Improved Observation and Communication with a Distributed Compound Vision Surveillance System

Nikolay Semenov, Kimberly Newman
The University of Denver
{nsemenov, kinewman}@du.edu

Abstract

Video surveillance systems are used in a variety of applications from home health monitoring to commercial security systems. A fixed or rotating camera has limitations as to the coverage area and resolution of the captured image. In this article we propose a new concept for the implementation of surveillance that shows significant improvement over the performance of distributed single cameras system by providing multiple angles of image capture without motion at each location. Our concept is called distributed compound vision.

1. Introduction

Video surveillance is widely used by government and commercial organizations to increase safety and security. The most basic surveillance system consists of one or more video cameras, a remote monitor with a recording device at the security post, and links between these modules for communication. Commercially available systems differ in camera resolution and degree of freedom as well as the number of cameras and their placement locations.

The major security weakness of traditional surveillance systems comes from camera installations. These have clearly identifiable locations, limited field of view (FOV), and a limited number of angles from which objects can be seen. Limited FOV is a security threat because an intruder can avoid detection by determining the limits of the field of view. Modern technology offers various ways to overcome FOV limitations. Traditional cameras are often mounted on rotational turrets to increase the total available FOV, but the instantaneous FOV is still limited. Fisheye cameras have wide FOV but there is a tradeoff in the image quality [1]. Catadioptric systems combine cameras and mirrors to create a 360-degree panoramic view around the installation, but the sector just in front of camera is hidden by the mirror [2]. Complex systems provide several cameras with overlapping fields of view, but within this complex zone of monitoring, security officers still have a choice of a small number (usually 1 to 3) of perspectives for an object in the monitored area. Damage of any camera decreases this number by 33% to 100%, which makes it a substantial loss of security. Increasing the number of cameras is a traditional way to solve this problem, but it is often not possible to install a large number of hanging cameras in an area of interest. Some areas, especially those of cultural value, like museums, do not want a large number of camera installations because they detract from appearance of the location.

In addition to concerns about the sensitivity and resolution of camera systems, vulnerability to attack is another issue for home or business security systems. In most cases, surveillance cameras are identifiable, and hence vulnerable. Cameras can be destroyed by an accurate shot or blocked by a layer of paint sprayed over the lens. Systems that can be embedded into the architecture or fabricated in such a way to reduce their profile provide better camouflage and protection from identification and subsequent attack.

Traditional cameras are camouflaged by painting them the color of the wall where the camera is mounted. However, the camera detection by an attacker is only delayed by this method. Another way to reduce the chance of detection is to use solutions that imitate compound eyes of insects or crustaceans. These eyes are arrays of very narrow-directional optical sensors. The artificial compound eyes can be made much smaller than traditional turret-based cameras, what helps them to avoid detection. The tradeoff is low image quality, so these units are generally used mostly for motion tracking [3], [4].

In order to address these issues, a biomimetic solution was envisioned. Our proposal is to combine features of traditional cameras and compound eye systems. The proposed system has a very wide FOV, and can see simultaneously in many directions. If distributed over a large surface, such optic systems can solve the vulnerability problems described above.
The major issue with this type of system is the signal processing required to reconstruct an image for observation.

2. Concept

Our proposed design concept for a video surveillance system is a distributed compound eye made of small traditional CMOS video cameras. The proposed construction of a module is shown in figure 1. Each module consists of five micro-cameras (object #1 on fig. 1). One camera is frontal, and the others form a symmetrical cross.

![Figure 1: Element design, side-cut and front view](image)

These cameras can be similar to those incorporated in modern laptops, i.e. iSight in Apple MacBook Pro, which are 6.35x6.35 mm. These cameras are capable of providing from 640x480 to 1280x1024 24-bit color auto-exposure video at 30 Hz, which makes a data stream of 26.37 to 112.5 MByte/s (object #3 on fig. 1). Video streams from all five cameras, respectively, are combined by a small chip (object #2 on fig. 1) to generate a stream of 131.84 to 562.5 MByte/s, which is within the capability of a single IEEE 1394b (Firewire 800) cable (object #4 on fig. 1). Each camera has an approximate FOV of 60 degrees in the vertical and horizontal planes. Therefore, each module has a maximum FOV (area #5 on fig. 1) of approximately 180 degrees in the vertical and horizontal planes. There are some miniscule blind spots (area #6 on fig. 1) in the closest proximity to the module. However, these do not limit the FOV due to their sizes and multiplicity of modules. The proposed module construction is immobile so motors or turrets are not required.

An array of micro-camera modules is incorporated in a panel as shown in figure 2. Regions of the design include an appearance layer (object #1 on fig. 2) that can be either paint, or a transparent panel, a textured decorative block, etc.

![Figure 2: Panel Design Concept](image)

The modules are connected through IEEE 1394b cables (object #2 on fig. 2) to data preprocessing elements (object #3 on fig. 2). Camera modules and preprocessing elements are powered by the power grid (object #4 on fig. 2), which is under the surface layer. After preprocessing, data is collected and transmitted through centralized cabling (object #5 on fig. 2). The whole panel is powered by an external plug (object #6 on fig. 2). There are very small blind spots (area #7 on fig. 2), near the peripheral zone of the panel. Also there are miniscule zones with reduced data acquisition overlap (area #8 on fig. 2). The FOV (area #9 on fig. 2) of panel is very wide, and the majority of points inside the FOV are viewed from many perspectives.

3. Efficiency estimations

The key features of the concept are high protection from physical attack and wide set of angles-of-view on objects in FOV. Some estimation will be given in this section.

The proposed system is intended to give a set of viewpoints on the objects in sight. The viewpoints are defined by three parameters – panel size, panel configuration and camera module density. For close distance surveillance (i.e. monitoring people who come to use ATMs), base size does not have to be large, since monitored objects (faces) are close and compact. Since 95% of people have a head breadth less than 165 mm [5] and estimating the average range between the user and ATM at 400 mm, a flat
A panel of approximately 630 mm wide will be able to monitor faces within ± 30° from the front, which gives a total angle of 150 degrees. A panel of this size could be made in the shape of a thin strip attached above or below the ATM interface. For high-distance surveillance, panels can be made much wider, or use other form factors like an arc with a large radius.

To estimate the safety efficiency of the system, a sample layout is chosen. For example, a panel shape is a rectangle 1 m high and 2 m wide, the area is 2 sq.m. The rectangle is covered with 50x50 mm square grid. Figure 1 demonstrates the concept. Note that not every grid line is drawn. Small squares represent the location of camera modules in the corners of the grid, and there are no camera modules on the edges of the panel. The chosen sample configuration provides 38 modules in each horizontal level and 18 modules in each vertical level for a total of 684 modules over 2 sq.m. The average density is 342 modules/sq.m.

Imagine a physical attack on the system. Since the modules are compact, the probability of a successful bullet shot is limited. This probability depends on the marksmanship of the attacker. Law enforcement statistics [6] say that in a gunfight hit rates are between 18% and 24%, aiming over 9.15 m away in 93% of the cases. Since the size of a typical silhouette is slightly less than 2 sq.m., the probability that a bullet hits the panel is slightly higher than 18-24%. With our concept, we estimate a module size equal to approximate 23.7 mm, having frontal projection area of approximate 5.62*10^-4 sq.m. All 684 modules together have an area of approximately 0.38 sq.m., which is less than one fifth of total panel area. Due to module compactness, it is almost impossible that a bullet destroys more than one module at once. Hence, the probability that a bullet destroys a single module is estimated as 4-5%. It means that to destroy at least 33% of the camera modules, the attacker needs to precisely land 226 bullets. With respect to hit rate, the attacker will need to fire more than 4500 bullets at the panel. And such tremendous number will keep 66% of the camera modules intact. Explosions or shotgun rounds, however, can potentially disable more than one camera module at once, and thus, destroy the panel more quickly.

4. Data processing

The system is intended to collect a very large amount of data. In order to make the system more usable, this data needs to be processed into information that can be displayed for decision by an operator. The collected data might be used for motion tracking, face or figure recognition, object position estimation, etc. In most simple configurations, a user could switch between the camera views, taking views of the same object from different angles. Depending on the density of camera modules over the area, the viewpoint shift can be potentially very smooth, as if the viewpoint slides around the object. With motion tracking software, the system can automatically switch between camera channels to keep an object in the center of frame. Some work has been done in motion tracking for compound video systems [3], [7], [8].

If image recognition is included, the system could detect separate objects and show multiple pictures. This could be used for classification and autonomous operation of the system. With a very high level of image recognition, it is also possible to form a stereo image. This would allow objects to be examined from multiple angles similar to a hologram.

The huge stream of video data that comes from a single panel is too high for a remote system to process, so the data processing should be optimized by distributing the load among computational centers. There are research groups that work on artificial neural networks for compound eye video processing, trying to imitate the nervous system of insects [3], [7], [9].

A neural network can also greatly help in the video data compression in preprocessing elements and decompressing on the remote security post. Cramer et. al. [10] has developed a real-time method which allows a compression ratio of over 1000 : 1 for full-color video sequences with the addition of the standard 4:1:1 spatial subsampling ratios in the
chrominance images. It should greatly help the data collection.

We envision the image preprocessing elements will be controlled from a remote security post and capable of filtering data from multiple module sources. In motion tracking mode the modules will transmit to the remote post only the data from the cameras that detect motion or will soon observe object moving from FOVs of neighboring cameras. If the panel is equipped with cameras that operate in various spectral intervals, only the data from cameras of the necessary spectral regions will be transmitted.

Instantaneous acquisition of data on the same object from different angles should greatly assist such tasks as face or figure recognition. These tasks require a significant amount of data to be used, and multiple cameras can speed up the data collection process. Small angles between the viewpoints will help to determine the actual 3D shape of monitored objects by measuring relative displacement of object points during the viewpoint displacement. These extracted 3D key points can then be applied on a 3D model [11].

The system is also capable of giving raw data for object position estimation. We propose a method for it: any object in the FOV of the whole system is observed by multiple cameras. Borders of their FOVs outline a domain in 3D space, which contains the observed object. The closer an object is to the center of the panel, the more accurate the estimation since more cameras track the target and the domain becomes smaller. It is possible to estimate object size by estimating positions of extreme points. This method might be helpful for crime scene investigations. Multiple viewpoints also perfectly suit the needs of stereo vision systems. These systems analyze disparity of images taken from location with known distance between them [12].

5. Discussion

A distributed array of camera modules is capable of providing covert video surveillance, with a wide FOV. Distributed modules acquire images of an object from a set of angles, which would vary based on the panel form factor. If combined with a special processing block, the system can potentially generate a stereo video stream.

The array has two parts – core and peripheral. Blockade or damage to any of the core modules does not affect FOV, since FOVs of neighbor modules overlap the FOV of the malfunctioning one. It affects only the image quality, but it should not lose information from malfunction of any single module. In order to disable the proposed system it would be necessary to block or damage a substantial part of the core modules to create a blind spot for the system. However, the peripheral modules are more vulnerable, because they have fewer neighbors. The situation might be improved by having higher concentration of modules in peripheral areas or by wrapping the device so that there is no visible seam or edge.

As a tradeoff, the concept requires more image processing than traditional systems. To increase performance, we plan to incorporate image processing elements in panels to preprocess the acquired data and reduce cabling from the panels to the security post. There are several ways to use these preprocessing elements. One option is to assign a small element per video module, which might make the system too expensive. Another way is to use more advanced elements with areas of panel assigned to them. This solution is a security risk, since malfunction of any preprocessing module creates a blind spot. Another solution is to connect every module to a set of distributed preprocessing elements. This seems to be the most safe and efficient solution.

This scheme is present in the most advanced of biological compound eyes. With it, insects can not only navigate in very dark conditions, but also get light that flickers with a frequency up to 300 Hz, and distinguish light polarization [13]. The set of camera modules allows creation of combined-vision panels. Such panels are intended to be composed of camera modules with different sensor matrices, so the same objects are observed in visible light, infrared, ultraviolet, and light with different polarizations. This can give very detailed information on observed objects.

6. Summary and conclusions

Insect vision systems provide a novel approach to the design and implementation of cameras for video surveillance. The system proposed in this paper combines the advantages of insect and compound vision with analysis of existing vision systems to eliminate the major weak points and provide more
robust methods for video data acquisition. Preliminary analysis of risk shows that the system design is less vulnerable to attack and has fewer blind spots. Further analysis and testing will be performed on this system as it develops.

7. References


