Event-Handling Based Smart Video Surveillance System

Fadhlan Hafiz fadhlan_hafiz@yahoo.co.uk
Faculty of Engineering
International Islamic University Malaysia
53100 Kuala Lumpur, Malaysia

A.A. Shafie aashafie@iiu.edu.my
Faculty of Engineering
International Islamic University Malaysia
53100 Kuala Lumpur, Malaysia

M.H. Ali hazratalidu07@yahoo.com
Faculty of Engineering
International Islamic University Malaysia
53100 Kuala Lumpur, Malaysia

Othman Khalifa khalifa@iiu.edu.my
Faculty of Engineering
International Islamic University Malaysia
53100 Kuala Lumpur, Malaysia

Abstract

Smart video surveillance is well suited for a broad range of applications. Moving object classification in the field of video surveillance is a key component of smart surveillance software. In this paper, we have proposed reliable software with its large features for people and object classification which works well in challenging real-world constraints, including the presence of shadows, low resolution imagery, occlusion, perspective distortions, arbitrary camera viewpoints, and groups of people. We have discussed a generic model of smart video surveillance systems that can meet requirements of strong commercial applications and also shown the implication of the software for the security purposes which made the whole system as a smart video surveillance. Smart surveillance systems use automatic image understanding techniques to extract information from the surveillance data and handling events and stored data efficiently.

Keywords: Smart system, Human detection, Video surveillance, Object classification, Event handling

1. INTRODUCTION

Many applications of smart camera networks relate to surveillance and security system. Although applications of smart camera networks are fairly broad, they rely on only a few elementary estimation tasks. These tasks are object detection, object localization, target tracking, and object classification. Many surveillance methods are based on a general pipeline-based framework;

moving objects are first detected, then they are classified and tracked over a certain number of frames, and, finally, the resulting paths are used to distinguish “normal” objects from “abnormal” ones. In general, these methods contain a training phase during which a probabilistic model is built using paths followed by “normal” objects [1].Smart video surveillance systems achieve more than motion detection. The common objectives of smart video surveillance systems are to detect, classify, track, localize and interpret behaviors of objects of interest in the environment [2]. Some countries have implemented cabin video-surveillance systems which require pilot’s involvement in their operation as there is no intelligence built in them. Moreover, after 20 minutes of surveillance, in all such non-automated vigilance systems, the human attention to the video details degenerate into an unacceptable level and the video surveillance becomes meaningless. Thus there is an increasing demand for intelligent video surveillance systems with automated tracking and alerting mechanism [3].

Relevant work in this area include shape-base techniques which exploit features like size, compactness, aspect ratio, and simple shape descriptors obtained from the segmented object[4, 5].The smart camera delivers a new video quality and better video analysis results, if it is compared to existing solutions. Beside these qualitative arguments and from a system architecture point of view, the smart camera is an important concept in future digital and heterogeneous third generation visual surveillance systems [6]. Similarly, visual object classification is a key component of smart surveillance systems. The ability to automatically recognize objects in images is essential for a variety of surveillance applications, such as the recognition of products in retails stores for loss prevention, automatic identification of vehicle license plates, and many others.

In this paper, we address a simplified two-class object recognition problem: given a moving object in the scene, our goal is to classify the object into either a person (including groups of people) or a vehicle. This is a very important problem in city surveillance, as many existing cameras are pointing to areas where the majority of moving objects are either humans or vehicles. In our system, this classification module generates metadata for higher-level tasks, such as event detection (e.g., cars speeding, people loitering) and search (e.g., finding red cars in the video). We assume static cameras, and thus benefit from background modeling algorithms to detect moving objects [7]. Wolf et al. identified smart camera design as a leading-edge application for embedded systems research [8].In spite of these simplifications, the classification problem still remains very challenging, as we desire to satisfy the following requirements:

(a) Real-time processing and low memory consumption
(b) The system should work for arbitrary camera views
(c) Correct discrimination under different illumination conditions and strong shadow effects
(d) Able to distinguish similar objects (such as vehicles and groups of people).

Our approach to address these issues consists of three elements [7]:

(a) Discriminative features,
(b) An adaptation process, and
(c) An interactive interface

We also have defined 3 challenges that need to be overcome:

The multi-scale challenge

This is one of the biggest challenges of a smart surveillance system. .Multi-scale techniques open up a whole new area of research, including camera control, processing video from moving object, resource allocation, and task-based camera management in addition to challenges in performance modeling and evaluation.

The contextual event detection challenge
This challenge is mostly on using knowledge of time and deployment conditions to improve video analysis, using geometric models of the environment and other object and activity models to interpret events, and using learning techniques to improve system performance and detect unusual events.

The large system deployment challenge
It has several challenges include minimizing the cost of wiring, meeting the need for low-power hardware for battery-operated camera installations, meeting the need for automatic calibration of cameras and automatic fault detection, and developing system management tools.

2. METHODOLOGY

2.1 Camera Selection
Since the multi-scale challenge incorporates the widest range of technical challenges, we present the generic architecture of a multi-scale tracking system. The architecture presented here provides a view of the interactions between the various components of such a system. We present the concepts that underlie several of the key techniques, including detection of moving objects in video, tracking, and object classification..

Selections of the surveillance features are very important for a smart surveillance software or smart surveillance network. Camera and its components selection depends on the users. We have classified the camera based on the following criteria summarized in Table 1 and our selected hardware in Table 2.

Area		Complexity		Number of Cameras		Environment		Types of Camera	
Small	Large	Simple	Complex	Single	Multiple	Day/Bright	Night/Dark	IP based	Smart

Table 1: Classification of the camera

Components	Description	Quantity
1.Outdoor Box Camera	1/3" Sony Super HAD, 380TVL / 0lx / 3.6mm / 22IR LEDs / 12VDC	3
2.Dome Camera	1/3" Sony Super HAD	1
3. DVR Card	4 Channel 3rd Party DVR Card	1
4. Cable	Lay coaxial cable	50 ft
5.DVR	DVR configuration	1
7.Main Server	High Speed CPU	1

Table 2: The hardware used for our system

2.2 Camera Placement
We have developed our software to detect the movements and display features on the monitor as we want to view. A way to use the sun positions to determine the focal length, zenith and azimuth angles of a camera. The idea is to minimize the re-projection error, that is, minimize the distance between the sun labels and the sun position projected onto the image plane. If we have several observations, we can know the solution that gives the best data in a least-squares sense as in the following

$$\min_{f_c,\theta_c,\phi_c} \sum_{i=1}^{N} ||\mathbf{p}_i - \mathcal{P}(R^{-1}\mathbf{s}_i)||^2, \qquad \text{where } \mathbf{p}_i = \begin{bmatrix} u_i & v_i \end{bmatrix}^T$$

Where, P_i is the projection operator. Our goal is to find the camera parameters (fc, θc, Φc) that will best align the sun labels pi with the rotated sun position. The following equation is used to find those parameters and summarized in Table 3:

$$\min_{f_c,\theta_c,\phi_c} \sum_{i=1}^{N} \left(u_i - \frac{-f y'_{s,i}}{x'_{s,i}} \right)^2 + \left(v_i - \frac{f z'_{s,i}}{x'_{s,i}} \right)^2 ,$$

Name	Symbol	Default Value
Focal Length	fc	1000 px
Number of Image	N	20
Camera Zenith	θc	90
Camera Azimuth	Φc	0

Table 3: Necessary camera parameters

Camera positioning is very important for clear video image. Our test camera positioning is shown in the following figures. Camera angles are very important parameters to be considered. Depending on the application we can choose either 45 degree or 90 degree angle's camera. The focus point will determine the area under surveillance. Figure 1 and 2 bellow shows the front view and side view of the testing area. In this area we used four different cameras.1 is dome camera and the rest 3 are box cameras.

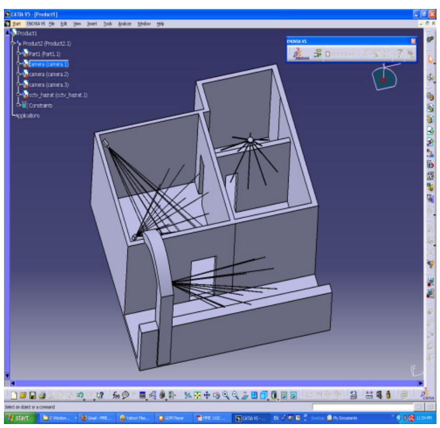

Figure 1 : Front view of the test area

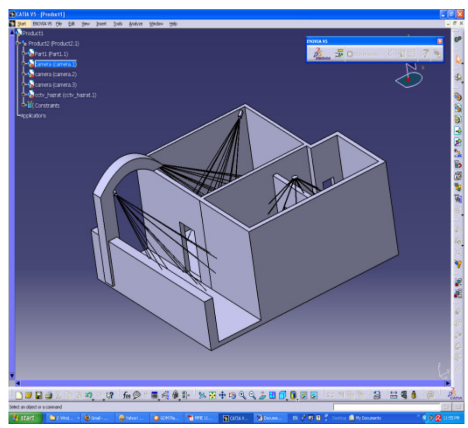

Figure 2 : Side view of the test area

2.3 Smart camera video sequences handling

Figure 3 shows the internal structure of our smart surveillance system. From the figure we can see how it works in the practical field. First the video will be recorded and then will be saved in the specific destination in hard drive, next it will use the object detection algorithm, after that it will follow the multi object tracking and classification method. Using this method we can sort the types of the objects detected by the camera. Finally Event classification will be used to index the data in the index storage from where we can retrieve the active object efficiently.

2.4 PC and software selection

The PC used for this project is Intel Quad Core Q9550, 2.83GHz with 4GB of memory and the software used for the system development are:

- Windows Vista.
- Microsoft Visual Studio 2008
- SQL Server Database
- AForge Image Processing Library (open source)

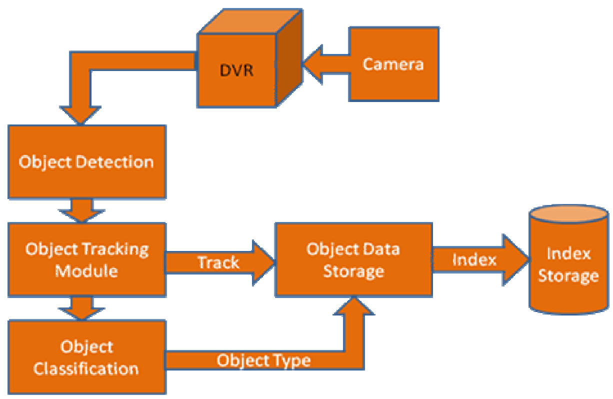

Figure 3: Overall internal structure of smart surveillance system

3. DETECTION, EVENT CLASSIFICATION AND DATA STORAGE

3.1 Detection algorithm

The overall human detection technique proposed in this system is discussed in details in [9] and it can be classified into two parts, (1) Image Pre-processing and (2) Segmentation. Image pre-processing includes frame processing, foreground segmentation, and binarization. While Segmentation includes Shadow removal, morphological operation, noise removal, and size filtering. The overall steps employed are summarized in Figure 4 and also illustrated with example in Figure 5.

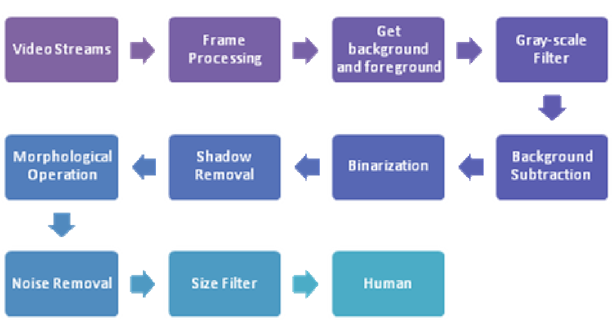

Figure 4: Human detection process flow

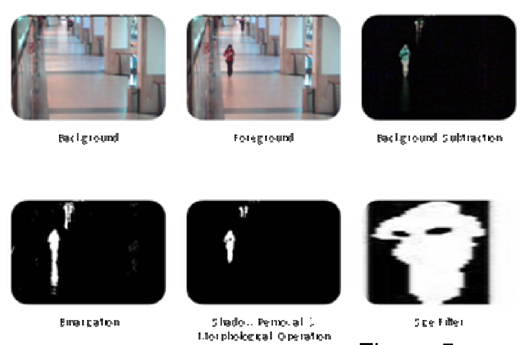

Figure 5

Graphical overview

3.2 Event classification

This classification system classifies the events that occur within a space that may be monitored by one or more cameras. The classification and associated information data processing are carried out in real-time. The following is descriptions of few elements in event classification:

- Event: An event is defined as any detection, movement or event that occurs within the camera's secured area.
- Event ID: This is a unique number which identifies a specific event. Main index to Video ID and Human ID. Used in *Indexing and Data Retrieval*
- Time: Time at which the event occurs.
- Location: Location where the event occurs
- Video ID: This is a unique number which identifies a specific video recorded active camera and is the index to the physical video storage information.
- Human ID: This is a unique number which identifies a specific human who appears within camera focus area. Index to the image, appearance information, and tracking data

3.3 Data Storage System

The storage system in this surveillance system comprises of two components: (a) physical storage and (b) database. Video data from the cameras are saved into the hard disk and this falls into physical storage category. The video files are saved in frames, and not in usual video format like *.wmv* or *.avi* format. This is because the system, which is developed in .NET framework, cannot save the video data while analyzing the video streams for detection purpose. This feat is applicable in COM but not in .NET. As a solution, we copy the image frame, one at a time, and save the images on the hard disk just before the detection class processes the images. These images are saved in JPEG format. All the data associated to that video such as the physical location of the images, the time stamp, duration, and frame speed (fps), are saved in the video database and each video is assigned a unique ID as the key for that information. This information is saved in the database using SQL Server Database.

3.4 Data Indexing and Retrieval

This is the core and the brain of this surveillance system. Those IDs which were assigned earlier in real-time, such as Event ID, Video ID and Human ID are used to index all the data in the database and for automated data retrieval when requested by the user. Every ID in this system is linked to each other so that user can access all necessary information like human appearance, time, track, and video data in one comprehensive and inter-related search. The flow of *Data Indexing and Retrieval* is illustrated in Figure 6. Once process objects are identified, the data extracted from objects are stored together with the unique ID which indexed earlier into database. Then the data retrieval will handle the request for data using the assigned ID as the key for retrieval of data from database and then retrieve all the associated data.

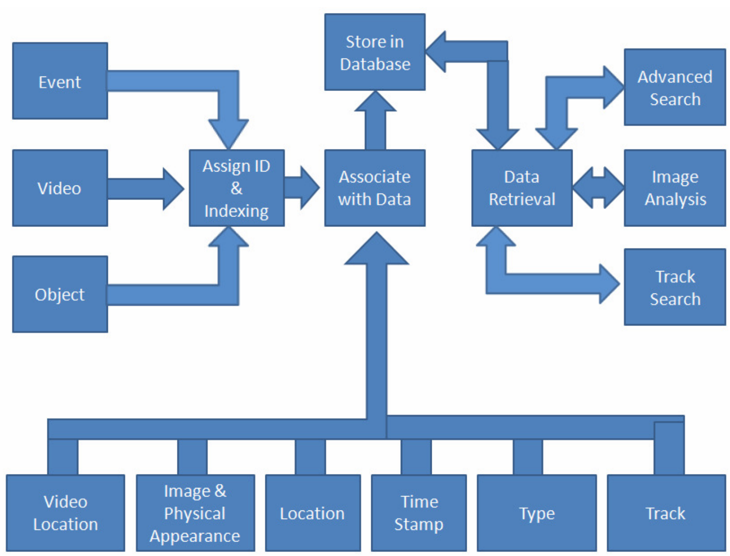

Figure 6 : Data Indexing and Retrieval process flow

4. SYSTEM OVERVIEW

Our developed system gives more advance result compared to the existing camera system. It has varieties of displaying options which will easily help to analyze data in any time in the region of the cameras. The functions can be divided into 5 categories as follows:

1. Surveillance Mode

In this mode, user can connect to all available cameras, and change the surveillance settings for each camera, such as motion detection, human detection, security settings and secured parameters. This mode also enables administrative control such as system lock, password management and profile selection. It summarizes all the running components and the result of each component such as number of events, number of human detected, camera frames per second, and intruder detection. It also has full screen mode. This mode is illustrated in Figure 7 and Figure 8.

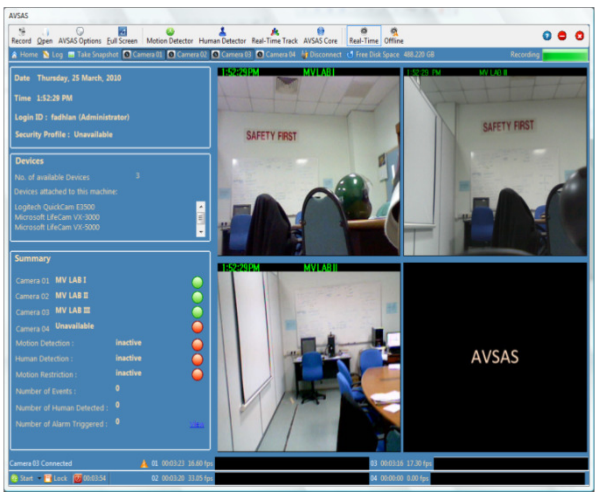

Figure 7: Surveillance in windowed mode

Figure 8: Surveillance in full screen

2. Core Databases

All events and videos are saved into the hard disk and their details and locations are indexed into the event and video database. This database will keep track of any object such as human, detected by the surveillance systems and link the events to corresponding video. Human and video database are shown in Figure 9 and 10 respectively. This will enable the user to easily search for any event in interest such as human at specific time and place automatically without the hassle of looking through terabytes of video data.

3. Advanced Search

Advanced Search is a sophisticated search engine for searching any event, human or object at any time and place. The resulting event from this search can be used to directly open the corresponding video in playback mode or to view the track of the requested human. Each event bears a unique ID and unique time stamp so that a user can directly access the video of the event using the *Data Indexing and Retrieval* system discussed earlier. *Advanced Search* with human images result is illustrated in Figure 11

4. Playback Mode

Playback Mode as in Figure 12 can carry out several video analysis such as automated video searching according to time constraints. It has playback feature from normal surveillance system such as play, pause, stop, fast forward, and rewind with adjustable magnitude. The user can open up to maximum 4 video at a time where each video window have the same time stamp but corresponds to different camera locations.

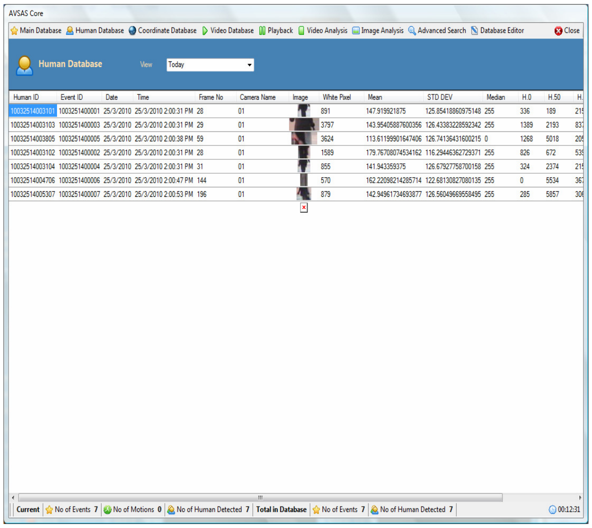

Figure 9: Human Database

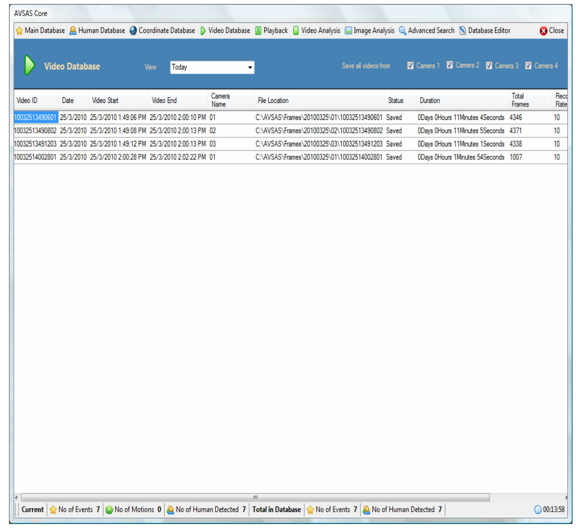

Figure 10: Video Database

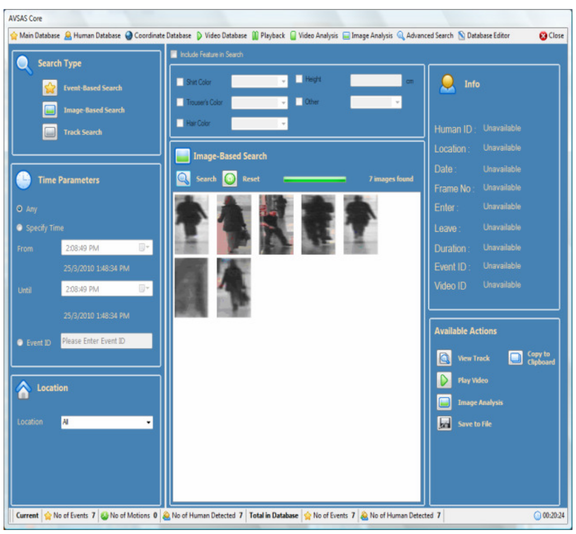

Figure 11: Advanced Search

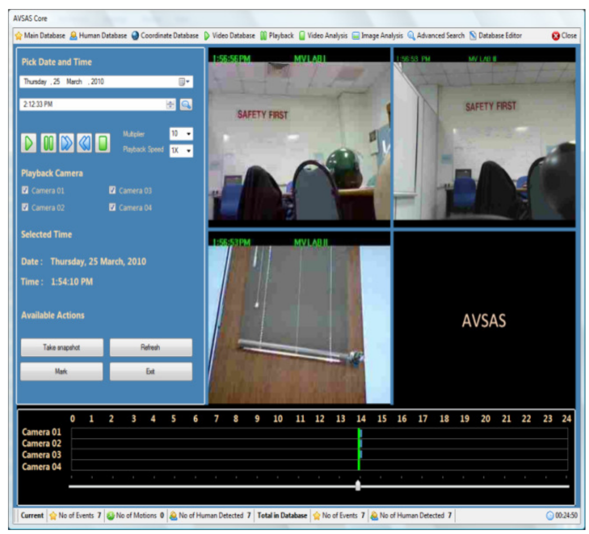

Figure 12: Playback Mode

5. Video and Track Analysis

Video Analysis is linked to *Advanced Search* and it handles the playback of the video containing the event searched in *Advanced Search*. Any video opened in this mode will list all the events contained in the video as indexed by the system beforehand. User can click any event listed to go directly to the video at the time it happens. The video controls of this mode are much less the same with *Playback Mode*. This mode is illustrated in Figure 13. While *Track Analysis* enables user to view previously recorded track of any human that appears in the video. This mode is illustrated in Figure 14.

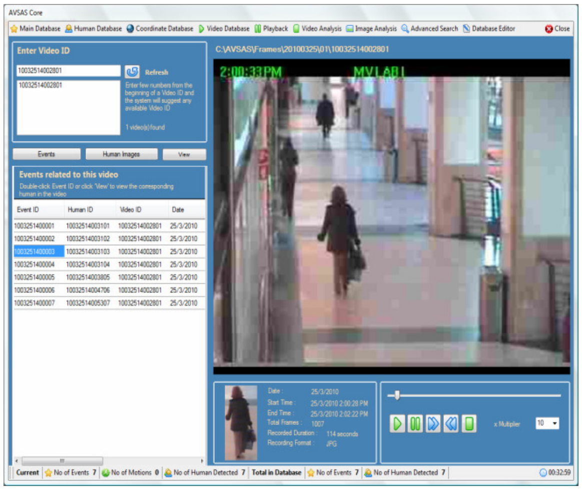

Figure 13: Video Analysis

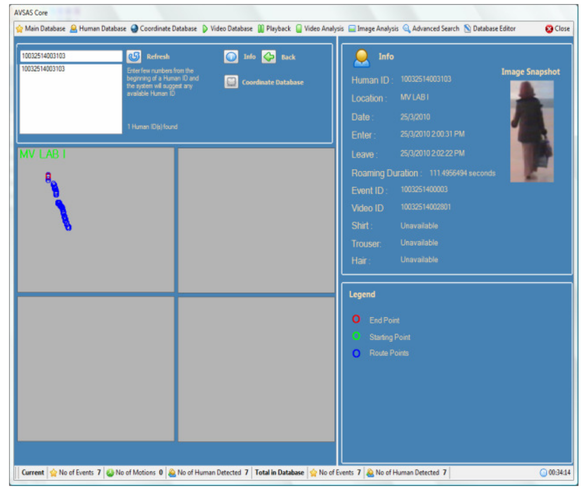

Figure 14: Track Analysis

5. RESULT AND DISCUSSION

Our developed software works very efficiently with the real time video surveillance system. The software is a great tool to classify and track the movement of any object under the camera's secured area. The software has details display mode like time, place, human or non-human, how many object, trajectory path of the moving object, video retrieval using *Data Indexing and Retrieval,* playing past videos from the hard drive and so on. The details of those features have been discussed in details in the system overview section. This software is based on the following key video analysis technologies:

- Human and Object Detection: This software can detect moving human and/or objects in a video sequence generated by a static camera. The detection techniques are invariant to changes in natural lighting, reasonable changes in the weather, distraction movements and camera shake. Several algorithms are available in this software including adaptive background subtraction with healing which assumes a stationary background and treats all changes in the scene as objects of interest and salient motion detection [10] which assumes that a scene will have many different types of motion, of which some types are of interest from a surveillance perspective. The result of human detection is illustrated in Figure 15

- Human and Object Tracking: This software can track the position of multiple objects as they move around a space that is monitored by a static camera as illustrated in Figure 16.

- Object Classification: This software uses various properties of an object especially human including shape, size and movement to assign an event and object type label to the objects. Our system fulfills the following criteria for *Advanced Search.*

Searching capability and smart indexing of data made this software handles data efficiently and ease the management of large video data. The searching itself comprises of combinations of logics as follows:

- Search by *Event Type* retrieves all event matches the requested type
- Search by *Time* retrieves all events that occurred during a specified time interval.
- Search by *Location* retrieves all objects within a specified area in a camera.

- Search by *Image* retrieves all images of human or objects that has appeared within the camera's secured area.
- Joint Search combines one or more of the above criteria as specified by the user

Figure 15: Result of human detection carried out by this system

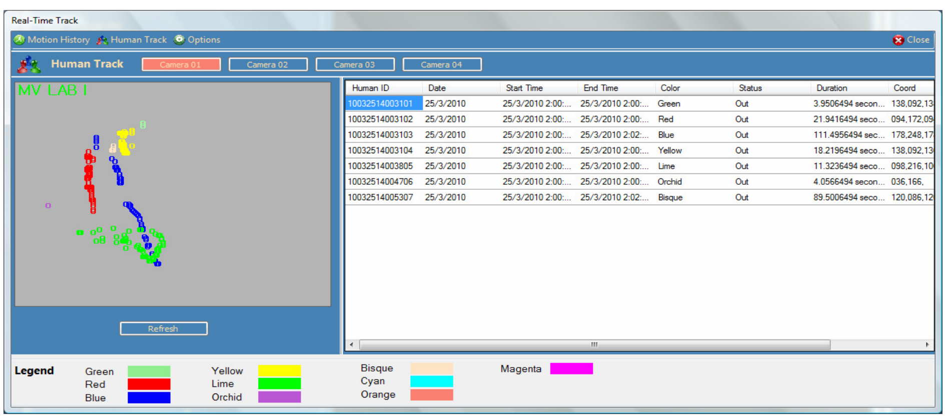

Figure 16: Multiple human tracking

6. CONCLUSION AND FUTURE WORKS

We have presented our smart video surveillance software for the surveillance purposes. Also we have introduced a generic smart video surveillance systems model, which relates the computer vision algorithms. All these methods have been linked with the smart surveillance system. From the practical point of view we found the developed software is more effective compared to the traditional surveillance system as well as it has a details display mode which helps us to track the moving object in an easier way and smart data handling for indexing of massive video data.

Future Works: The *event classification* and *Data Indexing and Retrieval* system would be improved further to enable wider search criteria to be implemented such as human/object specific appearance search and also to enhance the detection subsystem to be able to detect more objects. The system also will be improved to be able to classify human behavior from human body posture, movement speed and frequency of appearances. Moreover, the partial body occlusion is the main drawback from using shape-based human detection in this detection class as the system cannot determine correctly the boundary of human body in case of occlusion. But we will overcome this problem using head detection to correctly determine the number of human presents in the scene and therefore can locate the bounding boxes accurately based on human-shape model. The code for the software will also be revised to implement *multi-threading* for better performance with multi-core processors.

7. REFERENCES

1. W. Hu, T. Tab, L. Wang, and S. Maybank, "A Survey on Visual Surveillance of Object Motion and Behaviors," IEEE Trans. Syst. Man Cybern˙-- Part C: App. and Reviews, vol. 34, no. 3, pp. 334–352, 2004.

2. D. M. Gavrila, "The Visual Analysis of Human Movement: A survey," Computer Vision and Image Understanding, vol. 73, 1999, pp. 82-98.

3. C. S. Regazzoni, V. Ramesh, and G. L. Foresti**, "**Distributed Embedded Smart Cameras for Surveillance Applications**",** *Proceedings of the IEEE*, 39(2), Oct 2006.

4. L. Brown, "View Independent Vehicle/Person Classification", *VSSN'04*, October 15, 2004, New York, USA.

5. B. Bose and E. Grimson, "Improving object classification in far-field video", CVPR, 2004.

6. Case Study by Seagate, "Smart Surveillance Systems Tap Reliability and Performance of Seagate SV35 Series Drives." CS509,October 2006.

7. L. Chen et al, "An Integrated System for Moving Object Classification in Surveillance Videos" IEEE Fifth International Conference on Advanced Video and Signal Based Surveillance.2008.

8. W. Wolf, B. Ozer, and T. Lv, "Smart Cameras as Embedded Systems", *IEEE Computer Science Journal* , 35(9):48–53, Sep 2002.

9. Fadhlan Hafiz et.al, "Human Detection For Real-Time Video Surveillance", In Proceedings of International Conference on Engineering Technology 2010, Kuala Lumpur, Malaysia, 2009

10. Tian, Y.I. Hampapur, "A Robust Salient Motion Detection with Complex Background for Real-Time Video Surveillance", In: IEEE Computer Society Workshop on Motion and Video Computing, Breckenridge, Colorado, January 5 and 6, 2005