2011 International Conference on Digital Convergence (ICDC 2011)

Distributed Video Surveillance Using Mobile Agents

Praveen Kumar
Department of Computer Science, GRIET
Hyderabad, India
Praveen.kverma@gmail.com

Ankush Mittal
Department of Computer Science, COER
Roorkee, India
dr.ankush.mittal@gmail.com

Amit Pande
Department of Computer Science, UC Davis,
California, USA
amit@cs.ucdavis.edu

Abhishek Mudgal
Institute for Transportation, Iowa State University
Ames, IA, USA
abhishek@iastate.edu

Prasant Mohapatra
Department of Computer Science, UC Davis,
California, USA
prasant@cs.ucdavis.edu

Abstract— **This paper presents a framework for distributed multimedia processing using mobile agents in Distributed Video Surveillance (DVS) network architecture. The increased computational costs due to video processing tasks like object segmentation and tracking are shared by the cameras and a local base station called as Processing Proxy Server (PPS). However, in a distributed scenario like traffic surveillance, where moving objects is tracked using multiple cameras, the processing tasks needs to be dynamically distributed. This is done intelligently using mobile agents by migrating from one PPS to another for tracking an individual case object and transmitting required information to the end users. Although the authors propose a specific implementation of DVS systems, the general ideas in design of such systems exemplify the way information can be intelligently transmitted in any ubiquitous multimedia applications along with the use of mobile agents for real-time processing and retrieval of content from real-time video streams.**

Keywords- video surveillance; mobile agents; distributed processing;

I. INTRODUCTION

The demand for distributed multimedia services like distributed video surveillance [1, 2], real-time remote laboratory [3, 4] etc., are rapidly increasing and the expectation of quality for these services is becoming higher and higher. Even though the technology is continuously progressing and pushing up the bandwidth limit and reducing the transmission, storage cost, still the channel bandwidths and storage capabilities are limited and relatively expensive in comparison with the volume of multimedia data required in such applications. The rapid emergence of wireless networks and proliferation of networked digital video cameras have favorably increased the opportunity of developing a low cost distributed Video Surveillance (DVS) framework.

Availability of network IP-based digital video camera improved accessibility for both collection and distribution of video data. Despite their significant advantages, IP cameras with wired networking still suffered with limited flexibility and high installation cost. But the development of wireless infrastructure has provided great flexibility in connecting IP cameras to a wireless network with considerable reduction in infrastructure, maintenance and operational cost. Many commercial companies now offer IP-based surveillance solutions. Company's likeSony and Intel have designed equipments like smart cameras; Cisco provides many networking devices for video surveillance, Texas Instrument DSPs and Xilinx Spartan series of FPGAs can be used to implement many multimedia compression modules such as motion detection and object tracking etc. This has accelerated the latest step in the evolution of video-surveillance systems i.e. migration to digital IP-based surveillance and more recently to wireless interconnection network so as to increase the scalability of large scale deployment of video surveillance systems at reduced cost.

This has led to the evolution of third generation of surveillance system aimed towards the design of large distributed and heterogeneous (with fixed, PTZ, and smart cameras) surveillance systems for wide area surveillance like monitoring movement of military vehicles on borders, surveillance of public transport etc. For example the Defense Advanced Research Projection Agency (DARPA) supported the Visual Surveillance and Monitoring (VSAM) project [5], whose purpose was to develop automatic video understanding technologies that enable a single human operator to monitor behaviors over complex areas such as

This work is partially supported by the National Science Foundation under Grant #1019343 to the Computing Research Association for the CIFellows Project.

battlefields and civilian scenes. The usual design approach of these vision systems is to build a wide network of cooperative multiple cameras and sensors to enlarge the field of view. From an image processing point of view, they are based on the distribution of processing capacities over the network and the use of embedded signal processing devices to give the advantages of scalability and robustness potential of distributed systems. Intelligent cameras are a novelty in these surveillance systems. They perform a statically defined set of low-level image processing operations on the captured frames onboard to improve the video compression and intelligent host efficiency [6].

However, most video processing and analysis in current surveillance systems is executed at a central host using standard workstation racks. For example, in traffic surveillance, typical video analysis tasks include video compression, detection of stationary vehicles and wrong-way drivers, and computation of traffic statistics such as average speed, lane occupancy, and vehicle classification [7]. The system designer statically assigns the analysis of a camera's video input to a specific workstation. Modifying or reconfiguring this assignment during the surveillance system's operation is difficult. However, such static surveillance system configuration is no longer feasible in the ongoing paradigm shift from a central to a distributed control surveillance system. The main motivation for this shift is increasing the surveillance system's functionality, availability, and autonomy. Such surveillance systems can react autonomously to changes in the system's environment and to detected events in the monitored scenes. Therefore, the system architecture must support reconfiguration, migration, quality of service, and power adaptation of analysis tasks.

This paper addresses the requirement of reconfiguration, migration, quality of service, and dynamic adaptation of analysis tasks to support large scale ubiquitous deployment of video surveillance system. Use of mobile agents is illustrated with the example of tracking a moving vehicle migrating into different camera regions for traffic surveillance purpose.

II. Generalized DVS Network Architecture

Figure 1 shows a DVS network architecture, where there are several (over a thousand) video cameras distributed over a wide area, with smaller groups/cluster of cameras under a local base station called Processing proxy server (PPS) deployed at a monitoring position. These cluster of cameras/servers are connected with each other through the backbone internet network and also with different users like a monitoring station with security manager viewing many areas simultaneously, a mobile guard with a cell phone/PDA for local action or to a video recording server for data storage and retrieval tasks. Note that these users have different requirements in terms of image quality and preference for visual objects. The network resources available with each user are also different (like mobile guard has a low bandwidth GSM connection etc). The base station must send the appropriate video frames at appropriate level of compression depending on user preference and bandwidth availability. An example in case is the Chicago or New York Police Department used video surveillance cameras. In Chicago, the police capture feeds from around 10000 cameras mounted in different locations across the city. Different cameras inputs are transmitted to the PPS. There are few PPS across the city at some centers where the video data is being fed and inputted. Now, the PPS are all connected to the internet using a high bandwidth connection. The police officials near the PPS may be able to directly access videos using a dedicated network. However, in case of a patrol party, they may use their laptops or mobile devices to create an authenticated communication channel with multiple PPSs and request for the video feeds through internet. In this case (which is more likely the case), the video bandwidth will be varying and we need to ensure the delivery of requested feeds to the patrol party.

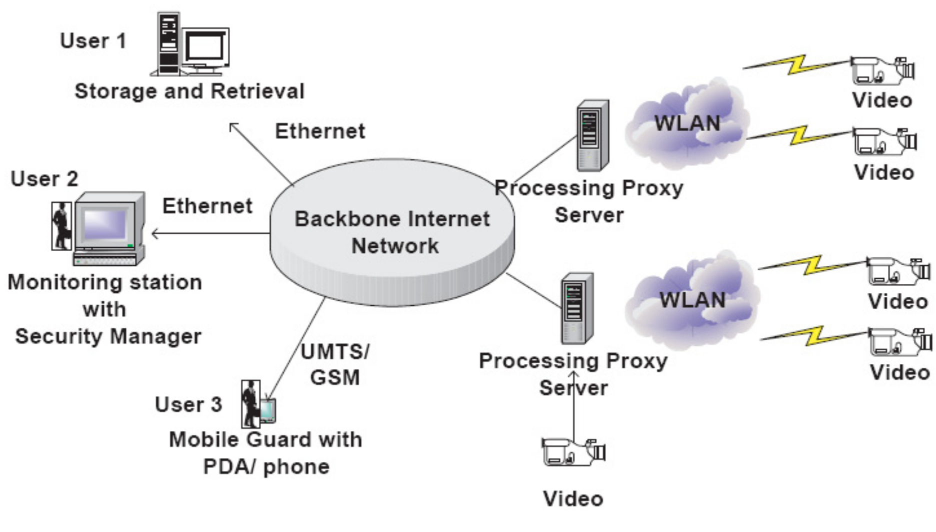

Figure 1. A Generalized DVS network architecture

III. Distributed Processing Framework using Mobile Agents

In addition to the sensing and communication tasks, there are several video surveillance tasks like object detection, object recognition, tracking, video compression etc., and it is usually not feasible to do all the computations at the camera end. In that case, the cameras do the sensing and some low-level processing and transmit the video to the PPS which does the other video surveillance and communication tasks. If smart cameras are deployed which are equipped with powerful DSP processors then some of the surveillance tasks can be done on the camera itself. But even then, some of the tasks that take input from multiple cameras have to be done on the PPS. For example, to extract different attributes of a vehicle of interest viz. shape, size, speed, direction etc. would require different cameras to be set up in a particular view angle and position. For instance, camera can be placed facing the direction of the vehicle in order to record front size and shape whereas cameras placed at an aerial position may be useful in measuring speed and direction of vehicle. Thus, we treat the PPS together with the group of connecting cameras as a single entity in the design of the framework for executing the sensing, surveillance, communication tasks for a particular location.

Figure 2 shows the distribution of processing and task execution on a PPS/Camera unit with the use of mobile agents. There are some low level pre-processing modules that are done on each camera's video input. The DWT module enables us to get a scalable bitstream which is later used for CEZW compression for transmission over different bandwidth network. The Object Recognition (OR) module has a list of predefined objects and this module tries to identify them. For example – in case of police department cameras on street we want to check for 3 objects – human beings, cars and unidentified moving objects. Human beings have the attribute such as dress color or skin color or specks etc., which are extracted by the VO module when it detects a human subject. Similarly car has attributes such as car-type (sedan or SUV or compact), color, number plate etc which the OR module tries to capture as much as possible and store in VO (Visual Object) features.

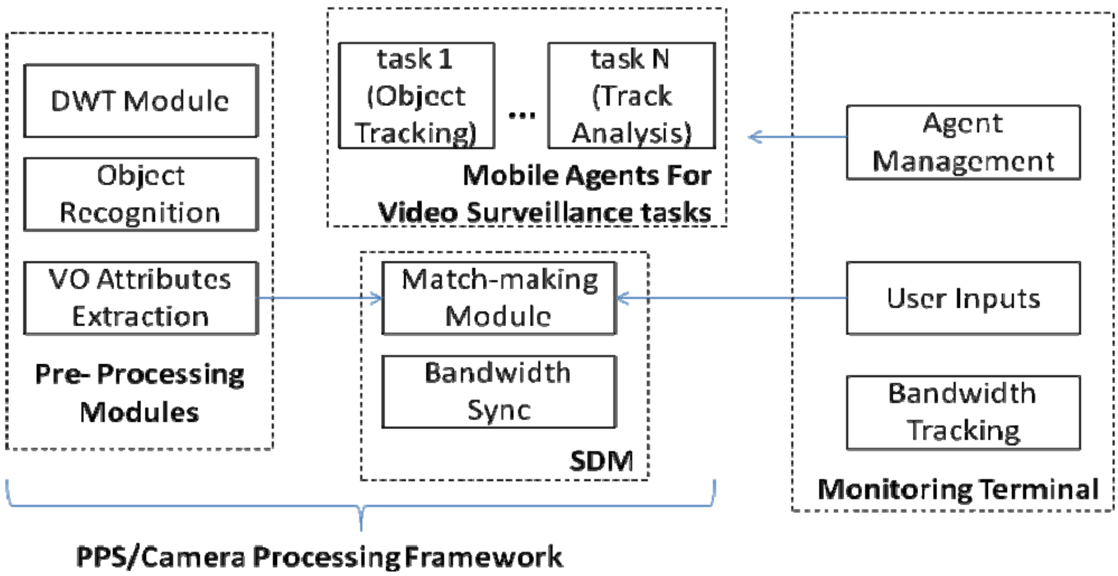

Figure 2. DVS software architecture consisting of camera/PPS processing framework using mobile agents.

When a user on the monitoring terminal, for example a policeman in our Chicago police example, observes a suspected car and wants to track it, a mobile agent is generated for tracking that particular object on the Camera/PPS where it was first identified and the parameters/attributes of the object is passed as the data part in the mobile agent. The SDM module on the PPS gets input from a number of cameras with different video contents. It has VO attributes and it checks for these attributes with those requested by the user. For example – the patrol party wants to find out a speeding pink car with IL number plate. SDM matches this with VO attributes and decides whether this camera output is of interest to this particular user or not. Thus, it looks for relevant VOs and transmits them to the end user according to their motion and importance.

Now in a distributed scenario, where different PPS/cluster of cameras are monitoring different regions of the area under surveillance, the object of interest like a moving vehicle migrates from one region to another. For example, if a vehicle is being tracked by PPS1 and it leaves its region and enters into the region of PPS2, then the object tracking task on PPS1 needs to be migrated to PPS2. We implement this dynamic configuration of tasks onto various PPS using a mobile agent system (MAS). In our MAS, agents represent surveillance tasks that can migrate dynamically between the PPS. Significant changes in the observed environment or in the available resources trigger a task migration. Mobile agents are most suitable for this distributed application because we can encapsulate each surveillance task within a mobile agent, which can then migrate between cameras. Thus, MAS supports autonomous operation of the surveillance tasks. Moreover, this approach is highly scalable and flexible

To determine the neighboring PPS/Camera unit to which the mobile agent should migrate, different approaches have been proposed in the literature [8, 9]. In [8], an algorithm to determine the neighboring video camera is discussed in the context of automatic human tracking system in a building. The neighbor camera nodes differ by the difference of view distance and view overlap of the video camera even if video camera's locations are same. The algorithm determines the neighbor camera node by the location and view distance of the video camera. In another approach [9], the author uses a master-slave approach for the tracked object handover. The master tracker identifies the object in its field of view and tracks its positions. As soon as the object enters a migration region, the master tracker creates slave trackers on every smart camera assigned to that migration region. The master tracker initializes these slave trackers with the object's identified features. When the slave tracker identifies the tracked object in its field of view, it terminates the master tracker and other slave trackers, and becomes the master tracker.

Figure 3 shows the schematic diagram of the internal blocks of mobile agent architecture designed in this case. The resource manager application which is equipped with a User Interface (UI) coordinates monitoring and control policies relating to the MA's. Active agent processes are discovered by the manager, which maintains and dynamically updates the list. The mobile code repository is a collection of binary executables of different surveillance tasks that are downloaded from the network processor to one of the DSPs on the cameras. The MAS module manages the agent's integration, communication, migration etc. with its environment.

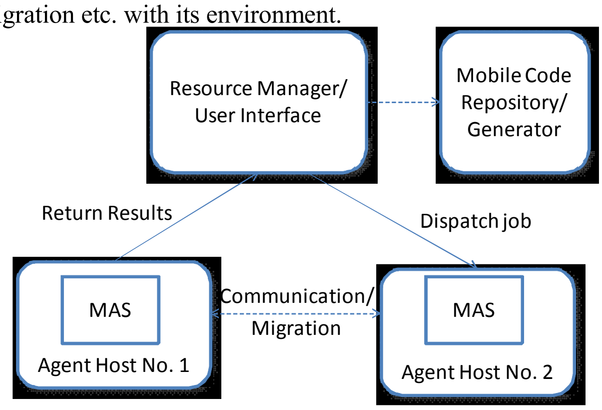

Figure 3: The Mobile agent based architecture

IV. Experimental Study - Traffic Surveillance

We illustrate the application of the above DVS framework for traffic surveillance and transportation systems. The following are some of the important areas where it can be applied.

- Estimating demand and traffic parameters at a given site for improving traffic operation: The increase in traffic demand has challenged the traditional methods of obtaining traffic parameters. The methods highlighted in this work would be useful in estimating traffic parameters in a more dynamic way.
- Tracking Red light running and other traffic violations and catching violators: Red-light running cameras have been successful in recording the licence plate of the offender but they would not be efficient in identifying any crashes leading to red-light running. While, in the proposed system, once the offender (or violator) is identified, the system can track the violator making use of more resources than it would have done otherwise.
- Reconstructing crashes at high-crash locations: The proposed method would provide higher accuracy in identifying the micro-events leading to crash and after crash damages by making use of the system resources in an intelligent way.
- Managing traffic at traffic work-zones: Traffic workzones have been identified as locations where both traffic operation and safety are at stake. Workzones are also highly vulnerable to crashes and for that reason any violation at workzones are penalized more severely. This explains the importance of a surveillance system which is proactive in identifying the violations and alarming the workers working downstream of the road. In addition, the proposed system would also track the offender with greater efficiency and make a more accurate record of his or her traffic offenses.

Many of these tasks require cooperative and dynamic interaction between multiple cameras. For example, in a system for detecting traffic violations, when a violator is identified by the camera at the signal, it passes the information of the violator vehicle to the next camera and so forth. In this way the violator is tracked with greater accuracy. A camera kept at a high crash location may not be able to capture the full event. So, when it identifies a collision it tags the vehicles involved and passes this information in the attributes to the nearby cameras. This makes it easier to reconstruct crashes when the crash occurs outside the field of a single camera. In work zones, this method can be used for tracking late mergers.

As a part of pilot study, two cameras were positioned at the work zone (I-35N in Iowa, USA). One points to upstream traffic and another to tapered zone is a workzone where one of the lanes merges (here right lane merges to left). In this way late merge scenario can be identified and also the behavior of a particular vehicle can be tracked in a workzone. An aggressive merger (late merger trying to get in) is more inclined to commit traffic offenses and can easily be tracked by the other cameras if the mobile agent with sufficient information about VO corresponding to voilater migrates to one PPS to another. Once the vehicle is tagged as important, it is further tracked by the camera in order to accurately measure its attributes. These attributes are then supplied to the next camera for further tracking. The attributes that identify a particular vehicle are color, size (length), speed, shape factor (rectangular or round). These are required when the camera is unable to record the license plate number of the violator's vehicle.

The figure 4 (a) and (b) show two situations from the two cameras where a vehicle of interest is a truck and a car. In the field of view of the camera 3, the user detects an object and initializes a mobile agent for tracking on PPS1/Camera1 using the attributes of the object that are extracted from the VO module. The mobile agent continues to track the object and as the object leaves its field of view, its attributes carried by the mobile agent is passed on to the neighboring camera 4/PPS2 unit. On PPS2, the mobile agent tries to match these attributes with the attributes of the different VOs detected by OR module and when it identifies the object, the tracking process continues. In this way the tracker agent migrates from one unit to another to track the object and a complete analysis can be made to identify any possible violation or abnormal behavior, for which an alarm can be raised.

V. CONCLUSIONS AND FUTURE WORK

A generalized DVS framework was presented highlighting the requirement of reconfiguration, migration, quality of service, and dynamic adaptation of analysis tasks to support large scale ubiquitous deployment of system. Use of mobile agents was illustrated with the example of tracking a moving vehicle migrating into different camera regions for traffic surveillance purpose.

The DVS framework could potentially be deployed in applications such as smart environments, intelligent infrastructures, and pervasive computing. Augmenting the system with embedded smart cameras could transform the system into fully scalable and dynamically configurable high performance system.

REFERENCES

[1] *Intelligent distributed surveillance systems: a review.* Valera, M. and Velastin, S.A. 2005, In Image and Signal Processing, IEE Proceedings, Vol. 152, pp. 192-204.

[2] *A co-operative multicamera system for video-surveillance of parking lots.* Micheloni, C., Foresti, G. L. and Snidaro, L. 2003, IEE Symposium Intelligent Distributed Surveillance Systems, pp. 5/1--5/5.

[3] *Content-based network resource allocation for real time remote laboratory applications.* Mittal, A., Pande, A. and Kumar, P. 2, s.l. : Springer, 2010, Signal, Image and Video Processing, Vol. 4, pp. 263-272.

[4] *DVTS-based remote laboratory across the Pacific over the gigabit network.* Kikuchi, T., et al. 1, 2004, IEEE Transactions on Education, Vol. 47.

[5] Collins, R. T., et al. *{VSAM}: A system for video surveillance and monitoring.* Carnegie Mellon Univ., Pittsburgh. 2000. PA, Technical Report.

[6] *Smartcam: Devices for embedded intelligent cameras.* Caarls, W., Jonker, PP and Corporaal, H. 2002. Citeseer. pp. 14-17.

[7] *A survey of video processing techniques for traffic applications.* Kastrinaki, V., Zervakis, M. and Kalaitzakis, K. 4, s.l. : Elsevier, 2003, Image and Vision Computing, Vol. 21, pp. 359-381.

[8] *An Algorithm to Determine Neighbor Nodes for Automatic Human Tracking System.* Kakiuchi, H., et al. 2009. pp. 96-102.

[9] *Distributed embedded smart cameras for surveillance applications.* Bramberger, M., et al. 2, 2006, Computer, Vol. 39, pp. 68-7

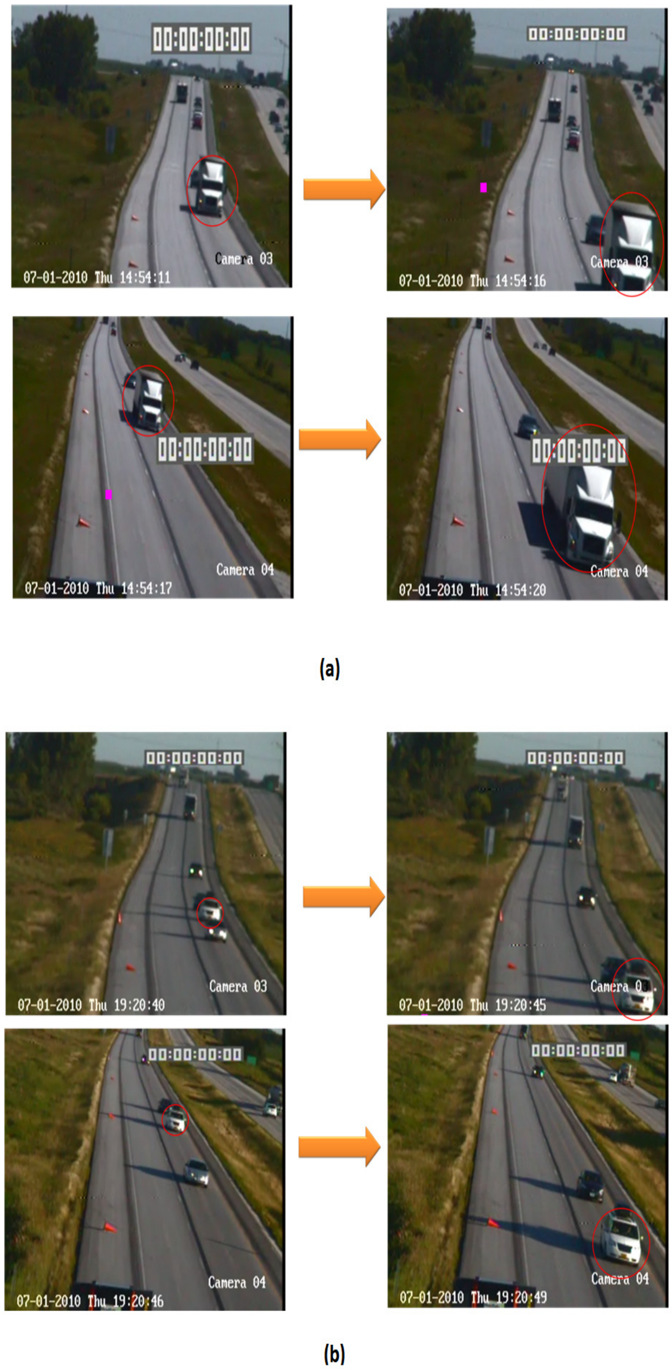

Figure 4 (a) Tracking a truck (b) Tracking a car, between succeeding cameras (here camera 3 and 4) on a highway