分 类 号 _____ 密　　级 _____

U D C _____ 单位代码 10151

基于 3G 的手机视频监控系统研究

包格

指 导 教 师	王新年	职 称	副教授
学位授予单位	大连海事大学		
申请学位级别	工学硕士	学科（专业）	信息与通信工程
论文完成日期	2012 年 5 月	答辩日期	2012 年 6 月

答辩委员会主席 _____

Research on Video Surveillance System with 3G Mobile Phone

A thesis Submitted to

Dalian Maritime University

In partial fulfillment of the requirements for the degree of

Master of Engineering

by

Ngoma Yeke Borgea Dumas

(Information and Communication Engineering)

Thesis Supervisor: Associate Professor Wang Xinnian

June 2012

ABSTRACT

During past six decades, video surveillance systems, using closed circuit television (CCTV), has undergone various technological advances. Indeed, the video surveillance provides a solid physical and logical fundamental security in many fields of daily life. Nowadays, video surveillance is increasingly based on new technologies of information and communication. It is possible to implement a video surveillance system based on 3G networks, because the technology of 3G has been developed.

The system mainly includes three functions, which are monitoring, video acquisition and transmission, and intrusion detection. The last one is the most important. In this thesis, a 3G mobile phone based video surveillance system was designed and implemented. The main contributions are as follows.

1) An Overview of surveillance systems is given

Based on analysis of existing surveillance systems, evolution of the systems is given, and main elements of an intelligent system are listed out. An intelligent system needs supports from technologies of communication, computer, image processing, and even robot vision.

2) A video surveillance system based on 3G mobile phone is designed

The designed system has the features of mobility, intelligence and real time. It contains six parts, which are management module, module of input and output, coding and compression module, communication module, database, and intrusion detection module.

3) An intrusion detection algorithm is proposed

After analyzing algorithms based on differences between current image and background, an intrusion detection algorithm based on adaptive Gaussian mixture model is proposed, and the background can be adaptively changed. According to the experiment results, the algorithm works well.

4) A video surveillance system based on 3G mobile phone is implemented

Based on the design of a system and algorithm, a 3G mobile phone based video surveillance system is implemented under the developing environment of

Microsoft .net 2003, and SQL server is used to manage relevant database. The tests have shown that each module of the system can converge to the expected requirements and has the performance of mobility, intelligence and real time.

Key words: Video surveillance; 3G mobile phone; Intrusion detection; Real time

Contents

Chapter 1 Introduction

1.1Problem Description

The video surveillance began in England in 1970s; they were designed to give a declaration when some suspicious people existed in the places under observation[1]. At first, the video surveillance systems were simply made up of cameras and monitors, and of course, monitored by human called guardian. The installation of cameras had a goal to detect dangerous situations, for example, terrorists' sudden attacks, and once the situation was detected, the guardians would call police and security. Usually there would be many cameras and that meant there would be many monitors at the same time. It was a tough job for the watchers, which made the early systems low efficiency and high cost ones[2].

Later on, the technology of Computer Science and Image Processing made the systems more and more intelligential. There might be more cameras and monitors, however fewer guardians were needed. A computer with special software would be enough to play the part of human's role. They have made a great success supported by the public, because the systems were meant to maintain the safety of public's possessions[3][4].

Now, people gradually care about the safety of both public and private[5]. Individual possessions are also important. Therefore, a system applied to factories and families is needed. First, the designers of the system should figure out that, well, there are many big differences between safety in our houses and offices. For example, in the factory, the bosses can build a team to ensure the systems' running well. But many of us are busy working everyday and have no extra money to build up our own safety office. A video surveillance system that can report the situation in a style of real time, mobile and automation will complete the mission.

1.2Taken Approach

A decade before, scientists have already made the system smarter. All we should do is to make the system with a higher IQ and give a wing of mobility to the system.

In order to update its brain, we should take approaches from the point of humans sitting before monitors. When they are on guard to detect human intrusion, what are they thinking about?

1) Are there any differences between normal background and images in the monitors?

2) If there are any differences, are they brought by objects moving?

3) If the objects are moving, are they humans?

4) If they are humans, a declaration should be given out, the time of intrusion written down, and images stored.

We can design algorithms to detect intrusion, and implement an application to act as the observers based on above theories. One autonomous application[6][7] should be like the followings. When it is installed, at first, a user should configure operating paramcters. Whcn thc system is working well, it responds automatically different messages according to different situations. Autonomous system should have autonomous algorithms for intrusion detection and declaration. Meanwhile the system should be basically stable.

But how can we make the system mobile? Thanks to today's technology of mobility. 3G is a new technology that combines the wireless communication technology with the Internet or other multimedia technology[8]. 3G technology service offers more convenience to current applications, including the needs for higher speed of data transmission. 3G promotes the transmission bandwidth which makes wireless communication more easily [9] [10].Due to the development of networks, new video applications, superior to the traditional video, are widely used in our life. And those technologies can be used in video surveillance to make it as a more convenient system. Compared to the design of a classical system, the specific system that we proposed is concentrated on its mobility and automation[11][12].

The mobility of the system is to ensure the fact that customers can visit the system from anywhere at any time[13][14][15]. So one of our tasks is to exploit the 3G network, for transferring data between users, and design the architecture of system.

Above all, a 3G mobile video surveillance system at least maintains three parts[16]: one part is to capture images as its inputs, another to detect the intrusions as its judging part, and last one to send interesting information to users via 3G and Internet as system outputs, and the system should be stable, accurate and real time.

1.3Relevance

Even though we live a happy life, there are still many criminals around the world [17]. Laws protect us in everyway, but it's impossible to sentence or even catch the bad guys without effective evidences. Cameras are very useful tools to urge strategies. Our society needs video monitoring, and the needs breed the industry of cameras. Now, cameras can be produced maturely, perfectly and with high resolution[18].

As mentioned in 1.1, Computer Science develops so fast that anyone can implement applications with a laptop as long as they keep working on them. Besides, the technology of Computer Vision even can rebuild 3D object from 2D photos[19], and that make the intrusion detection more easily to be implemented.

Today's 3G and high speed of CPUs' calculation guarantee the system running in real time. All in all, no matter in theory, technical supports of data transferring or hardware items, today, the mobile video surveillance system based on 3G can be achieved.

1.4Methodology

The main goal of this paper is to design a mobile video surveillance system based on 3G by gathering useful information and analyzing the functions of the system, and then implement the system.

The purpose of this paper is based on three questions:

1) How to acquire videos and transmit the video flow to computer?
2) How to make the computer detect intrusion like a human?
3) How to transmit alarm signals to the customers by 3G network and Internet?

To answer these questions, our work has been organized as follows.

1) Firstly, in order to absorb some experiences from the existed video surveillance systems, we should overview the principles and architectures of them, especially how they detect the moving objects.

2) Then, according to the knowledge we've got, we propose a design of the mobile video surveillance system based on 3G. The design includes a detail description of functions, architecture of the system, and how the data flows among modules.

3) Thirdly, based on the design of this system, we give the implementation and some results from experiments.

4) Finally, we lay out our works shortly, draw a conclusion on the research, and identify the work in future.

1.5 Outline

In Chapter 1 Introduction, the thesis mainly gives out how the system develops in history, and according to the existed technologies and systems, the thesis aims to implement the system by analyzing, studying relevant knowledge and designing the system.

In Chapter 2 Relegated Work, this paper firstly gives out an overview on different video surveillance systems from the point of their developing history and typical components of them. Then the thesis analyzes several kinds of methods in motion detection.

In Chapter 3 the Design of Mobile video surveillance system, the thesis firstly provides functions of the system, and then designs the system's architecture and data flow. After that, the system's work environments, software and hardware, are stated.

In Chapter 4 Implementation, the thesis firstly gives out how the system is implemented and then lists out the results generated by core modules.

In Chapter 5 Summary and work in future, the thesis lays out the finished works shortly, draws a conclusion on the research, and discusses the work in future.

1.6Summary

In this chapter, we aim to find out how the system has developed in history, how to implement the system based on the existed technologies, and make the readers to get an overview of this paper.

Chapter 2 Related Works

Nowadays, several research projects are oriented in the field of video surveillance. That is the reason why it is more and more widely used. The interesting aspects of reported documents, in surveillance system, include 3D object modeling, human behavior analysis, real-time video surveillance system, motion detection and human tracking [20]. This chapter mainly gives out the existing style of video surveillance systems and methods in motion detection, which is important to make the system smart.

2.1 Overview of Surveillance Systems

Video surveillance is to remotely monitor public and private locations, using cameras, often motorized, which transmit the captured images to control equipment which records or reproduces them on a screen (Figure 1). As section 1.1 says, many systems are on guard to prevent suspicious people from the places under observation, in order to decrease the possibility of dangers.

Due to usage of cameras, we can get as many images as we want in the system. Cameras have done a lot for the safety of our society. There is no doubt that as the supplier of videos, camera becomes one of the most important industry products related to security. So cameras are basic parts of the system, acting as input items. Besides, generally, the video surveillance systems have another three parts[21]: One part to control the video, another to show out the images transmitted from cameras, and the last to give out declarations. Of course, there are many differences among systems in their functions, architectures, technologies and goals.

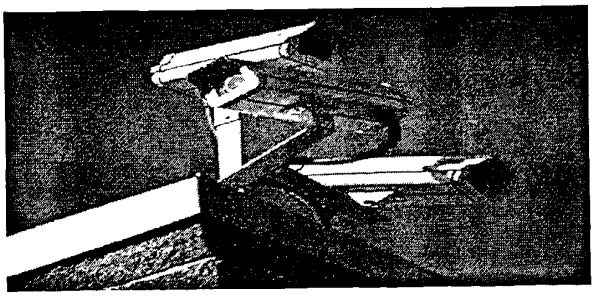

Figure2.1 Surveillance cameras

According to different applications, there are three kinds of systems.

1) Active systems: mostly adapted in public places in order to give emergency calls to police or hospital when there're accidents happening. Usually, one guardian (may be humans or computers) monitors less than 5 sections to guarantee the accuracy, for safety is the highest priority.

2) Passive systems: mostly adapted in factories in order to find out the hidden dangers so that the process of manufacturing won't be interrupted. Usually, there will be a department to take care of it, and one guardian monitors many sections ,for the possibility of danger is low.

3) Record systems: mostly adapted in places where there are many accidents or illegal behaviors. They are built up to record evidences if necessary. When a criminal appears and has something actions breaking the laws, the police will catch him according to the record. Usually, the monitors are located in local police stations or security rooms. In some places, one guardian (may be humans or computers) monitors less than 5 sections to guarantee that the declarations and arrests can be in time.

2.1.1History of Video Surveillance System

According to the types of signal flow developing in video surveillance systems, they can be separated into three generations.

2.1.1.1First Generation: Analog System

The systems in first generation were strongly limited by the undeveloped and developing technology. In the network of CCTV, also called the traditional analog networks, analog cameras are connected with screens by coaxial cables (one cable per camera) for monitoring and storing data, for example, the VCRs that can record videos into tapes. The system has limitations because, only by the connecting lines ,can the monitor restore the image emitted by the source, and multiple cameras meant multiple monitors and lines. However, there were multiplexers used to put multiple cameras together so that video streams could come into a single processor, and then the data could be transmitted to a VCR or an analog monitor. It will display 4, 9 or 16 video signals on a single screen, or record them on the same catalogue system. At the beginning, there was no compression during the video signals' transferring. Under the limitation of immature video compression technology, when a VCR was recorded,, in

order to save space on the video tapes, they had no choice but to cut down the frame rate as long as the videos could satisfied monitors' needs.

The first generation systems were hard to build up, not only because there were many lines and hardwares, but also it was hard to update the systems. VCRs really cost many tapes, but they are still used today because the data losses in tapes are the minimal.

2.1.1.2Second Generation: the Hybrid System

As digital signal processing developed, many new chips, used to do the work of DSP ,started invading the analog system's markets. The digital video recorder (DVR) gradually replaced VCR, which represented a change from analog to digital in recorders. This came to a conversion in system. Appeared in 1990s, digital video recorders were mainly used in storing videos on hard disks. Meanwhile, computers showed their great ability on calculation. With those two hardware's' existence, the hybrid systems usually had multiple analog cameras to capture analog videos as their inputs, and converted data into digital stored in the cache area in computers. By software, DVRs could be shown on screens. For the digital systems could be easily operated and achieved with lower price, the hybrid systems replaced both the analog multiplexer and VCR in the first generation.

The second generation systems were aimed to make the system easily to store and change data into the form that could be managed quickly. As the analog cameras cost less than digital ones, currently there are still many such systems in the world. These systems are equipped with an Ethernet port, allowing to be connected with the network and the digital videos can be played or downloaded by remote access. That means the visitors who have the authority can watch videos at any time in any place, which makes it a real time system. Of course, the video transmissions are under some network protocols, for example TCP/IP.

Besides above, the systems represented an effective solution for upgrading a video surveillance system while maintaining existing analog cameras in place.

2.1.1.3Third Generation: Digital System

In recent years, the technology of DSP has developed a lot. Computers and many chips with good performance are widely used almost in every field. Compared to the original technologies, video compression and transition are more easily to be achieved, and of course, with higher speed. For our system, a video surveillance in the third generation is fully digitized, and that's to say all its components are digital and all transmissions are performed under the networks' protocol. Nowadays, webcams are the most widely used cameras. They are connected via network switchers with servers (laptop), equipped with a video management software. The video processing is done in the server. Recording is done on a server or in network used for video recording. This allows the system, firstly to be able to store large amounts of images without loss of quality, while being able to be selected quickly with processing software, and secondly, a computerized video surveillance system can take good advantage of Internet instead of intranet.

Because of the evolution of mobile phones, nowadays the system can exist as a mobile video surveillance system that provides customers the possibility to get the videos via mobile phone, based on the technology of 3G as long as the phone can be connected with Internet.

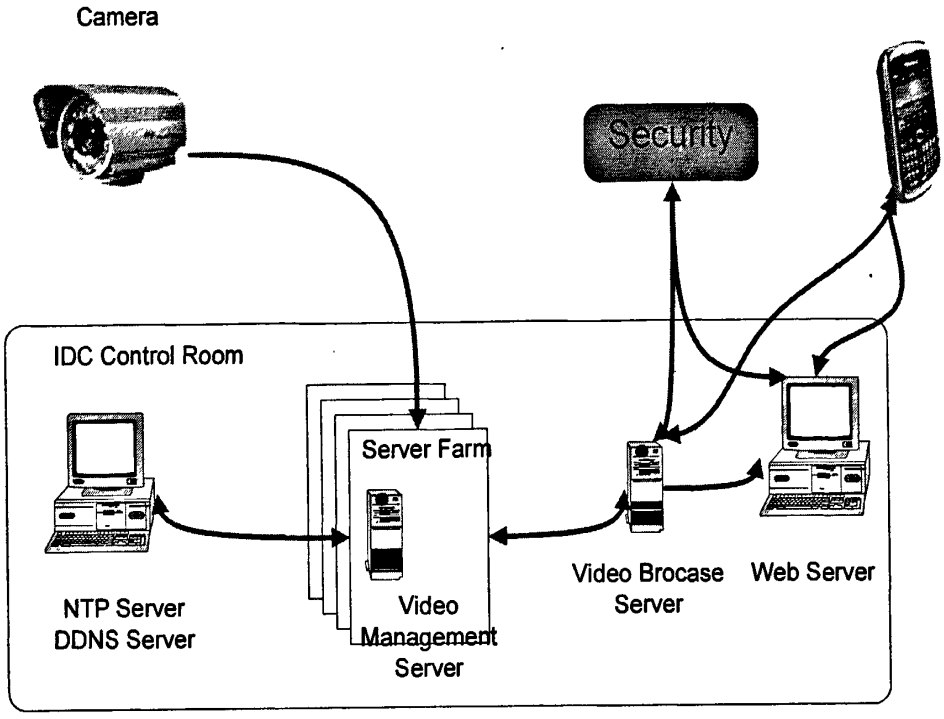

Figure 2.2 A digital video surveillance system

Because the digital data shows its advantages (for example, easy to be stored and managed, etc), the third generation system is a human-free system so that the systems can be smart and automatic. So in some way, we can call these systems smart systems. Through software, the system can automatically identify some objects with specific behaviors in video sequences. When it detects something wrong, it can send warning messages to the supervisor who will then make a response.

The smart video surveillance systems use mathematical algorithms to detect moving objects in images and filter out irrelevant movements. They can create databases that record the attributes of all detected objects and their movements. The decision made by the system or events of interest in archived footage is based on settled rules (e.g.: if a person crosses a limit, send a warning) beforehand.

All in all, the systems in the third generation have the highest speed and they can be made into smart ones. In fact, there are many smart ones in the second generation, but as the technology develops, digital cameras have taken the place of analog ones in the second generation.

2.1.1.4The Evolution of Video Surveillance System

In every generation of the video surveillance systems, some equipment became the most obvious features. For example, DVR meant the disappearance of first generation with VCRs. When the softwares that can be used to get videos under the protocol of TCP/IP emerged, the third generation began. Computers started to take place of the guardians which marked the third generation had developed to a high level. Until now, softwares to complete certain jobs are implemented one after one. The systems are more and more intelligent. The evolution of main equipments in video surveillance systems were shown in figure 2.3.

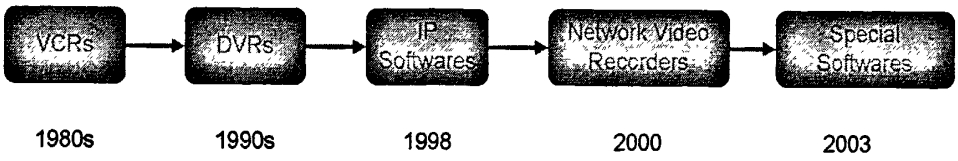

Figure 2.3 Evolution of video surveillance equipments

2.1.2Components of Video Surveillance Systems

In this section, we are going to briefly present the components of video surveillance systems in third generation, and of course, it is a smart one.

2.1.2.1Acquisition

There are a variety of camera models for different needs of monitoring. Either analog or digital, they can be adopted by needs. More specifically, we classified cameras into following types according to their features.

1) Fixed cameras[22]: Fixed in one direction, they cover a defined area (one static office, one static road or park, etc). They belong to the traditional surveillance systems. Sometime they are excellent choices when the designers want most of the images from cameras to be predictable.

2) PTZ (Pan-Tilt-Zoom) cameras[23]: they can be both operated by human and machines (computers or some special controllers) in moving their pans tilted or zoomed in. They are mostly used to do the jobs of objects' tracking, for example, tracking the individuals' moving in the scene, or zooming in regions of interest (for example, a registration plate), etc.

3) Dome cameras [24]: covered with a hemispherical chamber, they are usually unobtrusive. The dome can be in different models just like lizards to hide themselves in different environments, and the invisibility makes them do the work of criminal resistance more easily. And of course, the domes are made them waterproof, and the domes can be fixed and mobile. They can cover a very wide area, because they can be rotated horizontally 360° and vertically 180°. They can be called "watchtower", and one of them can replace 10 fixed cameras to sweep the area needed monitoring.

4) Mega pixel cameras [25]: with a higher resolution than the standard ones, they all have a resolution ranging from 1 to 16 mega pixels. It can not only capture a more detailed image, but also cover a wider field of vision. In this way, their installation can reduce the amount of cameras to cover the same fields. They are used with a wide vision field angle, ranging from 140° to 360°. The softwares to control the cameras offer the ability to control the cameras to zoom in and out. They also can be substituted by PTZ cameras' taking image sequences and use special softwares to do the image mosaic job. Their high resolutions help the system in improving the performance of detection algorithms and recognition, because the data, with a high level, would be managed in detail, for example in the character recognition systems or face recognition systems.

5) Infrared and thermal cameras [26] [27]: sensitive to infrared radiation (IR), they can produce images with high quality according to the temperature of objects instead of lightning. Due to the feature of this kind of cameras, they are widely used in monitoring at night or military. But they can't generate images with real colors, which is their common imperfection. Some infrared cameras are equipped with their own source of IR light, and start to work when the light level drops below a certain threshold. Separated IR projectors (light or LED) can also be used.

6) Pan cameras [28]: with the help of a special lens, they can offer a vision field angle of 360 °. The main panoramic technologies for monitoring are technologies called fisheye. However, the resolution of these cameras is often insufficient for requirements analysis of a high level of details. So one application to generate panoramic photos is usually tied up with cameras.

7) Webcam [29]: as the most widely used cameras in daily life, a webcam is a small

camera connected to a computer. It is made up of a base and lenses that uses CCD (Charged Couple Device), electrical sensor which converts an optical image into digital information, and then transmit data to computer via USB.

2.1.2.2Transmission

The videos captured by surveillance cameras shall be provided to our systems for recording, processing and viewing. This transmission can be done by cable lines (coaxial cable or optical fiber, copper twisted pair) or wireless styles (infrared signals, radio transmission, etc).

Mostly, video are transmitted by wires. In this way, the system provides greater bandwidth and higher reliability than wireless connections with a lower cost. However, the wireless video transmission is sometimes necessary, for example one system needs monitoring large areas where wiring installation would be too expensive, or when the areas to be monitored are impossible to reach by wires. The video signal can be analog or digital. Even today, most of the video transmissions for surveillance are analog. However, computer networks (LAN, WAN or Internet) are increasingly used to carry the video under the protocol of TCP/IP protocol.

2.1.2.3Coding and Compression

The digitized videos contain a large amount of data which have to be transmitted and archived. Storing a uncompressed digital movie file which lasts an hour can take up more than 165 megabits in the hard disk. That is why the surveillance video must be coded, or compressed. By coding and decoding, algorithms can reduce the amount of data in eliminating redundancies. These methods mainly do their jobs by finding out the relationship among images in the same sequence, and the fact that details in images are invisible to the human's eyes. In order to make the system with real time , these methods are required to run the codec easily, also quickly.

The goal of coding and compression is to make the video easily to be recorded, stored and played. Every application to record, store and play the video, in fact, is a video coding/decoding tool. In order to uniform the world's standards for coding and decoding, there are two main groups of international standards to draw up image/video compression. JPEG[30] is created by Joint Photographic Experts Group, and MPEG[31]

is developed by Moving Photographic Experts Group. In the first group, we have JPEG format for fixed images and MJPEG for video clips. The second group includes MPEG-1, MPEG-2, MPEG-4 and H.264 formats.

In fact, nowadays cameras usually carry one or two tools with them, some in the form of software and some in hardware.

1) Video recorder (VR) does coding and decoding jobs in order to convert signal into data that the system can be managed, and then store them in some standards. It is usually a function of cameras that can't be connected in net as an internal hard drive. The outputs of VR usually exist as video files. Different cameras may produce different types of files. According to the types of its inputs, there are three kinds of VRs: digital video recorders (DVR), whose inputs are digital, analog video recorders (VCR), whose inputs are analog, and hybrid video recorders (HVR), whose inputs can be both analog and digital. But in IP cameras, VRs are usually soft wares which obey the protocol of TCP/IP or others. They get the data from IP cameras via the net, and complete the job in computers.

2) Videos' storage (VS) is usually done by both hardwares, that connected cameras with systems, and softwares that code the videos into files. For the system, on average, organizations maintain the videos as evidence for 30 to 90 days. In fact, the pictures taken from the cameras with high resolution still occupy much space in the storage device though they are even compressed.

According to the storage devices' locations, storage solutions in the systems can be classified into two types.

(1)Internal: The Hard disks can be integrated with DVR or servers, which is the most common form of archiving. It can provide up to four terabytes of space. Some IP cameras even have a memory card or a USB drive that can save videos last for hours, even days. This kind of solution is suitable for the video surveillance systems that have less than 50 cameras.

(2)Attached: Archiving is done on external devices for recording or video servers. NAS (Network Attached Storage) or SAN (Storage Area Network). These systems provide a shared storage between different network clients. On a system of NAS, a file

is stored in a single hard disk. Whereas with SAN, a file can be saved in fragments spread across multiple storage media.

These attached storage solutions are particularly advantageous for large video surveillance network with great numbers of cameras. Although archiving systems are more expensive than internal,, these solutions have the merits of scalability, flexibility and redundancy.

3) Videos' display

Much of the video captured by surveillance cameras is never viewed. It is simply archived in case that a viewing is required following an incident. Traditionally, video surveillance was used primarily as an investigative tool. However, in many cases, surveillance systems, which had security guards to view the monitor, were real time to show images from the surveillance cameras.

Video can be viewed on different devices. In small facilities, it is possible to watch the video directly from a recorder, meanwhile it is simultaneously recorded. More generally, the images will be viewed remotely on a computer, in a mobile phone or a portable device as long as the device has a software to decode the videos.

Major centers of security operations, overseeing hundreds of cameras, often use a wall of video screens. They offer a large viewing surface and can display different video streams.

2.1.2.4Processing

After the needed data comes into the system, the system should run the core part of it, the processing part, to do the job of images processing mainly including videos' analyzing and researching useful information. They usually exist as softwares, for they can be implemented to complete different missions. This part can be updated by the designers as long as the hardwares of the system maintain unchanged.

2.1.3Areas of Applications of Video Surveillance System

At the beginning, the video surveillance systems are no more than cameras, and all they are designed just to capture videos. The videos captured by cameras were firstly used for public services (police, transport, etc). The systems later was adopted by companies in order to protect strategic assets such as refineries, power plants, dams,

the food processing plants and pharmaceutical factories. For the purpose of earning money, the casinos became the pioneers to install video surveillance systems. Today, surveillance video systems can be found in both public and private places, including shops, parking lots, airports, banks, highways, undergrounds, train stations, etc.

2.2 Overview of Methods in Motion Detection

Motion detection is the basic processing for any automatic system. The tech aims to detect moving objects by image sequences. At each moment, each pixel must be labeled by an identifier bit (fixed/mobile). When camera is fixed, it can be detected by temporary differences computed among continuous frames. The motion detection algorithms can be divided into three classes: the ones based on optical flow, another ones based on clutter and the rest based on feature points. These three kinds of methods have been widely explored, and each of them has both advantages and disadvantages.

2.2.1 Technical of Optical Flow in Motion Detection

The optical flow[32] allows the implementation of a comprehensive analysis of the movement using an equation to join the variation of intensity light at a point with the speed of this point. This technique allows to analyze scenes that the totality of the picture is moving, and to compare the moving objects to each other. This technique is used in weather service (analysis and measurement of movements of clouds, cyclones), and in helping driving cars (detection and analysis of moving objects in front of the vehicle: other cars, pedestrians, etc). Figure 2.4 can show the results generated by the analysis of optical flow.

This technique is not robust to the noise and requires additional treatments, such as filtering, to operate satisfactorily.

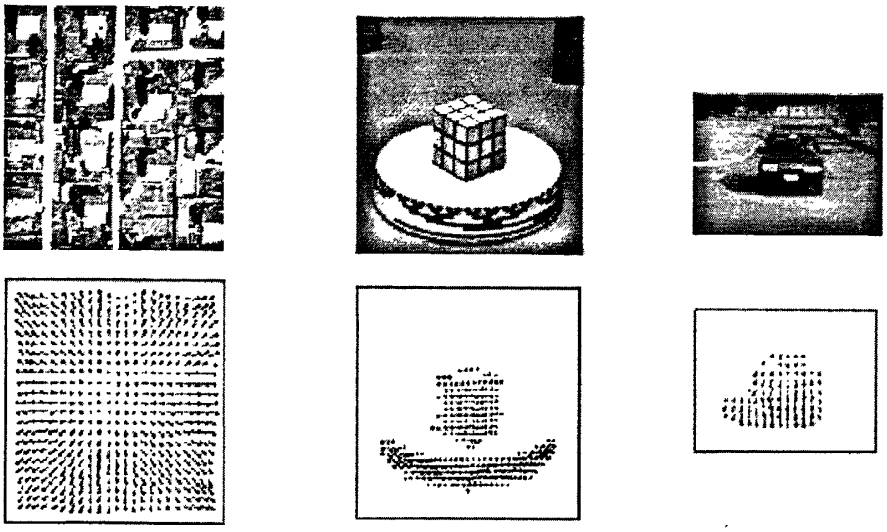

Figure 2.4 Example of optical flow

2.2.2Technical of clutter in Motion Detection

This technique of clutter is widely used by several applications. It detects the movement regions pixel by pixel to subtract the current image to the background image. It is limited to cameras in a fixed position. However, the technique allows, through the establishment and continuous updating of an image of a blank background, to distinguish moving objects with a background different from the blank background.

There are three ways to create the background image:

1) The averaging calculation of images.

2) Choose the background image that has little change and in which, there is no movement during a certain time.

3) Select the first image as the background in the video stream.

Similarly the technique is not robust to the noise and requires additional treatments, such as filtering, to operate satisfactorily.

2.2.3Technical of feature points in Motion Detection

This technical focuses on finding feature points [33] [34] (corner points, SIFT points, etc) in the images and then find the most closed to the found ones. If the points' locations have changed in different images, there is no doubt that motion emerges.

This kind of technique is robust to noise but the time and memory consumption will be amazingly large.

2.2.4Motion detection and its Constraints

The challenge of detecting motion is that when the systems are made up by fixed sensors, it's hard to guarantee the ability to perform a good segmentation of moving objects, regardless of their size, velocities, or the complexity of background. In this research, when designing the algorithms for motion detection in video surveillance systems, the following constraints should be taken into account.

1) The algorithms should be stable to make the system stable. In order to make the system an automatic one, the system must be able to be operated without human's intervention for a long time, and be able to take into account gradual changes, or sudden changes such as illumination or the presence of new static objects in the scene. The system must be temporally adaptive.

2) The algorithms should be highly robust to the noise. In order to make the system an effective one, the system must have the ability to ignore the effects caused by all uninteresting movements, for example, noises generated by a river waving or grasses blowing in the wind. And of course, the system should be robust to small movements of the sensors.

3) The algorithms shouldn't be too complicated. In order to make the system a real time one, there will not be much time to run the part of motion detection.

2.3Summary

In this part of our work, we presented an overview of surveillance systems, introduced three generations of systems, listed out how they evolved. After analyzing the components of different systems, we get the useful information to design our own system. Then we brought information about the motion detection in a sequence of images to a video surveillance system. Lastly, we concluded some constraints in motion detection in order to make our system smarter.

Chapter 3 the Design of Mobile Video Surveillance System

The video surveillance systems involve monitors and remote sites (rooms, warehouses, conference rooms, etc). Nowadays with the progress of technologies in the areas of video surveillance and mobile telephony, we believe that we can build up a mobile video surveillance system, which can be real time and connected with portable terminals by the Internet and mobile telephony networks.

The mobile video surveillance offers to those who need to give a removal guard in case of anything wrong. That's to say when there's a suspicious person detected by the system, the system sends messages alarms in order to ensure the safety of commercial, industrial or residential premises.

The remote surveillance can be controlled by a laptop or desktop computer connected to Internet network (as showed in figure 3.1) by the means of a 3G modem, ADSL, etc.

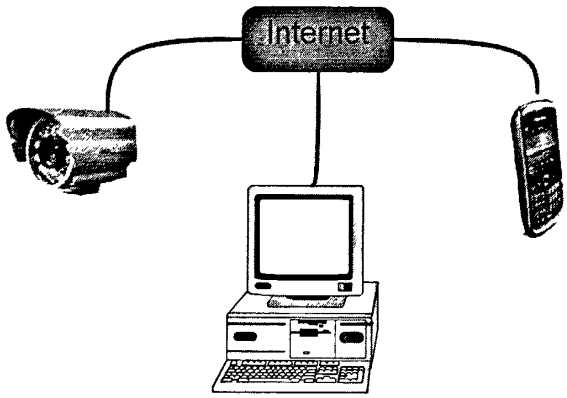

Figure 3.1 Mobile video surveillance systems

3.1Preliminary Design of the System

The system, of course, should have the basic items. Cameras play the role of input to the capture digital/analog videos. The part of data processing will do the intelligent work to give a declaration of intrusion. Via the network, the messages can be sent to users. The figure 3.2 shows the basic architecture of the mobile video surveillance system.

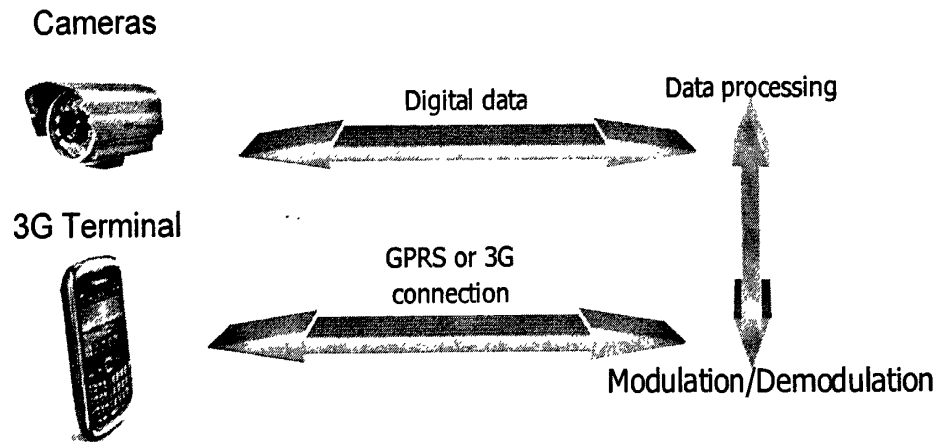

Figure 3.2 Global view of mobile video surveillance with GPRS or 3G connection

3. 2Needs of system

The system should be real time, mobile and automatic.

With the new generation of 3G mobile telephony, the operation of the bandwidth is wider, and allows passing more data than usual (audio, video, text). It can reach theoretical flow of 2Mbp/s, and commercial flow of 384 Kbps in mobile reception and 64 Kbps in emission. The technology guarantees the speed of data transmission. As long as the algorithms in the part of image processing allow, the system will be a real time one.

Indeed, the mobility of the user is to ensure users can visit the system at any place. So our solution is to exploit the 3G connection for transferring data between the user and the core system in order to make the system a mobile one.

The system must also be autonomous. When the system is installed, at first ,user should do some configurations of operating parameters. When the system is working well, it responds automatically different messages according to different situations. Autonomous system means autonomous algorithms for intrusion detection and declaration. Meanwhile the system should be stable basically.

3. 3Specifications functions of system

Our purpose in this work is to design a mobile video surveillance system that the users can log the system at any time to view the current images captured by cameras via 3G networks and get messages of warning generated by the system. So, the system has

three functions, real time monitoring, which is used to monitor the places need monitoring, video acquisition and transmission, which is used to send data, and intrusion detection, which is used to detect the intrusion.

3.3.1Monitoring

In fact, a monitoring system is an application to do the decoding jobs when the coded data arrives. This function usually exists in the phone like software. As software is decoding, this function pays more attention to the standards of videos and the memory of the terminal. 3G mobile phones are good choices, for there are many existed applications that can complete the mission in the platform, and usually they have got a large memory more than 1GB.

3.3.2Video Acquisition and Transmission

After the camera captures the images, the data should be converted into the digital, and one part to do the coding and decoding job should be in the system. It's easy to connect cameras with computers used as the signal processing part by USB. But how can the data be transferred between phones and the system? Of course there should be a communication module to identify commands from 3G terminals and send video stream or warnings to the phone. Figure 3.3 shows the function of the part.

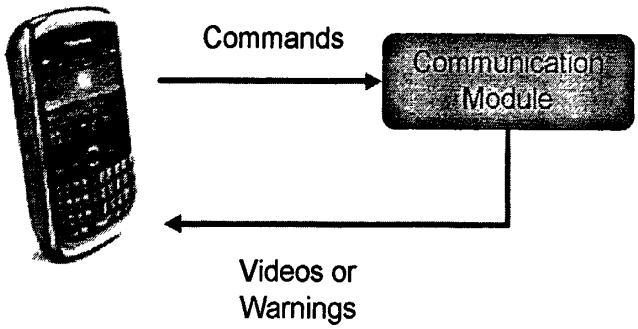

Figure 3.3 Communication Module in the system

3.3.3Intrusion Detection

This function makes the system smart which determines the importance of it. This function mainly does the guardians' job. It is usually exists in the computers like software. For example, we have a jewelry box under monitoring, and when a human presents and the camera captures pictures showing him, the intrusion detection software

should detect the behavior and send massages to 3G Terminal. In order to make sure we can catch the man, the less time it costs, the better. This process can be shown in figure 3.4.

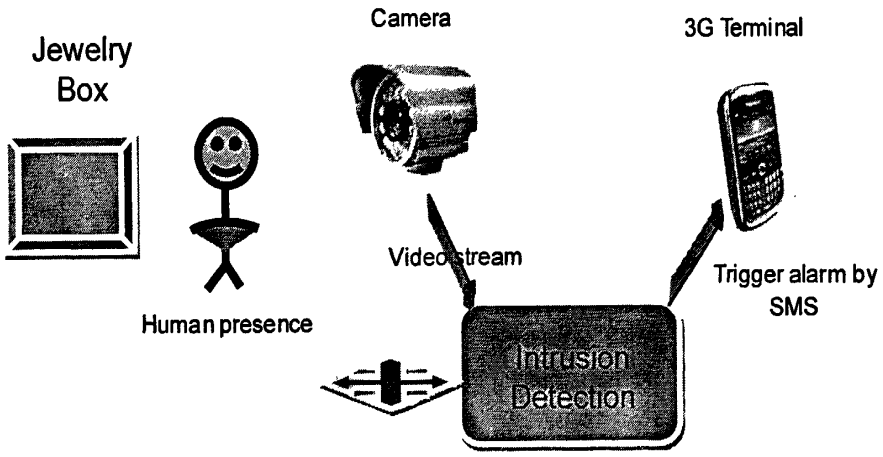

Figure 3.4 Example of intrusion detection

3.4Architecture of the Proposed System

According to the analysis of functions in the system, the proposed system should at least contain parts below, cameras and 3G terminal as module of input and output, the management module to set paramcters in the system, the intrusion detection module, the coding and compression module to convert videos, the communication module, and the database to give record on the time of warnings and videos.

The proposed parts in the system are illustrated in figure 3.5. It brings together the various modules that comprise the system.

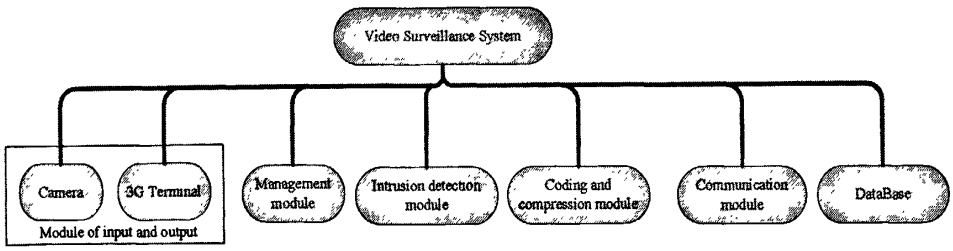

Figure 3.5 Modules of the proposed system

3.4.1View of each Module of the Architecture in System

The system architecture represented in the figure 3.5 is made up of four functional modules, the database, and module of input and output with one camera the 3G terminal.

3.4.2Management Module

In order to make the parameters can be configured easily; there is an interface for customers. This unit represents the persons machine interface (IPM). In general, we recognize the importance of friendly user interface, because it is clear that there is a strong relationship between the design of the (IPM) and the internal design of application. The design of the IPM is seen as an in-depth understanding of the constraints, priorities and clearly identifying system concepts.

But for most projects, very few resources are mobilized to implement actually the ergonomics (human engineering). On the other hand, the involvement of a specialist may be too expensive; they are then often developers, architects or clients themselves to design their own interface. For our system we need a simple interface that will help us set our parameters as a list of numbers to call in case of intrusion or adjusting the size and format of the video sequences to send.

3.4.3Module of Input and Output

The image acquisition module represents the input of our system; this task must be performed via a video capture device (camera, webcam). For this category of products, there is a variety of manufacturers that offer multiple solutions with different technologies.

1) Objective: The objective in this part is to make a good choice of hardware and software for video capture function in order to be faithful to the needs and constraints of the system. We also seek to use the product that allows us to use advanced features (synchronization, compression). Currently, cameras are becoming more advanced digital technologies to offer professional options.

2) Different camera models: First, we must not confuse a network camera and a webcam. A network camera connects directly to a network, it combines the functions of an optical camera and the ability of a small computer equipped with an internal web

server. The camera thus has a network RJ45 socket for direct connection to a HUB or Switch. While the webcam connects to a PC via the USB port and has only the optical features. Several options exist for selecting material for this module.

(1)Smart camera: Provides powerful optical features, it may include an intelligent image processing.

(2)IP camera: This camera can be connected directly to an IP network (network camera).

(3)Web camera: This camera connects to a PC via a USB port.

3) Analysis of the choices of camera: Smart cameras perform better than other type of cameras, they offer more advanced features. Such as option, motion detection, alarm procedures. These types of camera are relatively high priced, and they are very useful for security applications.

For our system, an IP camera is sufficient to provide the core functionality. Many products include an MPEG-4 (SONY SNC-M1, SNC-M1W, SNC-M3, SNC-M3W, SNC-P1, and AXIS 207). The cameras that incorporate a compression have relatively high prices, (The day of the test system, we will need internal use of the camera, and so we can operate a webcam). Indeed, it is characterized by simple installation on different operating systems. Latter simply substitute a webcam by an IP camera suitable for application.

In order to make our system an automatic and mobile one, there is no doubt that 3G mobile phones become our terminals.

3.4.4Coding and Compression Module

There are many standards for video coding or compression. How we choose the treatment? Video compression is an important consideration to ensure good quality of service in new generation networks. Indeed, one of the major problems encountered in the integration of advanced services in third generation network is the ability of mobile devices to decode and read large data. Since there is not a compression technique that can combine all the needs of each application, there are various types of compression that cover different requirements. MPEG-4 is a new standard that was originally designed for low flow rates. Today, its range of use has spread to various applications

such as DVD, digital television, video production, mobile multimedia, gaming and video telephony.

MPEG-4 has some advantages such as

(1)The increasing in compression performance.

(2)The scalability of a media object, i.e. instead of compressing a video sequence in blocks of pixels, a media object can represent an image part of various shapes (e.g. background, car), or an audio object (e.g. a voice or the sound of a musical instrument). Finally, these objects are encoded from a set elementary stream that will be successively fragmented and pocket size to form a stream of access units (AU, Access Unit).

(3)Robustness to errors.

(4)The encoding of multiple multimedia data.

The size of compressed video images is 176*144 pixels. Packet size MPEG4 video is not fixed, the average is 158 bytes. MPEG4 terminal consist of three layers, layer compression, layer synchronization and layer delivery.

The layer compression generates elementary streams (ES: Elementary Stream) organized access units (AU: Access Unit). The AUs are transmitted to the sync layer with time-stamping information.

The synchronization layer synchronizes the elementary stream level access units by encapsulating each elementary stream in a separate packet stream SL (Layer Synchronization). This layer is optional and may not be implemented. The layer consists of DMIF (Delivery Multimedia Integration Framework), it receives the packets through the SL interface DAI (DMIF application interface) and provides multiplexing by offering seamless access to the transport technology used.

We can reduce the complexity of the system by choosing from the outset,a camera module that integrates video compression.

3.4.5Communication Module

This module represents the interface input/output of system; its role is to transmit the video stream to the user and receiving the request if it is a monitoring application. Clearly this module has the function of bridging the gap between the requirements of the user and system applications; the input of the system includes receiving the user

command to enable its translation into the kernel system; the output of the system is to access the GPRS or 3G network, and then send the video using the address of the user terminal.

This feature requires a transfer mechanism in conjunction with the various entities of the system:

1) Related to intrusion detection module to allow conveying data immediately after they are analyzed in case of trouble.

2) Linked to module coding and compression to transmit data without necessarily analyzing when monitoring the application.

3) Linked to the database, which allows users to access archived to footage.

We are going to give some detail about the technical specifications of the communication module to make a good material choice.

(1)Objective: This module will be responsible for the transmission of data between the system and user. It's used for transferring data between a PC and 3G terminal. So we need a communication modem for 3G network access from the desktop. According to our needs, we will compare existing products on the market and select the most suitable product for our system.

(2)Technological solutions: A user access the third generation network via its integrated modem in his 3G terminal. However, the multitude of services and rates offered by the network grow the business and private estates. So now there are several ways to access 3G network through a personal computer.

① PCMCIA cards: The development of this type of card is intended primarily for laptop, they are characterized by a simple and inexpensive installation. This card is about some size of a credit card, simply insert it into the appropriate port and install the drivers.

② Adapter cards USB/PCMCIA: some companies need access to broadband networks via their intranets and enjoy the functionality of the PCMCIA card at the same time. In this case, the USB adapter to enable the function of the use of PCMCIA cards in desktop by connecting the card to the PC via the USB connector.

③ Adapter cards PCI/PCMCIA: This adapter also allows the use of PCMCIA

cards by users of fixed PC.

④ Terminal modem: Certain terminals can play a role of a modem with their USB connectivity to the desktop. Several manufacturers offer this solution.

⑤ 3G USB Modem: Today we see many 3G modem selled by the 3G modem manufacturers, we should know which type of 3G Modem that fit our need. 3G modem is a device that allow computer to connect to internet via 3G mobile network. We can separate 3G modem into three types and this type is currently available in the market, there are: The Data Card models, the USB modem Stick, and the last one is 3G USB dongle.

(3)The choice of modem: For our system, connectivity modem is preferably via USB. Respecting this condition, we can either use 3G USB modem or a USB/PCMCIA card adapted by a USB connector.

(4)Important note: For windows, all the solutions mentioned above are possible. The decision will now depend on our environment and the availability of product to choose. The installation of the modem requires a driver that enables communication with the machine. The material will be detected and can exchange data between input and output.

3.4.6Database

Storage of captured footage is useful for certain functionality. The user can request to see a saved file. For example in America, the surveillance cameras reveal some of the crimes and acts of violence. Sometimes filming accidents is useful for some authorities. The sequences are then stored in a database, compression of the incoming flow reduces the occupation of storage space. It must also release previous recordings to save space occupation.

Where the acquisition module does not perform the task of data compression, it becomes very cumbersome to save the footage in our PC. Faced with this problem we can store the data in an MPEG-4 digital recorder. This device allows to back it up via USB, which includes alarm inputs and an internal hard drive capacity (160-750 GB).

It is necessary to use management software database in order to manage the huge amount of sequences to be recorded. It is also preferred to create reports indexed from

video recordings. This archiving can reduce the space occupancy and facilitate the management of database.

3.4.7Intrusion Detection Module

The name we have chosen for this module means a security's measure to take against a danger (e.g. monitoring a room system or equipment for detection of moving regions in image). However, the feature may simply be an identification of areas (monitoring sites or road segments), an analysis of the trajectories of moving objects in the reference image or an optical zoom, etc.

So we can call it a module of analysis and calculation. This module focuses on three key functions:

1) Detection: By comparing two successive images to locate objects in motion or identify specific key areas.

2) Analysis: Make projections in a geographic reference; analyze a trajectory of an object in the reference image, etc.

3) Identification: Enlarge a image or make optical zoom on parts of the image to identify an object or person.

To detect an intrusion is to extract attributes which characterize the different images in a sequence, such as points of interest, colors and regions, etc. Then space-time information provided by the sequence is used to check for unwanted changes (abnormal movement or introduction of new objects).

Algorithms for motion detection fall into two broad classes: optical flow and correlation. These methods have been widely explored, and each has its advantages and disadvantages.

The first and oldest of all methods is a simple frame difference, where a notable variation of gray level is associated with a movement. Parallel algorithms are developed based on Markov fields in many areas in image processing (edge detection, segmentation of areas, and restoration of noisy images). If their mainly advantages are the robustness and quality of results, their main drawback is their slow execution speed due to the large volume of calculation.

The motion detection method that we adopt depends primarily on the needs of our client who offers a proper procedure with the feature of the site that wish to be monitored. On the other hand, the constraints of our system require us to adopt the method that requires a minimum of computing time because the response of the entire system must be conducted in real time.

So we will proceed with the methodology for comparing the gray level between two images.

3.5Data Flow in the System

By the diagram below, we are trying to show data flow procedure overview.

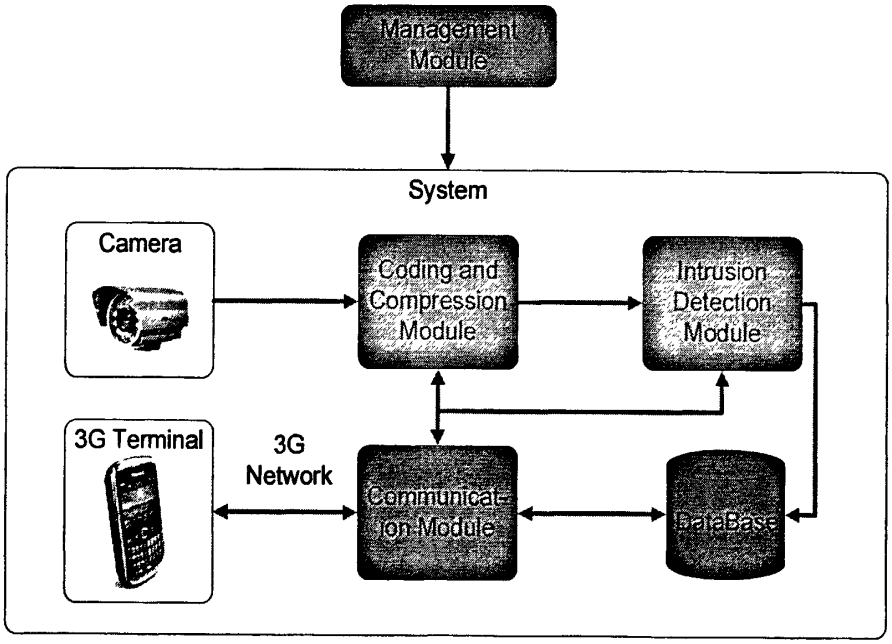

Figure 3.6 Overview about the data flow diagram of system

3.5.1Data Flow in Monitoring Application

To implement this procedure, we must find a way that allows the user to place an order through its terminal. The system's response to this command must be real time. Therefore, our idea is to operate a mobile videophone mechanism.

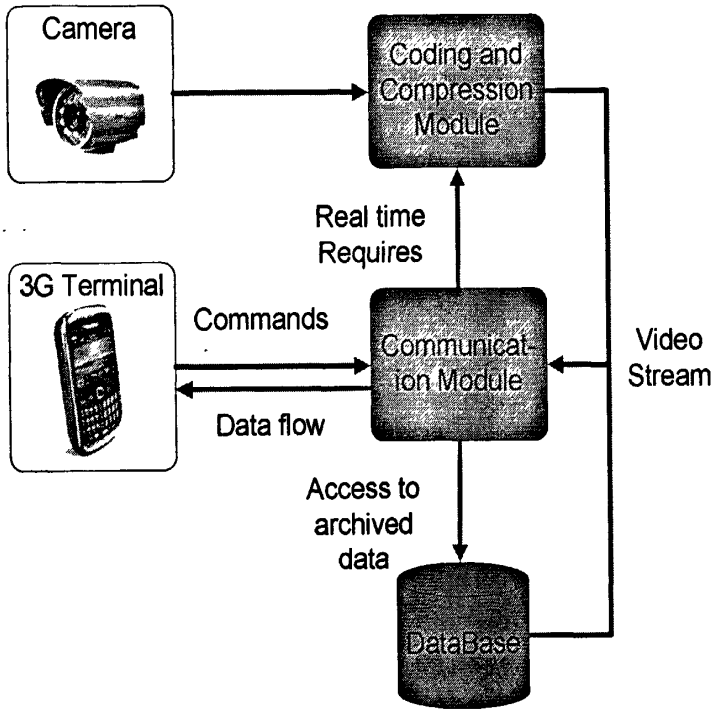

Figure 3.7 Monitoring application data flow diagram

Monitoring application and data flow procedure:

Commands are messages containing requests for access by the users. They can be either real time data, or a sequence archived at a given time or for a definite duration.

Real time access means:

1) Run the query request
2) Negotiation to establish a connection between two remote devices
3) Establishment of a media session between two remote devices
4) Transfer of video stream to the user
5) End of the session and the close of connection

Access to archived data: The procedure for this mode of use is similar to the part of output, except that the user does not need direct access to real time flows, its objectives is to access the old archived footage in the database.

This use presents a difference case compared to the first level of establishment phase of the media session. Indeed, the user must include ,in this phase, the media source to get what he wants and some parameters such as time and date of acquisition of

the sequence that we have desire to allow the system to identify and extract video from the database.

The procedure is as follows:

1) Run the request

2) Negotiation to establish a connection between two remote devices

3) Establishment of the media session and indexing to the database, to identify the parameters of the sequence and extract the required sequence

4) Transfer of video stream to the user

5) End of the session and close the connection.

3.5.2Data flow in Video Acquisition and Transmission Unit

In fact, data in this unit is just like a receiver and transmitter, so the data is flowing clearly in the system. Images or videos come into camera, and then the coding and compression modules compress them. After that, the data can be stored automatically into database, or transmitted to the terminal at the same time by communication module.

3.5.3Data Flow in Declaration of Intrusion

The system must first be a motion detection module. This module can process the received images to uncover and report intrusions that appear on these images. In intrusion detection, a procedure for triggering an alarm shall be executed.

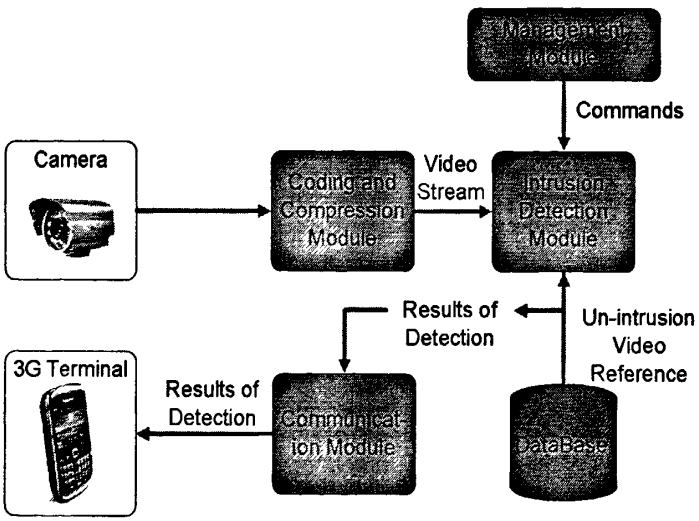

Figure 3.8 Declaration of intrusion data flow diagram

Declaration of intrusion procedure is shown in the following.

1) Images from the camera are analyzed using the adopted method. The goal here is to detect the presence of strange objects or people, movements or unusual phenomena in scroll images.

2) In case of intrusion, the system asks for a connection establishment with the user by the module of management module.

3) Establishment of a media session.

4) Transfer video stream in real time. The video stream which is transmitted can be accompanied by an alarm (voice, tracing trajectories, and delimitation of the intruding object ...). Note that graphics solutions require the use of powerful tools for image processing.

5) End of the session and close connection.

3. 6Work Environment

To carry out our work and ensure the functioning and operation of our system, we will create a development environment that is in fact described by the description of provisions (hardware and software).

3.6.1Hardware Environment

In this system, these items are needed: one laptop with good hard disk capacity (at least 320GB), used as the management module, coding and compression module, part of communication module, intrusion detection module and database; one webcam (with a resolution of at least 3 million pixels), used as the input module; a 3G mobile phone (with a good CPU), used as 3G terminal; one 3G USB communication modem, used as part of communication module; finally, wires , used to connect camera, computer and Internet.

3.6.2Software Environment

The choice of software environment is general depends on technology directions. Basically, the system should at least satisfy below items.

1) The operating system should be Windows XP sp1/ Win7/Vista.

2) The system should have libraries of Microsoft visual .Net 2005.

3) The net agreement should be TCP/IP.

4) There's sql server installed in the computer.

In fact, the digital video surveillance equipment must:

1) Ensure the supervision and streaming in real time.

2) Have adequate means of image processing: compression, specific processing.

3) Manage IT resources: storage of data.

For our case, we plan to work with two sever; the web server and the video server.

The web server provides the interface of user on two sides (camera side and surveillance side). Both sides use web browser to establish sites under surveillance. The video server provides the function of real time audio/video stream transmission and management.

3. 7Summary

In this part of work, we firstly analyzed the needs of the system, and then classified the functions in the system. According to the functions, we specified the various components of our system, and then figured out how the data flows between every module. To make a mobile video surveillance system obeying to the specifications laid out in this part and providing the requested features, we made some rules for the environment of the system. After the design, we will make it a fact next.

Chapter 4 Implementation of Mobile Video Surveillance System

Base on design of the system, implementation of our mobile video surveillance system is really clear. But for intrusion detection module, an algorithm to detect intrusion is needed. As long as intrusion detection module works well, system can be a intelligent one. So in this part, we will firstly organize the development process according to chapter 3, and secondly 3G network will be embedded into the system in order to meet the expectations of our system, and thirdly an algorithm for intrusion detection is proposed, and the last we will present the results already obtained with the application experiments, used to test performance of the intrusion detection module. Considering that we are faced with a programming problems and logistics to carry out this realization, system uses many existing softwares and hardwares. For example, receivers and transmitters in communication parts and applications used to display videos in 3G mobile phones.

4.1Module of input and output

Concerning the image acquisition, tests were performed with a Logitech webcam Pro HD C525. We show the characteristics of this webcam below.

Figure 4.1 Webcam Logitech Pro HD C525

Technical Specifications :

1) HD video calling (1280 x 720 pixels)
2) HD video capture: Up to 1280 x 720 pixels*
3) Logitech Fluid Crystal™ Technology
4) Logitech More HD technology
5) Auto focus
6) Photos: Up to 8 mega pixels (software enhanced)

7) Built-in mic with Logitech Right Sound™ technology

8) Hi-Speed USB 2.0 certified (recommended)

9) Universal clip fits laptops, LCD or CRT monitors

Basic requirements:

1) 1 GHz of crystal oscillators

2) 512 MB RAM or more

3) 200 MB hard drive space

4) Internet connection

5) USB 1.1 port (2.0 recommended)

General Requirements:

Windows® XP (SP2 or higher), Windows Vista® or Windows® 7 (32-bit or 64-bit)

Iphone-4s will be enough as the 3G mobile phone.

4.2Coding and compression module

In the system, the camera we used supplied us a short way to code the video. In the camera, there is a module for compression. The video coming out from the camera satisfies the standard of MPEG. That's to say, what we get is what we want.

4.3Communication module

About the communication module, the USB modem E153 of HUAWEI plays this role to ensure communication between System/User and User/System. Below are the specifications of the modem.

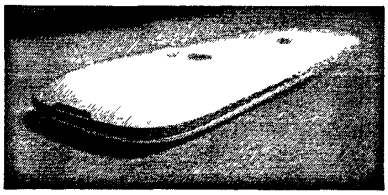

Figure 4.2 HUAWEI USB Modem E153, CE0682

Detail Product Description:

Huawei E153 modem 3G USB sticks

Speed: 3.6Mbps Wireless

UMTS 2100MHz HSDPA

Micro SD memory card slot

Windows 2000/XP/Vista/7, Mac

Key Features:

1) HSDPA/UMTS
2) EDGE GPRS/GSM 850/900/1800/1900MHZ
3) Equalizer
4) Micro SD card Slot
5) Data and SMS service
6) Plug and play

Systems requirements:

1) Windows 2000 SP4, Windows XP SP2/SP3;
2) Windows Vista SP1/SP2, Windows 7;
3) Mac OS X 10.4, 10.5 and 10.6 with latest upgrades;
4) Hardware system should meet the recommended requirements for the installed OS version;
5) Display resolution 800*600 or above;
6) Standard USB interface.

4. 4Database

In our system, there are two databases.

One is used for record the evidence of intrusions, including when the intrusion is detected, when the warning messages are sent and the pictures that show the intrusion. There is one table in the database, and the key value is generated by the detected time. For example, the time of intrusion detected is 2012-06-12 14:23:33, and the key value will be 20120622142333. And the table should contain the time column, the warning messages column to record the warning messages sent, and the files' path column to record the images showing the intrusion.

Another one is used for managing the recorded videos from cameras. There is one table in the database too. Besides the key value column, in the table, there are one

column to record the date, one column to mark whether the intrusion happens in the video, and one column to record the videos' path in the server.

4.5 Intrusion detection module

This module should be implemented by software, and this software can be embedded into the system.

In order to analyze and compare the image, we will focus on the levels of Red, Green and Blue (RGB) to each pixel of image.

4.5.1 Algorithm of Intrusion Detection

The invasion of the object detection algorithm in video images is divided into five steps to detect objects. The first step is background initialization and foreground detection. An online adaptive Gaussian mixture model is used to establish background template [35] [36]. The foreground pixels can be detected by using the background template and the current video image. The test process relies on the use of the background template, and can update the background template, in order to adapt to the dynamic scene change. Because of the camera noise and environmental impact, the foreground image must contain noises. So the second step is to eliminate noises through the image post-processing operations, and the filtered image can be got. The third step is to find the target area by labeling the connection areas, and then the rectangle boundary is calculated. The fourth step is to remove the remaining small areas through the relevant rules. The final step is to extract the target features such as center of mass and so on. The detailed descriptions are as follows:

1) Background initialization and foreground detection

A foreground pixel map is created and object features at every video frame are extracted by using a combination of a background model and image post-processing methods. Background models generally have two stages: initialization and update. An online adaptive Gaussian mixture model is used to detect foreground in this step. The basic algorithm is as follows:

Each pixel is characterized by its intensity in the RGB color space. Then, the probability of observing the current pixel value is considered given by the following formula (4.1) in the multidimensional case:

$$P(X_t) = \sum_{i=1}^{K} \omega_{i,t} \eta(X_t, \mu_{i,t}, \sum_{i,t}) \tag{4.1}$$

Where the parameters are K is the number of distributions, $\varpi_{i,t}$ is a weight associated to the 1th Gaussian at time t with mean $\mu_{i,t}$ and standard deviation $_{i,t}$. η is a Gaussian probability density function as formula (4.2):

$$\eta(X_t, \mu, \sum) = \frac{1}{(2\pi)^{n/2} \left| \sum \right|^{1/2}} e^{\frac{1}{2}(X_t - \mu)\sum^{-1}((X_t - \mu))} \tag{4.2}$$

K determined the multimodality of the background and by the available memory and computational power. It is proposed to set K from 3 to 5.

The initialization of the weight, the mean and the covariance matrix is made using an EM algorithm (Expectation maximization Algorithm). Stauffer and Grimson used the K-mean algorithm for real time consideration.

Once the parameters initialization is made, a first foreground detection can be made and then the parameters are updated. Firstly, Stauffer and Grimson used as criterion the ratio $\gamma_j = \varpi_j / \delta_j$ and ordered the K Gaussians following this ratio.

This ordering supposes that a background pixel corresponds to a high weight with a weak variance due to the fact that the background is more present than moving objects and that its value is practically constant. The first B Gaussian distributions which exceed certain threshold T are retained for a background distribution, as formula (4.3):

$$B = \arg\min_b (\sum_{i=1}^{b} \omega_{i,t} > T) \tag{4.3}$$

The other distributions are considered to represent a foreground distribution. Then, when the new frame incomes at times t+1, a match test is made for each pixel. A pixel matches a Gaussian distribution if the Mahalanobis distance, as formula (4.4).

$$sqrt((X_{t+1} - \mu_{i,t})^T \sum_{i,t}^{-1} (X_{t+1} - \mu_{i,t})) < k\sigma_{i,t} \tag{4.4}$$

Where k is a constant threshold, then two cases can occur:

Case 1: A match is found with one of the K Gaussians. In this case, if the Gaussian distribution is identified as a background one, the pixel is classified as background else the pixel is classified as foreground.

Case 2: No match is found with any of the K Gaussians. In this case, the pixel is classified as foreground.

At this step, a binary mask is obtained. Then, to make the next foreground detection, the parameters must be updated. Using the match test, two cases can occur like in the foreground detection:

Case 1: A match is found with one of the K Gaussians. For the matched component, the update is done as follows as formula (4.5):

$$\omega_{i,t+1} = (1-\alpha)\omega_{i,t} + \alpha \tag{4.5}$$

Where, α is a constant learning rate

$$\begin{aligned} \mu_{i,t+1} &= (1-\rho)\mu_{i,t} + \rho X_{t+1} \\ \sigma^2_{i,t+1} &= (1-\rho)\sigma^2_{i,t} + \rho(X_{t+1} - \mu_{i,t+1})(X_{t+1} - \mu_{i,t+1})^T \\ \rho &= \alpha\eta(X_{t+1}, \mu_i, \sum_i) \end{aligned} \tag{4.6}$$

For the unmatched components, μ and . are unchanged, only the weight is replaced by formula (4.6):

$$\varpi_{j,t+1} = (1-\alpha)\omega_{j,t} \tag{4.7}$$

Case 2: No match is found with any of the K Gaussians. In this case, the least probable distribution k is replaced with a new one with parameters. Once the parameters maintenance is made, foreground detection can be made and so on.

2)Image post-processing operations

The outputs of foreground region detection algorithms explained in previous generally contain noises, and therefore it is necessary to conduct special post-processing. There are various noises in foreground detection such as camera noise, reflectance noise, and background colored object noise, shadows and sudden illumination change and so on. Morphological operations are applied to the foreground pixel map in order to remove noise that is caused by the first three kinds of noises. In this paper, the last kind of noise is neglected.

3)Connected regions detection

After detecting foreground regions and applying post-processing operations to remove noise and shadow regions, the filtered foreground pixels are labeled by connected regions, then the bounding boxes of these regions are calculated.

4)Region level post-processing

After removing noise, some small regions remain due to inaccurate object segmentation. In order to eliminate these regions, the average region size (n) in terms of pixels is calculated for each frame and regions that have smaller sizes than a fraction (α) of the average region size ($Size(region) < \alpha * \gamma$) are deleted from the foreground pixel map. Due to segmentation errors, some parts of the objects are found as disconnected from the main body. In order to correct this defect, the bounding boxes of regions that are close to each other are merged together and the region labels are adjusted.

5)Extracting object features

Once regions have been segmented, features of the corresponding objects from the current image can be extracted. These features are size (S), center of mass (C_m), and so on. The definition of C_m is as follows:

$C_m(xC_m, yC_m)$, of an object O, using the following equation (4.7), the center of mass can be calculated:

$$xC_m = \frac{\sum_i^n x_i}{n}, yC_m = \frac{\sum_i^n y_i}{n} \tag{4.8}$$

Where n is the number of pixels in O.

4.5.2Management of Intrusion Detection

The intrusion detection management is working in two steps, the processing of video sequence and the sending SMS alert.

1) The processing of video sequence

Here the most important thing is to be able to keep track of what 's going on during an intrusion effectively happened. And at the same time, the module should be able to save sequence of video footages that can be used as evidence in case of theft. At a

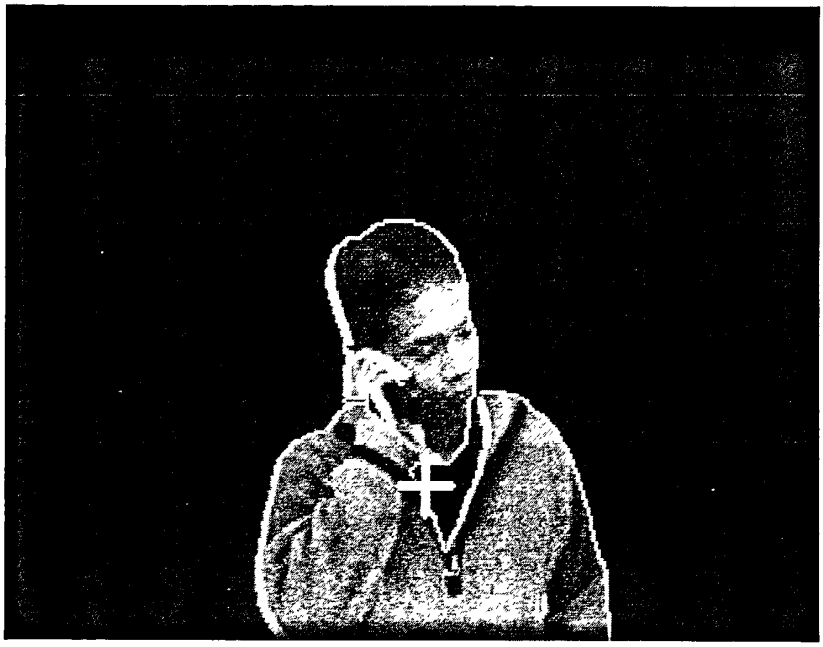

Figure 4.11 Motion detection procedure result

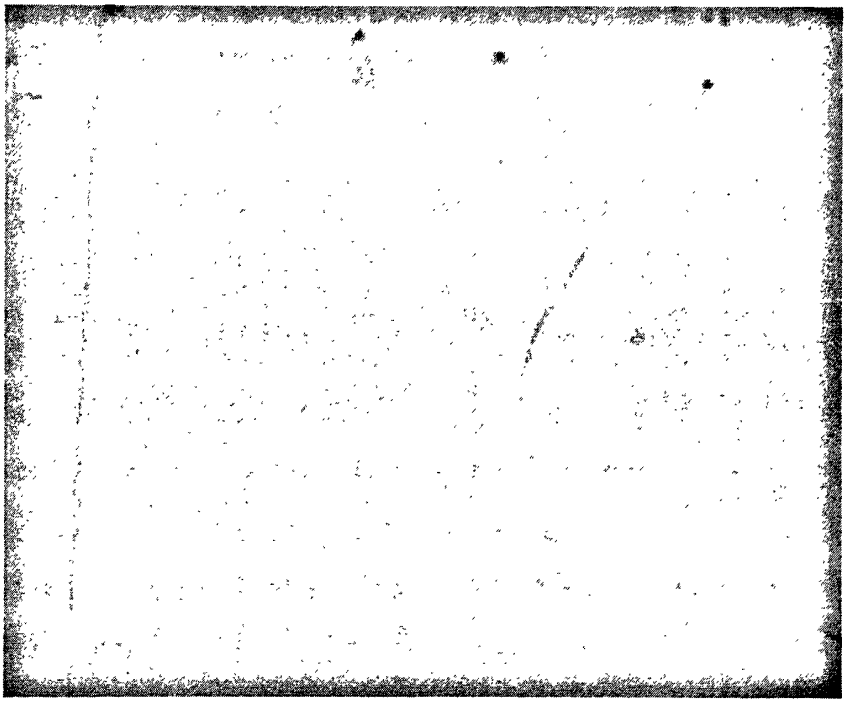

Figure 4.12 Background image chosen

Figure 4.13 Motion detection procedure result

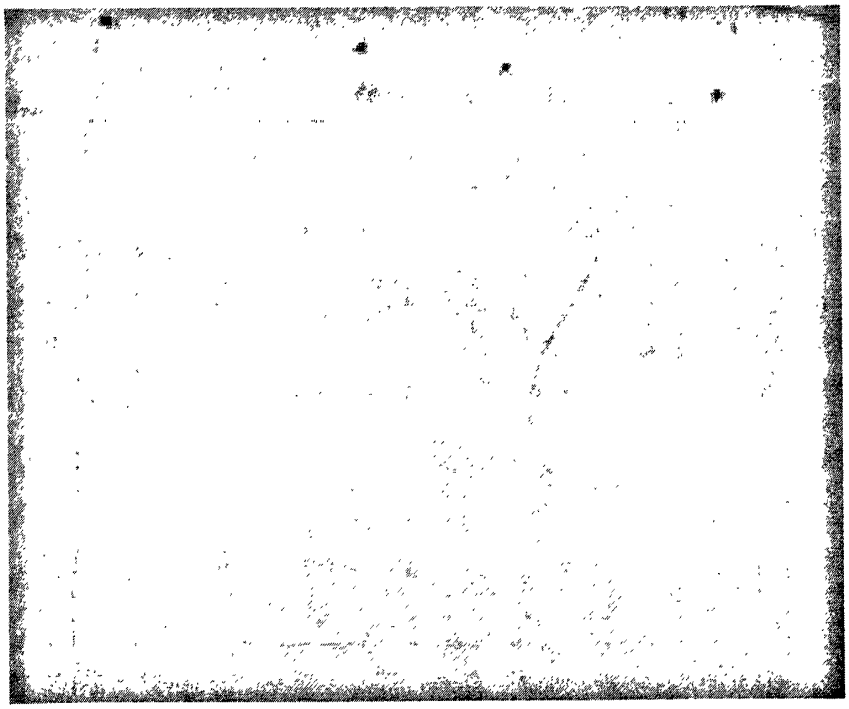

Figure 4.14Background image chosen

moment of a movement detected, the images are captured and sent across the network and at the same time the image in a directory is saved automatically for a certain period of time.

2) The sending SMS alert: About sending SMS, it will be enough that the user may have a subscription with a mobile phone operator providing the receiving SMS service.

4.5.3Experiment results of intrusion detection module

Here we have chosen a background image on what the motion detection algorithm will make its calculations, to see if a change has occurred in the surveillance zone.

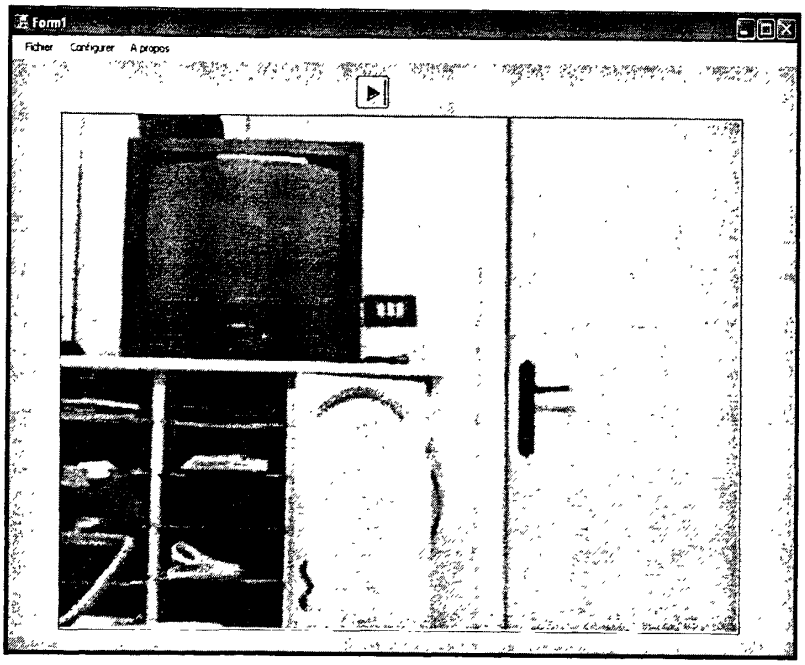

Figure 4.8 Background image chosen

Figure 4.9 Declaration of intrusion and triggering alarm

As the figure 4.9 shows, we can see the intrusion of a person in the monitored site and automatically triggering an alarm that will be followed by sending SMS to the user. Figure 4.10-4.19 are the same results.

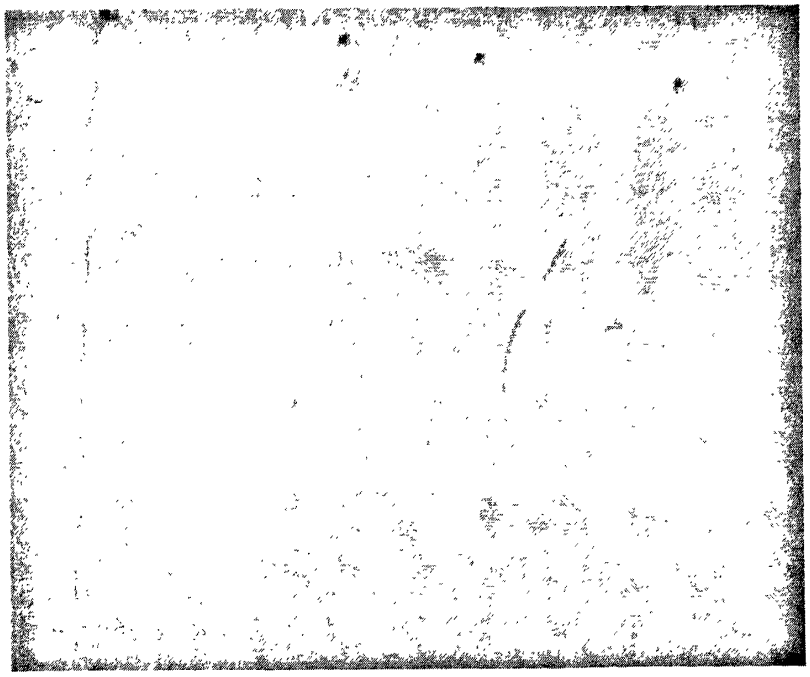

Figure 4.10 Background image chosen

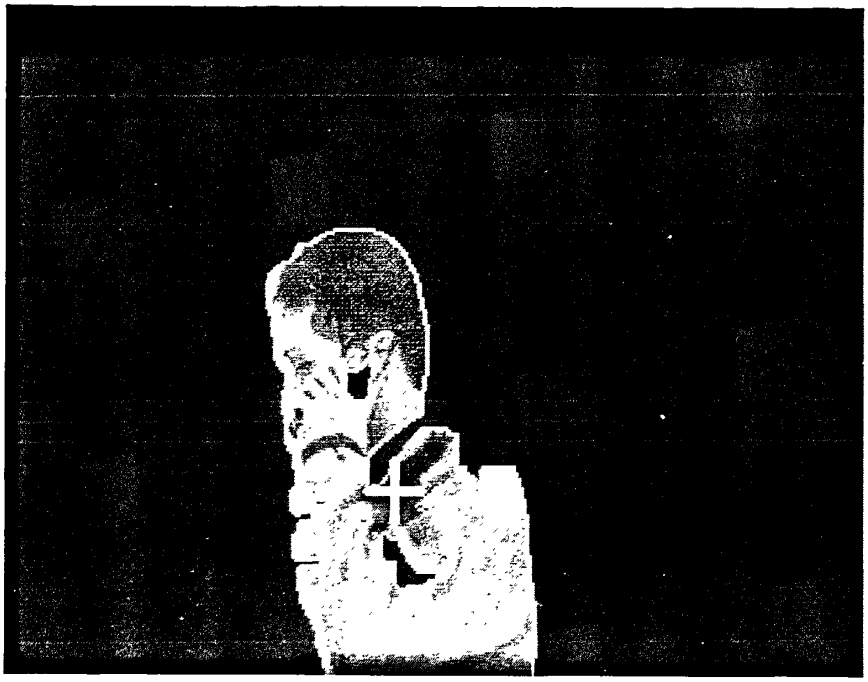

Figure 4.15Motion detection procedure result

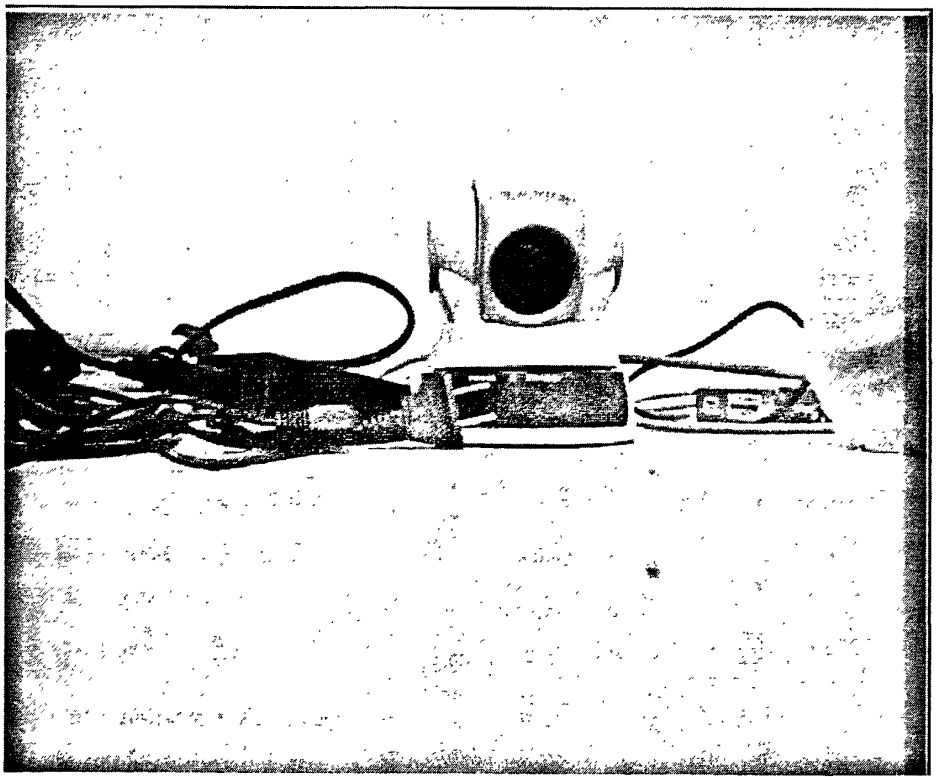

Figure 4.16 Background image chosen

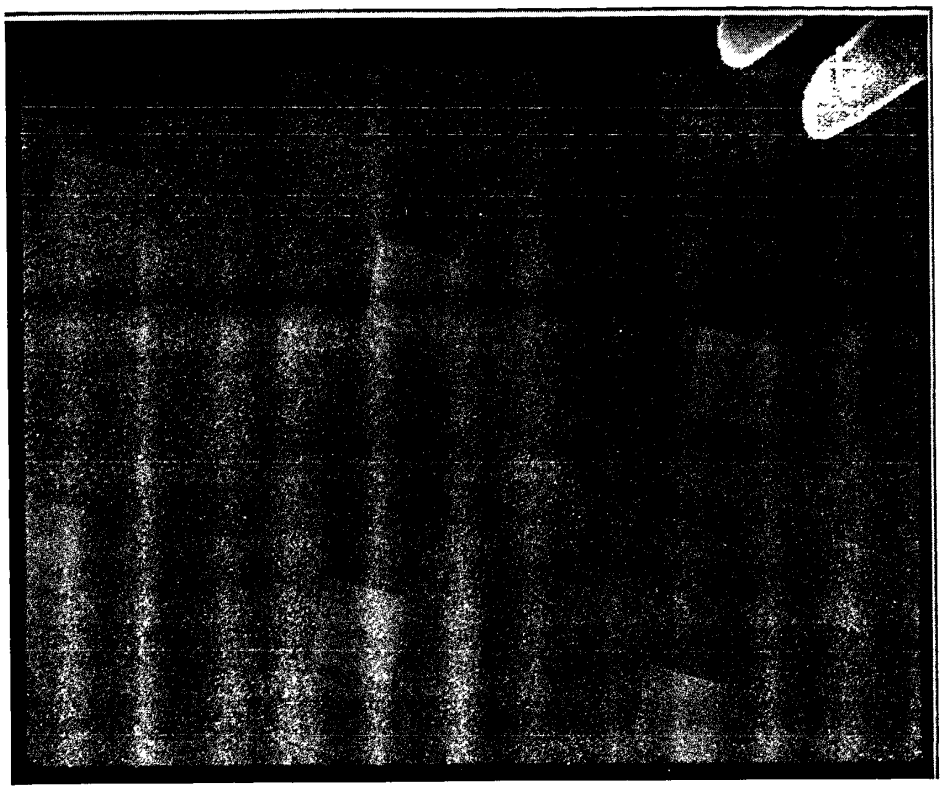

Figure 4.17 Motion detection procedure result

The first image of each couple respective images shows, the motion detection taking place in the same environments but with different sceneries. The second image of each couple image shows the motion masks generated from the first images in case of motion detection.

4.6 Presentation of management module

4.6.1 Software environment

1) Operating system Windows XP SP2

2) Development environment

(1)Microsoft visual .Net 2005;

(2)The Microsoft office XP 2003;

(3)Power designer 12.1, for modeling UML.

4.6.2Choice of programming language

The programming language of this mobile video surveillance system was implemented with C# on the platform Dot.Net 2005.

This platform was chosen because:

1) It allows object-oriented programming

2) It is known for the optimization of its code and also its speed running availability

3) It has a compiler for C#

4) It offers advanced tools for new technologies

5) It provides the C# language, a language derived from C++ but more powerful, flexible and easy to be used.

4.6.3Main interface of application

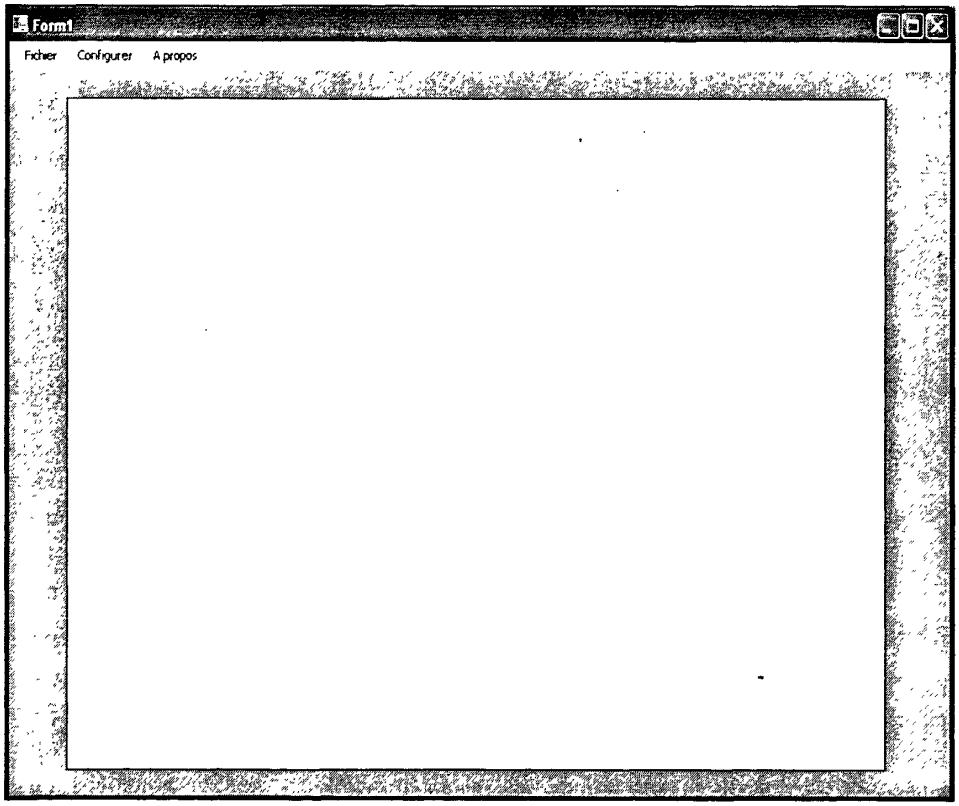

Figure 4.4 Interface of application

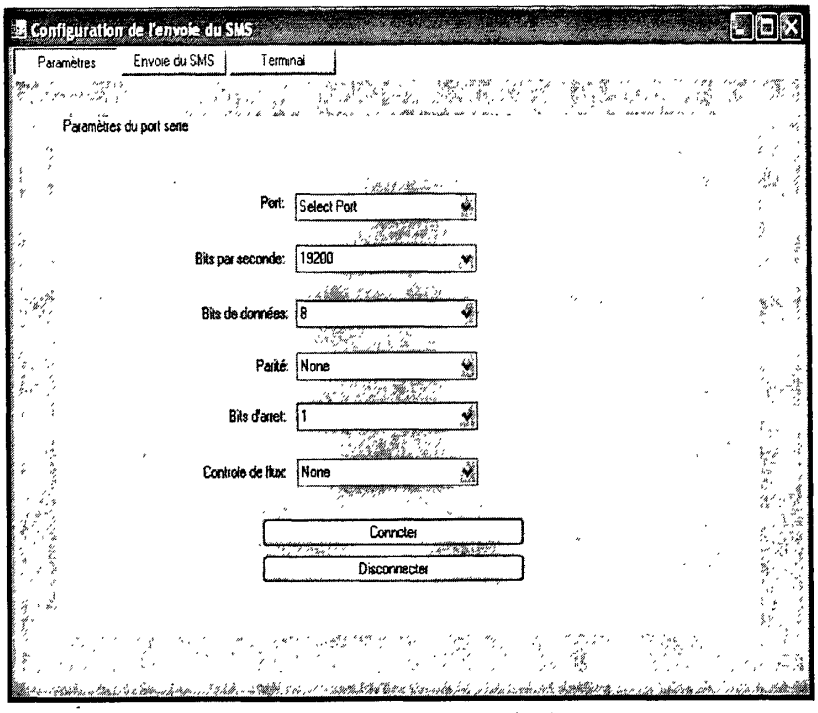

Figure 4.5 Configuring communication port

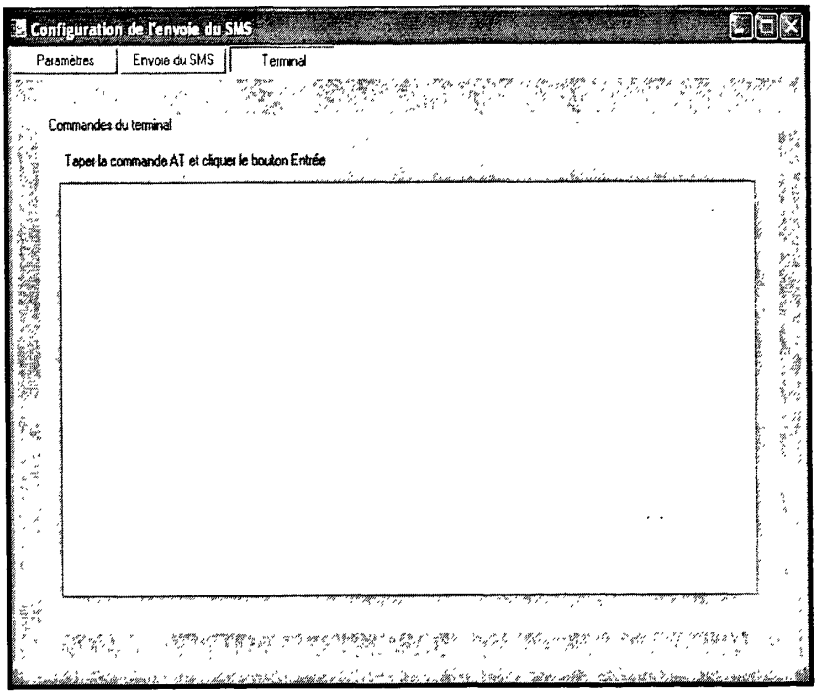

Figure 4.6 Terminal interface

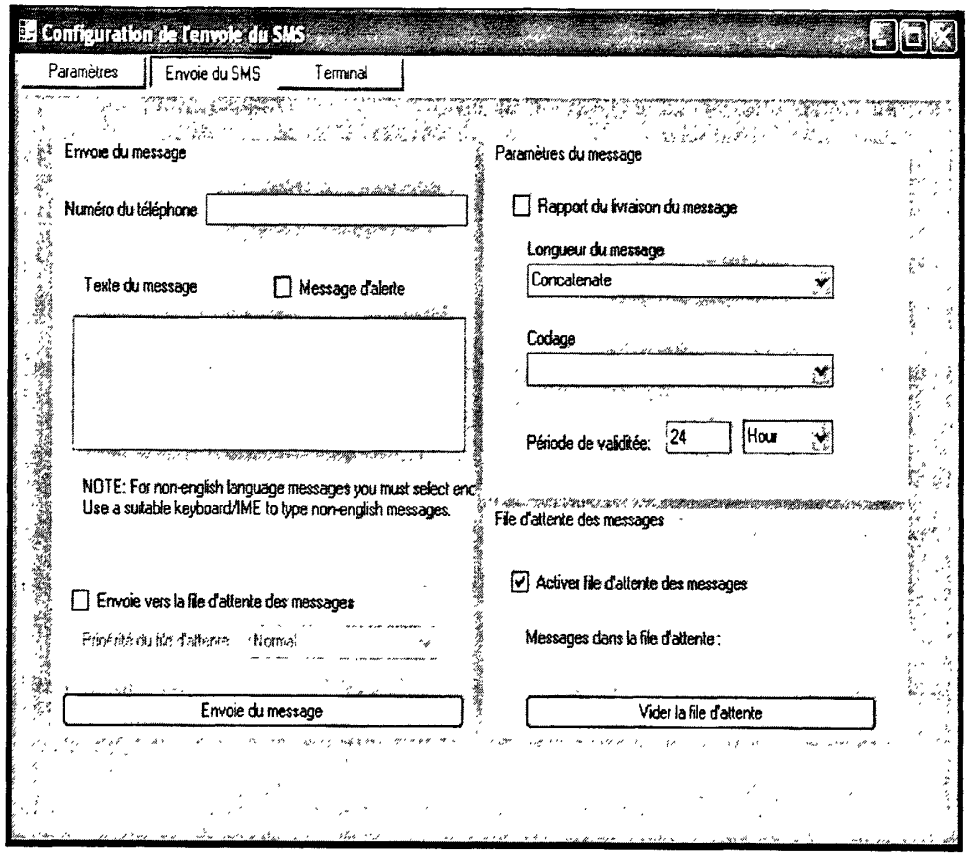

Figure 4.7 Sending SMS interface

4. 7Summary

This chapter showed the implementation of every part in the system, including module of input and output, which are built by existing hardware, coding and compression module, which is supplied in the camera, communication module, which determines the style of communication, the database, which is built by sql server 2003, intrusion detection module, which is implemented into a software, and as the core of the system, we firstly designed an intrusion detection algorithm based on the adaptive Gaussian mixture model, secondly tested its function by several experiments. After that, we showed the video surveillance interface, and the system is proved work well. We remember that, there are several ways to achieve the same results obtained in this part of the work, and we choose the simplest and most detailed to describe the outcome of realization.

Chapter 5 Summary and Work in Future

In this work that we called: Research on a video surveillance system with 3G mobile phone, we had the main purpose to design a system that can provide real time video surveillance of a place or any object by means of a mobile phone.

The proposed system allows the user, for example being able to keep an eye out for a precious object despite of the constraints of distance that often arises when we are moving. The system also allows the user access to the internet provided through mobile phone operators, to be able to receive a SMS alerts on its mobile phone in case of intrusion under surveillance in real time.

We built this system from design to its implementation. We analyzed nowadays video surveillance systems and designed one. After obtaining the functions of the system, we divided the system into 6 parts: input module and output module, management module, coding and compression module, communication module, database, intrusion detection module. We built them from one to another. As the core of the system, we designed the intrusion detection algorithm and tested it. According to the results we got, the system can work well.

Looking toward the future, with technological advances in the fields of information, communication and mobile telephony, the exploitation of this kind of system with several image acquisition sources could bring something more to the resolution of the some conflicts of interests.

Because of the limitation of time and material and financial resources that we met for the realization of this research, our future research may be focused on a system that can handle several images acquisition sources with more complicated backgrounds and more moving objects.

References

[1] M. Shah, O. Javed, and K. Shafique, Automated visual surveillance in realistic scenarios, IEEE Multimedia [J], 2007, 14 (1) : 30-39.

[2] M. W. Green. The appropriate and e ective use of security technologies in U.S. schools: a guide for schools and law enforcement agencies. National Institute of Justice, 1999:12-13

[3] Nilsson. F, Intelligent Network Video: understanding modern video surveillance system, Boca Raton ,CRC Press,2009.

[4] Taaghol, R. Evans, B. G, Buracchini. E, De Gaudezi, R. Gallinaro, G. Lee, J. H, Kang, and C. G: satellite UMTS / IMT-2000 W-CDMA air interfaces. IEEE communication magazine, 1999, 37 (9) .

[5] I. Pavlidis, V. Morellas, P. Tsiamyrtzis, and S. Harp, Urban surveillance systems: from the laboratory to the commercial world, Proceedings of the IEEE, 2001, 89 (10) : 1478-1497.

[6] Zhan. B, Monekosso.D, N, Remagnito. P Velastin. S. A. Xu. L. Q. crowd analysis: a survey. Machine vision and applications,2008, 19: 5-6.

[7] Benoit Barraque, Advantages of UMTS and the current state of its deployment in the world. Claude Bernard University, 2004.

[8] Margherita Pagan, Charles H. Fine, Value network dynamics in 3G - 4G wireless communicationnns: A systems thinking approach to strategic value assessment, Journal of Business Research,2008, 61:1102 - 1112

[9] Weijia Jia Haohuan Fu and Ji Shen, Object-oriented Design and Implementations of 3G-324M protocol Stack , department of computer science city university of Hong Kong, 2003.

[10] L. Marchesotti, L. Marcenaro, and C. Regazzoni, Dual camera system for face detection in unconstrained environments in Proc, of the

international conference on Image Processing (ICIP), Barcelona, Catalonia, Spain, 2003

[11] Korhonen, Juha. Introduction to 3G Mobile Communications (second edition), Norwood, MA, USA: Artech house, incorporated, 2003.

[12] Benoit Barraque, Advantages of UMTS and the Current State of its Deployment in the World. Claude Bernard University, 2004.

[13] Transport Canada. Reseau de television en circuit ferme: manuel de reference pour l`utilisation dans des applications de securite, 2008.

[14] DI Axiotis, Services in interworking 3G and WLAN Environment, IEEE wireless communications, 2004:1536-1284,.

[15] Mitrevski. K, Evolution to 3G, graduation thesis, ETF, Belgrade, 2000.

[16] WCDMA, Wireless network design and optimize, post and telecom Press, 2007.

[17] W. M. Thames. From Eye to Electron - Management Problems of the Combat Surveillance Research and Development Field. IRE Transactions on Military Electronics, 1960, 13-4(4):548 - 551.

[18] T Komatsu, K Aizawa, T Igarashi, Signal-processing based method for acquiring very high resolution images with multiple cameras and its theoretical analysis, IEEE proceedings Communications, 1993, 140:19-24.

[19] Robert M Haralick, Linda D Shapiro, Computer and Robot Vision, Addison-Wesley Longman Publishing Co., Inc. Boston, MA, USA, 1992.

[20] Ellis, T. Multi-camera video surveillance system, Proceedings of 36th Annual 2002 International Carnahan Conference on Security Technology, 2002:228-233.

[21] S. Hampapur, A. Brown, L. Connell, J. Ekin, A. Haas, N. Lu, M. Merkl, H. Pankanti. Smart Video Surveillance: Exploring the Concept of Multiscale

Spatiotemporal Tracking. IEEE Signal Processing Magazine, 2005, 22 (2):38 - 51.

[22] Cai, Q. Tracking human motion using multiple cameras, Proceedings of the 13th International Conference on Pattern Recognition, 1996:25-29.

[23] Sudipta N. Sinha, Marc Pollefeys. Pan - tilt - zoomcamera calibration and high-resolution mosaic generation, Computer Vision and Image Understanding, 2006, 103 (3):170 -183.

[24] Rainer Bernhardt, Video surveillance apparatus with a dome camera and a protective housing, United States Patent Application Publication, 2001:1-4.

[25] J. Peipe, M. Stephani. Performance evaluation of a 5 megapixel digital metric camera for use in architecture photogrammetry, The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, XXXIV(5/W12): 259-261.

[26] Jones, B.F. A reappraisal of the use of infrared thermal image analysis in medicine, IEEE Transactions on Medical Imaging. 1998, 17(6):1019-1027.

[27] K. Chrzanowski, S.N. Park. Evaluation of thermalcameras for non-destructive thermal testing applications, Infrared Physics & Technology, 2001, 42(2):101 -105.

[28] David, J. Calibrating pan-tilt cameras in wide-area surveillance networks, Proceedings 9th IEEE International Conference on Computer Vision, 2003, 1:144-149.

[29] Andrew D. Richardson, Julian P. Jenkins, Bobby H. Braswell, David Y. Hollinger, Scott V. Ollinger and Marie-Louise Smith. Use of digital webcam images to track spring green-up in a deciduous broadleaf forest, Oecologia. 2007, 152(2): 323-334.

[30] Skodras, A. Christopoulos, C. Ebrahimi, T. The JPEG 2000 still image compression standard, Signal Processing Magazine, IEEE. 2001, 18(5):36-38.

[31] Didier Le Gall. MPEG: a video compression standard for multimedia applications, Commu- -nication of the ACM. 1991, 34(4):46-58.

[32] Hans-Hellmut Nagel. On the estimation of opticalflow: Relations between different approaches and some new results, Artificial Intelligence. 1987, 33(3): 299- 324.

[33] T. Blasaka, Rachid Deriche. Recovering and Characterizing Image Features: Using an Efficient Model Based Approach[Technical Report 2422], INRIA, 1994.

[34] David G Lowe. Object recognition from local scale-invariant features. InICCV, 1998:1150- 1157.

[35] Yigithan Dedeoglu. Moving object detection, tracking and classification for smart video surveillance, 2004.

[36] T. Bouwmans, F. EI Baf, B. Vachon. Background modeling using mixture of gaussons for foreground detection-A survey. Computer Science, 2008, 1(3):219-237.

Acknowledgement

Getting academic degree in master program in Information and Communication Engineering is the main purpose of presentation, and this research project is performed in Dalian Maritime University. This project is called: Research on a video surveillance system with 3G mobile phone.

My previous studies in Electronics (Switching) which is one of the branch of telecommunication, and my current studies on Information and Communication Engineering were my inspiration source for this thesis project. The technology of digital information processing in real time develops fast and that brings me opportunities and availability of mobility to do some research in telecommunication fields.

First I would like to thank my supervisor associate Prof. Wang X. N for his experience, knowledge and time that he gave to me to carry out this work despite the difficulties of communication because of language.

I'm going to thank also, all of my classmates who supported me for the same cause.

Furthermore, I would like to thank my mother for her permanent and multifaceted support, my brother and my entire countrymen living in Dalian. I did not forget all of my friends who supported me in different ways.

Dalian,May.20.2012

Borgea Dumas Ngoma Yeke(包格)

研究生履历

姓　　名	Borgea Dumas Ngoma Yeke(包格)
性　　别	男
出生日期	1983 年 4 月 23 日
获学士学位专业及门类	Electronics/Engineering
获学士学位单位	Superior institute of applied technical
获硕士学位专业及门类	信息与通信工程/工学
获硕士学位单位	大连海事大学
通信地址	辽宁省大连市凌海路 1 号
邮政编码	116026
电子邮箱	nborgea@yahoo.fr

中文摘要

视频监控是通过视频对某一场所进行监控,它是公共安全防范系统的重要组成部分，是一种防范能力较强的综合系统。视频监控以其直观、准确、及时和信息内容丰富而广泛应用于许多场合。视频监控在日常生活中的各个领域提供了一个基于物理和逻辑安全的基础，随着科学技术的发展，视频监控越来越依赖于信息与通信新技术。在二十世纪七十年代，视频监控开始于英格兰，目的是防止恐怖分子袭击。此举取得了巨大的成功，并逐渐开始发展。安装视频监控系统的动机是多种多样的，然而，最主要的目的是保证公共或私人的安全以及保护个人财产和不动产。

经过多年的发展，视频监控技术已由早期模拟设备为主的第一代视频监控系统发展到目前的数字智能视频监控。随着3G技术难点的突破以及3G网络的发展，无线视频监控的实现成为了可能。

视频监控系统由摄像、传输、控制、显示、注册登记五大部分组成。首先由摄像机通过同轴视频电缆或网线将视频图像传输到控制主机，控制主机再将视频信号分配到各监视器及录像设备，同时可将需要传输的语音信号同步存储。通过控制主机，操作人员可发出指令，对云台的上、下、左、右的动作进行控制及对镜头进行调焦变倍的操作，并可通过控制主机实现在多路摄像机及云台之间的切换。利用特殊的录像处理模式，可对图像进行录入、回放、处理等操作，使录像效果达到最佳。

3G 是一项结合无线通信技术和互联网等多媒体技术的新技术，它在传输声音和数据的速度上有很大的提升，并能够在全球范围内更好地实现无线漫游，而且能够处理图像、音乐、视频流等多种媒体形式，提供包括网页浏览、电话会议、电子商务等多种信息服务。

为了确保用户可以从任何地方访问并监测系统，本文提出了基于 3G 的视频监控系统，该系统使用 3G 手机、相机、摄像头、互联网（ADSL）或 3G 网络进行本地或异地视频监控以及入侵者报警。本研究的目的是基于以下三个问题：(1)如何

使用户系统在任何地方获得实时的访问？(2)用户如何获得图像序列并通过什么样的资源传送？(3)如何在系统中执行入侵过程的检测？为了回答这些研究问题，本文首先对视频监控系统进行综述，进而设计了基于3G的视频监控系统，并给出了入侵检测算法，最后实现了基于3G的视频监控系统。主要工作如下：

1) 对视频监控系统进行综述

视频监控系统主要由前端音视频数据采集设备、传送介质、终端监看/监听设备和控制设备等四部分组成。视频监控运用传感技术、监控摄像技术、通讯技术和计算机技术等组成一个多功能智能化处理系统。该系统经过了模拟监控系统、硬盘录像机监控系统、网络监控系统、智能监控系统四代的发展历程，分别说明：

(1)模拟监控系统：全系统采用模拟信号，各个系统之间独立；存储介质为录像磁带，使用和保存成本高，特别不适宜全时段录像；从录像存储的磁带中检索历史图像资料难度大、成本高；图像清晰度受限制；功能单一。

(2)硬盘录像机监控系统：应用计算机技术，降低了存储和历史图像检索的成本，但仍未解决连点成片的网络问题，管理性和资源共享性差。网络监控系统：采用IP化全数字新技术构网，系统克服了前两代的缺点，显现出如下的优势：组网和应用不受地域限制、管理性强、图像清晰度高、存储成本降低、可实现的功能强、平台开放等。

(3)智能视频监控：智能视频监控是利用计算机视觉技术对视频信号进行处理、分析和理解，在不需要人为干预的情况下，通过对序列图像自动分析对监控场景中的变化进行定位、识别和跟踪，并在此基础上分析和判断目标的行为，能在异常情况发生时及时发出警报或提供有用信息，有效地协助安全人员处理危机，并最大限度地降低误报和漏报现象。其中，运动检测是智能视频监控系统的基础。运动目标检测是指在序列图像中检测出变化区域并将运动目标从背景图像中提取出来。目标分类、跟踪和行为理解等后处理过程仅仅考虑图像中对应于运动目标的像素区域。运动目标的正确检测与分割对于后期处理非常重要。场景的动态变化，如天气、光照、阴影和杂乱背景的干扰，使得运动目标检测和分割变得相当

困难。运动目标检测常用的方法有以下三种：①光流法：主要任务是计算光流场，即在适当的平滑性约束条件下，根据图像序列的时空梯度估算运动场，通过分析运动场的变化对运动目标和场景进行检测与分割。② 帧差法：基本原理是在图像序列相邻的两帧或者三帧采用基于像素的时间差 分通过阈值化来提取图像中的运动区域。③特征点检测法：其基本思想是计算两幅图像的特征点，将两幅图像的特征点进行匹配，如果匹配不上，说明相对位置发生变化，因而可以提取运动区域。

总的来说，随着通信技术以及计算机视觉技术的发展潮流，视频监控系统正朝着宽带化、IP化、移动化和智能化方向发展。

2) 设计了基于3G的视频监控系统

本文设计的基于3G的移动视频监控系统功能、模块以及数据结构流图描述分别如下：

(1)系统的主要功能包括：

①系统应该能够轻松而有效地与所有外设进行能够自动进行通讯。

②实时视频数据的流畅显示，用户的请求通过其3G手机接入系统后，用户必须能够实时查看视频序列。

③能够对静止监测点进行入侵者检测，并自动报警。

(2)为实现系统要求的功能，系统由五个模块组成，各模块功能分别为：

①管理模块

管理模块的功能是对于入侵检测的提醒功能设置参数，并设置图像的大小和格式以适应图像的传输。

②输入和输出模块

输入和输出模块功能是视频采集、视频编码、调用入侵检测模块进行入侵检测以及调用通讯模块进行视频传输。该模块由音视频采集设备（摄像机或网络摄像头、拾音器）、云台、视频编码卡以及相关软件组成。

③编码和压缩模块

编码和压缩模块功能是，将要传输的数据流进行压缩，以便于数据传输。

④通讯模块

通讯模块的功能由以下三部分组成：用户在有监控申请时，通过 3G 网络传送给视频采集端；采集端将得到的数据流进行压缩，传输给用户；当出现入侵检测时，采集端将报警信号传输给用户。

⑤数据库模块

数据库模块的功能是将检测到的入侵信号进行存储，方便用户调用。

⑥入侵检测模块

入侵检测模块的功能是确定检测区域中要检测的区域或运动物体，并将分析得到的运动轨迹传输给通讯模块。

(3) 系统的数据流程

基于 3G 的视频监控首先要将监控系统和用户使用的移动手机连接到 3G 网络，并通过管理模块设置各项参数。当用户端发送监控申请时，通讯模块通过 3G 网络将信息传送给视频采集端，视频采集端利用输入和输出模块的网络摄像头，采集视频信息，由通讯模块传输给用户；当目标入侵时，入侵检测模块将报警信息传送给通讯模块，通讯模块通过 3G 网络将信息传送给视频采集端。

3) 给出了入侵检测算法

视频图像的入侵物体检测算法分五个步骤来进行目标检测。第一步是背景初始化以及前景检测。本文采用了自适应在线高斯混合模型来建立背景模板。通过使用背景模板和当前视频图像来检测前景像素。该检测过程依赖于使用的背景模板，且能更新背景模版，以适应动态场景变化。第二步，由于照相机噪声和环境

影响使得检测到的前景图像含有噪声。本文采用膨胀、腐蚀等图像后处理操作来消除噪声，得到过滤后的前景图像。第三步，通过连通区域标记法找到各候选目标区域，计算边界矩形。第四步，通过相关规则，去掉残留的小区域。最后，提取出目标的质心等特征。各步骤详细描述如下：

(1)背景初始化以及前景检测

本文使用背景模版和图像后处理技术相结合的方法建立前景图像且提取每帧图像的物体特征。背景模板通常包含两个阶段：初始化和更新。本文主要使用自适应在线高斯混合模型来进行前景检测，基本算法如下：

每一个像素在RGB色彩空间由它的强度表征，当前观测到的可能的像素值，由下式表示：

$$P(X_t)=\sum_{i=1}^{K}\varpi_{i,t}\eta\left(X_t,\mu_{i,t},\sum_{i,t}\right) \tag{1}$$

其中，K表示高斯模型的个数，ϖ和时间t时第i个高斯模型有关的加权，$\mu_{i,t}$表示均值，$\sum_{i,t}$表示方差，η表示高斯概率密度函数。

所以，每一个像素值由K个混合高斯模型进行表征，一旦背景模型固定了，混合模型的不同参数也要初始化。

K决定了背景的多峰性，一般K为3到5的数字，均值和方差的初始化利用EM算法求得。

一旦初始参数设置完成，第一幅图像的背景构建就可以完成，然后更新参数，根据标准$\gamma_j=\varpi_j/\delta_j$，排列$K$个高斯模型的顺序，当方差较小时，加权值要大，因为背景图像比运动物体存在的更久，而且它的像素值几乎不变。

第一个超过一定阈值T的B被选为背景模版，如下式：

$$B=\arg\min_b\left(\sum_{i=1}^{b}\varpi_{i,t}>T\right) \tag{2}$$

其他的被认为表示前景分布，因此当$t+1$时刻，新的一帧到来的时侯，每一个像素就要进行匹配测试，满足以下式（3）就要匹配一个高斯分布：

$$sqrt\left(\left(X_{t+1}-\mu_{i,t}\right)^T\sum_{i,t}^{-1}\left(X_{i,t}-\mu_{i,t}\right)\right)<k\delta_{i,t} \tag{3}$$

k 是一常量，可以设置为2.5，这样就会出现两种情况：

第一种：匹配上一个高斯模型，如果这个高斯分布是被定义的背景图像，那么这个像素值就被归类为背景图像，否则，就为前景图像。

第二种：找不到可以匹配的高斯模型，这种情况下，就被定义为前景图像。

这一步之后，就可以找到一个二值化模版，然后，为了下一个的前景图像检测，参数必须更新。利用以上匹配测试公式，会出现两种情况：

第一种：有一个相匹配的高斯分布。

这种情况下，对于相匹配的部分，利用以下式（4）更新加权值：

$$\varpi_{i,t+1}=\left(1-\alpha\right)\varpi_{i,t}+\alpha \tag{4}$$

其中，α 是一个不断的学习率

$$\begin{aligned}&\mu_{i,t+1}=\left(1-\rho\right)\mu_{i,t}+\rho X_{t+1}\\&\delta_{i,t+1}^2=\left(1-\rho\right)\delta_{i,t}^2+\rho\left(X_{t+1}-\mu_{i,t+1}\right)\left(X_{t+1}-\mu_{i,t+1}\right)^T\\&\rho=\alpha\eta\left(X_{t+1},\mu_t,\sum_t\right)\end{aligned} \tag{5}$$

对于不相匹配的部分，μ 和Σ是不变的，只有加权值改变，如式（6），

$$\varpi_{j,t+1}=\left(1-\alpha\right)\varpi_{j,t} \tag{6}$$

第二种：没有相匹配的高斯模型

在这种情况下，最不可能的高斯分布就会被替代。

一旦参数制定，前景检测图像就可以得到。

通观整个高斯模型，它主要是有方差和均值两个参数决定，，对均值和方差的学习，采取不同的学习机制，将直接影响到模型的稳定性、精确性和收敛性。由于我们是对运动目标的背景提取建模，因此需要对高斯模型中方差和均值两个参数实时更新。

(2)图像后处理技术

由于前景区域检测算法的输入包含噪声，所以有必要增加图像后处理操作。引起前景图像检测的噪声有：相机噪声，反射噪声，有色物体噪声，阴影和照度突然变化的噪声。前三种噪声使用腐蚀、膨胀进行处理，而对于阴影和照度突然变化的噪声，本文不进行处理。

(3)相关区域检测

前景区域检测完成后，应用图像后处理操作来消除噪声和阴影区域。对过滤后的前景像素进行区域标记，并计算各区域的边界盒。

(4)区域水平的后处理

经过消除噪声，仍然残留下一些由于不精确物体分割的小区域。为了消除这类的小区域，根据像素，计算出每一帧的平均区域尺寸γ，若区域尺寸满足$size(region)<\alpha*\gamma$ α为小数，该区域将被删除。由于分割错误，一些对象的部分成分与主体分开。为了避免这种错误，区域的边界盒接近的区域被合并，且区域标签被更新。

(5)物体特征提取

分割区域后，从当前图像的对应物体中提取特征。特征有面积S，质心c_*，颜色直方图H_*。物体O，质心$C_m=(xC_m,yC_m)$，使用下面的公式（7），

$$xC_m=\frac{\sum_i^n x_i}{n},yC_m=\frac{\sum_i^n y_i}{n} \tag{7}$$

n为图像O的像素数。

4) 实现了基于3G的视频监控系统

实现基于3G的视频监控，系统各模块所需要的软硬件环境以及要满足的各项指标分别如下：

(1)图像采集和处理模块

对于图像的采集和处理，本文使用的网络摄像头型号为 Logitech Pro HD C525，它的高清视频捕获需要达到1280*720的像素值，图像可达到800万像素，且是自动聚焦的。

(2)通讯模块

对于通讯模块，本文使用的USB调制解调器型号为HUAWEI E153，它的无线网速要达到3.6Mbps，并且需要微型记忆卡槽，有数据和短信服务，即插即用。

(3)入侵检测模块

对于入侵检测，本文采用自适应在线高斯混合模型，检测入侵物体。对于入侵检测的管理可以分为两个步骤：视频序列的录制和发送短信提醒。

①录制视频序列：目的是保留入侵期间物体的移动轨迹，防止盗窃。基 于这一点，要实时进行运动检测，捕获图像，并通过网络发送，并自动保存图像。

②发送短信提醒：用户只需要订阅移动电话运营商提供的发送短信服务 就可以获得短信提醒服务。

3G 时代的到来开启了视频监控的新时代，随着 3G 网络的不断建设和完善，网络带宽和无线传输质量得到提高，从而使得视频监控达到更好的效果，更大的满足用户需求、提升用户感受。

本文设计并实现了基于3G的视频监控系统，通过利用自适应在线高斯混合模型，给出了入侵检测算法，并通过实验验证了算法以及系统的可行性。然而由于物力和财力的限制，本文设计的视频监控系统是基于单个图像采集源，不能同时观看多个采集源的视频信息，因而监控区域有限，不能实现无死角监控。因此我们未来的研究是基于开发一个移动视频监控应用程序，它可以处理多个图像采集源，该应用程序的实现，意味着我们能够在不同的站点进行监控，并可以跟随运动物体检测，也可以实现火灾检测。

关键词:视频监控；3G 手机；入侵检测；实时