per second, and later compressed in avi format, obtaining low quality images. Both videos contain 4 or more faces that appear at least on 30 frames each face.

The computer used was an Athlon 1.2 GHz. The frame rate obtained was about 8.5 frames per second. At the moment we are performing extensive tests of our system. The system produces correct samples of outputs for several kinds of inputs. The exactness of desired result is examined in this module. Each unit of code is verified before integrating it in to complete software. After the successful testing is produce the accurate result.

7. Performance Evaluation

Measuring the performance of smart surveillance systems is a very challenging task [5] due to the high degree of effort involved in gathering and annotating the ground truth as well as the challenges involved in defining metrics for performance measurement. Like any pattern recognition system, surveillance systems have two types of errors:

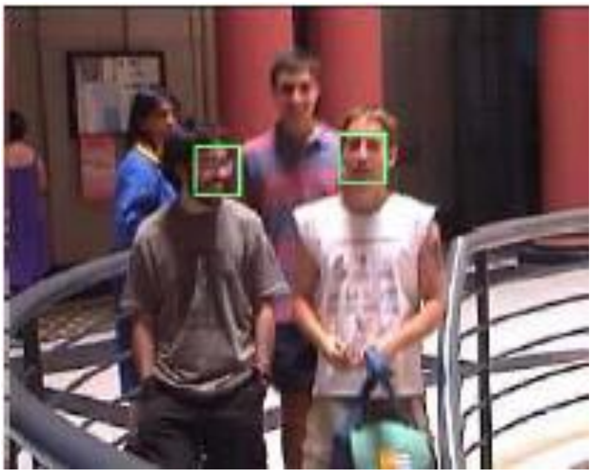
False Positives: These are errors that occur when the system falsely detects or recognizes a pattern that does not exist in the scene. For example, a system that is monitoring a secure area may detect motion in the area when there is no physical object but rather a change in the lighting.

False Negatives: These are errors that occur when the system does not detect or recognize a pattern that it is designed to detect. For example, a system monitoring a secure area may fail to detect a person wearing clothes similar to the scene background.

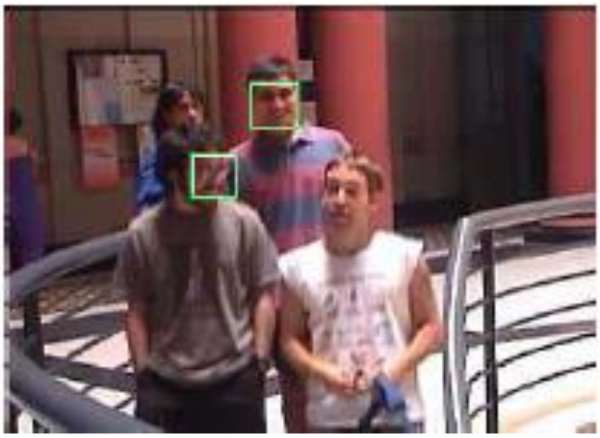
In this section, we present the various steps in evaluating an application like the real-time tracking of multiple human faces. The ultimate goal is to obtain good close-up head shots of people walking through the monitored space. The quality of the close-up face clips is a function of the accuracy of a number of underlying components.

8. Experimental Results

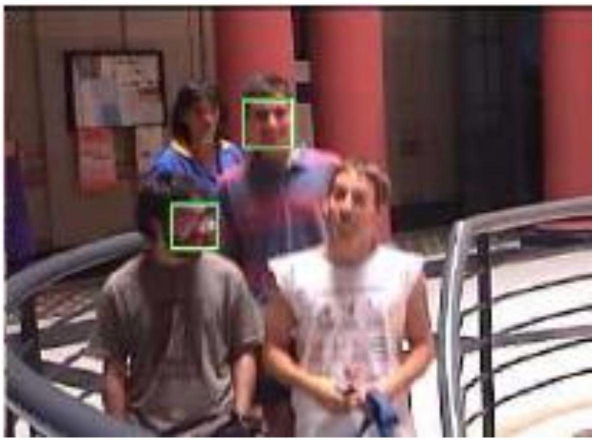
Performance of the proposed real-time tracking of face authentication system is evaluated for 25 subjects using a camera with a resolution of 160x120. Since humans rarely sit perfectly for a long time, motion information is used to estimate the face region. The estimated face region is not changed until there is a significant motion of the head in the video. The method tracks a multiple faces and is not sensitive to size of the face and lighting conditions. The location of the eyes and other facial regions are identified, and the feature vector is extracted from the face image only.



Frame 1



Frame 2



Frame 3

Frame 4

Figure 3: Experimental results of the proposed system on the video. Face detection windows are shown.

For testing the identity claim, we collected feature vectors from 10 face images of the subject with variation in size and tilt of the face one month after collecting the training data, and the average confidence value is estimated. The average confidence value is used to accept or reject the identity claim. In our experiment the identity claim of a subject is accepted if the average confidence value is greater than the threshold 0.9, and in general the threshold can be determined from the experimental studies. The average confidence values for 10 test subjects of the experiment against 10 corresponding models. High confidence value along the diagonal indicates that the system accepts the correct identity claim and low values in the off-diagonal indicate that the system rejects all the incorrect claims. The confidence values can be used to estimate the resemblance of a subject with other subjects.

The performance of the authentication system is invariant to the size and tilt of the face, and is also insensitive to variations in natural lighting conditions. The system tests the identity claim of a subject in real time, and it gives an equal error rate of less than 1%. The face detection and feature vector extraction techniques are computationally inexpensive, and testing 10 feature vector of a subject in the corresponding model requires less than 30 msec on a Pentium machine at 500 MHz's

8. CONCLUSIONS

The work represents an attempt to acknowledge and account for the presence on proposed system outperformed similar systems in the task of robust, real-world, real-time tracking of multiple persons. Work has to be done to improve the speed of the system, mainly in the face detector and in the blob identification block (color and movement) which are the slowest parts of the system. Real-time video analysis provides with the ability to react to an activity in real-time, thus acquiring relevant information at much higher resolution. The long-term operation of such systems provides the ability to analyze information in a spatiotemporal context. As such systems evolve, they will be integrated both with inputs from other types of sensing devices and also with information about the space in which the system is operating, thus providing a very rich mechanism for maintaining situation awareness. Further work can be done in terms of skin segmentation (automatic illumination adjustment, skin model and skin clustering), object tracking (overlapped object tracking and future position estimation) and human body detection.

REFERENCES

[1] A. K. Jain, R. Bolle, *"Biometric personal identification in networked society"* 1999, Norwell, MA: Kluwer.

[2] C.Hesher, A.Srivastava, G.Erlebacher, *"A novel technique for face recognition using range images"* in the Proceedings of Seventh International Symposium on Signal Processing and Its Application, 2003.

[3] HOME AFFAIRS JUSTICE, *"EU standard specifications for security features and biometrics in passports and travel documents"*, Technical report, European Union, 2006.

[4] ICAO, "Machine readable travel documents", Technical report, ICAO 2006.

[5] KLUGLER, D., *"Advance security mechanisms for machine readable travel documents, Technical report"*, Federal Office for Information Security (BSI), Germany, 2005.

[6] ICAO, "*Machine Readable Travel Documents",* Part 1 Machine Readable Passports. ICAO, Fifth Edition, 2003

[7] Riscure Security Lab, *"E-passport privacy attack"*, at the Cards Asia Singapore, April 2006.

[8] D. Monar, A. Juels, and D. Wagner, "*Security and privacy issues in e-passports"*, Cryptology ePrint Archive, Report 2005/095, 2005.

[9] ICAO, *"Biometrics Deployment of Machine Readable Travel Documents"*, Version 2.0, May 2004.

First Author Profile:

Mr. V.K. NARENDIRA KUMAR M.C.A., M.Phil., Assistant Professor, Department of Information Technology, Gobi Arts & Science College (Autonomous), Gobichettipalayam – 638 453, Erode District, Tamil Nadu, India. He received his M.Phil Degree in Computer Science from Bharathiar University in 2007. He has author more than 40 international journal article publications. He has authored or co-authored more than 60 technical papers and conference presentations. He is an editorial board member for several scientific international journals. His research interests are focused on Internet Security, Biometrics, Advanced Networking, Visual Human-Computer Interaction, and Multiple Biometrics Technologies.

Second Author Profile:

Dr. B. SRINIVASAN M.C.A., M.Phil., M.B.A., Ph.D., Associate Professor, PG & Research Department of Computer Science, Gobi Arts & Science College (Autonomous), Gobichettipalayam – 638 453, Erode District, Tamil Nadu, India. He received his Ph.D. Degree in Computer Science from Vinayaka Missions University in 11.11.2010. He has author or co-authored more than 30 international journal article publications. He has authored or co-authored more than 70 technical papers and conference presentations. He is a reviewer for several scientific e-journals. His research interests include automated biometrics, computer networking, Internet security, and performance evaluation.