The IBM Smart Surveillance System

Arun Hampapur, Lisa Brown, Jonathan Connell, Max Lu, Hans Merkl, Sharat Pankanti, Andrew Senior, Chiao-Fe Shu, Yingli Tian
19 Skyline Drive, Hawthorne, NY 10532
(914)-784-7440, arunh@us.ibm.com

ABSTRACT

The increasing need for sophisticated surveillance systems and the move to digital surveillance infrastructure has transformed surveillance into a large scale data analysis and management challenge. Smart surveillance systems use automatic image understanding techniques to extract information from the surveillance data. While the majority of the research and commercial systems have focused on the information extraction aspect of the challenge, very few systems have explored the use of extracted information in the search, retrieval, data management and investigation context. The IBM Smart Surveillance Systems demonstrates the capabilities of a fully integrated smart surveillance system, where the index data generated by the analytics is made searchable in real time.

1. INTRODUCTION

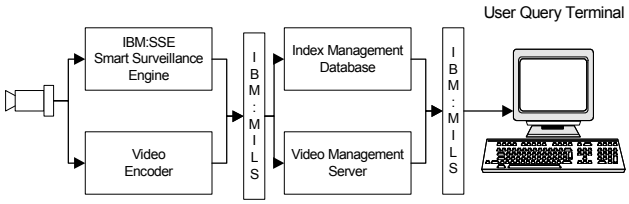

Figure 1: Block diagram of the IBM Smart Surveillance System. The video from the camera is simultaneously processed by the IBM:SSE (Smart Surveillance Engine) and encoded by the encoder. The index data generated by the IBM:SSE is loaded into a database server by IBM:MILS (Middleware for Large Scale Surveillance), which also provides the search functionality.

The IBM Smart Surveillance System has two key components, , 1) automatic video behavioral analysis and 2) data management for forensic retrieval of surveillance video. The IBM Smart Surveillance Engine (IBM:SSE) provides behavioral analysis based on automatic image analysis technologies. The IBM Middleware for Large Scale Surveillance (IBM:MILS) provides a set of services that support the management of surveillance data and content based retrieval of surveillance video. This document provides details on the functionalities of IBM:SSE and IBM:MILS

IBM: SSE Smart Surveillance Engine

The IBM Smart Surveillance Engine (SSE) is a software-only video based event detection technology which delivers "smart" video surveillance capabilities. The SSE is designed with the intent of making currently deployed surveillance systems "smart". The IBM SSE is based on the following key video analysis technologies [1,2,3].

- Object Detection: This set of technologies can detect moving objects in a video sequence generated by a static camera. The detection techniques are tolerant to changes in natural lighting, reasonable changes in the weather, distracting movements (like trees waving in the wind), and camera shake.
- Object Tracking: This set of technologies can track the shape and position of multiple objects as they move around a space that is monitored by a static camera. The techniques are designed to handle significant occlusions as objects interact with one another.
- Object Classification: These technologies use various properties of an object including shape, size and movement to assign a class label to the objects. Typical labels include, Person, Group and Vehicle.

The SSE provides the following functionalities.

- Real Time Video Based Alerts: These are alerts which depend solely on the movement properties of objects within the monitored space. Examples include 1) Motion Detection: 2) Directional Motion Detection 3) Abandoned Object Alert 4) Object Removal and 5) Intentional camera movement or blinding:
- Viewable Video Index (VVI): The SSE detects and tracks all moving objects within the cameras field of view. The SSE creates the VVI as a set of XML documents. The VVI encodes all the interesting activities in the video including, 1) Number of objects in a scene 2) Object Class, the current engine classifies objects in Single Person, Group of People and Vehicles, 3) Object appearance properties, including color, texture, shape, size and their changes over time, 4)Object movement properties, including position, velocity and trajectory, 5)Occlusion parameters when objects are in an occlusion, 6)Background changes due to changes in lighting and stopping of moving objects, 7)Event information: any events that may be flagged by the engine. The VVI index is also ideal for monitoring activities over a mobile wireless device such as a PDA, because of its extreme low bandwidth requirements (~10MB/hour). The index encodes

0-7695-2158-4/04 $20.00 (C) 2004 IEEE

all the activity that occurs in the video in terms of evolving background models, evolving object models and motion trajectories of each of the moving objects. The information contained in the VVI can be used by the an application to render the activity in the video independent of the original video stream.

IBM:MILS Middleware for Large Scale Surveillance:

MILS provides the data management services needed to build a large scale smart surveillance application. While MILS builds on the extensive capabilities of IBM's Content Manager and DB2 systems, it is essentially independent of these products and can be implemented on top of 3rd party relational databases. MILS provides the following functionalities

- Converting the VVI to Relational Tables: MILS is capable of taking the VVI generated by the IBM:SSE and inserting the data index into a relational database, thus allowing for SQL based querying of the index data by an application.
- Query Services for Surveillance Data: MILS provides applications with a set of services for querying the surveillance data. The queries supported by MILS cover a wide range of commonly used queries. The following are types of queries that will be supported by MILS. Each of these query types is illustrated by an example. The basic unit of storage in the database is a continuous movement event, for example, a car entering the camera field of view at time T1 and moving continuously to leave the field of view at T2, would be represented as a basic unit called an activity. All query responses are a set of activities.
    1. Time Queries: Show all activities between (7 PM Apr 17 2004) and (9 PM Apr 17 2004).
    2. Object Size Queries: Show all activities where the object size O is (S1 < O < S2)
    3. Object Class Queries: Show all activities with objects of class C1. Current set of classes are limited to Vehicles, Single Person and Group.
    4. Object Motion Queries: Show all activities where the object moved with speed V where, (V1 < V < V2) or where an object moved in a direction D (specified as an angle in the image plane).
    5. Context-based object content similarity Queries: Show all activities where there were "blue cars" or "cars similar to this car –here the user specifies an example car through an image".
    6. Queries based on region/location of interest: Show all activities that occurred within a specified polygonal region.
    7. Combination Query: A query which combines the one or more of the above specified query types.
    8. Interactive object/event classification based on relevance feedback
    9. Data Mining Services: MILS will support the following data mining applications.
        - Site activity summarization
        - Site activity pattern discovery
        - Automatic abnormal event/object detections

Description of Demo at CVPR 2004

The CVPR 2004 demo shows "live" query and retrieval of activities in the parking lot of IBM's TJ Watson Research Center. The user will makes queries which retrieve activities. Example queries include retrieve 1) the last 100 activities 2) all cars over a specified time interval 3)all people over a specified time interval 4)large cars and trucks over a specified interval and additional queries from the above list. Once an activity is retrieved from the database, the user will be able to play the associated clip of video. Depending on the privilege of the user, the video clip will be show either in its original form or in the enhanced privacy mode.

REFERENCES

1. A. Hampapur, L. Brown, J. Connell, S. Pankanti, A.W. Senior, and Y-L Tian , Smart Surveillance: Applications, Technologies and Implications, IEEE Pacific-Rim Conference On Multimedia, Singapore, Dec
2. PeopleVision: Demo videos of on going work at IBM research: www.research.ibm.com/peoplevision
3. A. Senior, A. Hampapur, Y-L. Tian, L. Brown, S. Pankanti, and R. Bolle. Appearance Models for Occlusion Handling. Proceedings of Workshop on Performance Evaluation of Tracking and Surveillance (PETS2001), December 2001.

.