Surveillance Systems for Urban Crisis Management

Jörgen Ahlberg and Lena Klasén
Div. of Sensor Technology
Swedish Defence Research Agency (FOI), Linköping, SWEDEN
{jorahl, lena}@foi.se

Abstract

We present a concept for combing 3D models and multiple heterogeneous sensors into a surveillance system enabling superior situation awareness. The concept has many military as well as civilian applications.

A key issue is the use of a 3D environment model of the area to be surveyed, typically an urban area. In addition to the 3D model, the area of interest is monitored over time using multiple heterogeneous sensors, such as optical, acoustic, and/or seismic sensors. Data and analysis results from the sensors are visualized in the 3D model, thus putting them in a common reference frame and making their spatial and temporal relations obvious.

The result is highlighted by an example where data from different sensor systems is integrated in a 3D model of a Swedish urban area.

1 Introduction

Visual surveillance systems are increasingly common in our society today. You can hardly take a walk in the center of a modern city without being recorded by several surveillance cameras, even less so inside shops. The traditional surveillance systems in urban areas consist of a set of CCTV cameras acquiring images that are recorded and monitored at a surveillance central. In the surveillance central a set of TV screens show the images from one or more cameras per screen. The problem with this approach is that each camera records micro events, and these micro events are hard to relate to other micro events recorded by other cameras. Thus, it is difficult to put the micro events in a correct spatial and temporal context, and also to get an overview of the entire situation, i.e., situation awareness.

A similar problem arises when a situation, for example a riot or an act of crime or terrorism, has already occurred, and available video (or still image) material is to be analyzed. The available material can be a mix of CCTV recordings, police recordings, and recordings from bypassers (with the advent of cellular phones equipped with video cameras, this is very likely). Forensic analysis of such material typically starts with placing all the imagery in a common time frame in order to facilitate the reconstruction of the situation. This step involves time-consuming manual work, and when the work is done, the following analysis is still difficult.

The solution to both problems is to use a three-dimensional computer model of the area. In this virtual environment, the cameras from the real environment are represented by projectors, that project the camera views onto the 3D model. This approach has several advantages:

1. The context in which each camera is placed is visualized and becomes obvious.
2. The spatial relation between different cameras become obvious.
3. Imagery from several cameras can be studied simultaneously, and an overview of the entire area is easily acquired.

We propose to exploit this approach in combination with heterogeneous sensors to create a framework for surveillance of urban areas.

2 Surveillance using heterogeneous sensors

The use of a 3D model enables heterogeneous sensor data not only to be presented together with the visual data, but also to be analyzed together.

For example, assume that we have a sensor that can localize gunfire. The position of the shooter can then immediately be marked in the 3D model, which gives several intresting possibilities:

- If the shooter is within the field of view of a camera, he is pointed out by marking the location of the shot in the 3D model. The shooter can then be tracked forwards and backwards in time, searching for pictures suitable for identification and also warn others in the area.
- Regardless if the shooter is within the field of view of a camera or not, the shooter's field of view can be marked in the 3D-model. The marked area is a risk area that should be avoided and warned for.

We use an acoustic sensor network to detect and localize gunfire, and also to track people and vehicles.

Additionally, we use passage detection sensors for determining when people and/or vehicles enter the surveyed area and the other sensors should be activated. Examples of passage detection sensors are:

- Ground alarms that react on pressure, i.e., when someone walks on the sensor (that consequently should be placed slightly below the ground's surface). Several types of such sensors exist, we use a fibertoptic pressure-sensitive cable.
- Laser detectors that react when someone breaks an (invisible) laser beam.
- Geophones, i.e., seismic sensors that register vibrataions in the ground.

Several other types of passage detectors are commercially available.

3 Experiments

3.1 Test site

To test the concept a industrial block in the city of Norrköping was chosen as a surveillance area. The area was deemed suitable for our purposes since it (a) was quite complex, thus more intresting than, e.g., an ordinary street crossing, (b) had a central open area, (c) was private-owned, which facilitates legal stuff about video surveillance, (d) had stairs and balconies where we could easily mount our sensors, and (e) had a 90 meter high tower.

3.2 3D models

The entire central part of Norrköping was already scanned by an airborne laser in a previous project. The laser scan data from the surveillance area was processed by the algorithm described in [1] to extract buildings and create a 3D model of the chosen block.

A second model was created manually based on data from a ground-based laser scanner. The reason for having two models was to compare the automatic/airborne and the manual/ground-based method.

3.3 Sensors

A multitude of sensors were installed in the surveillance area. The stationary sensors included cameras, lasers, cables, microphones, and geophones, as listed Table 1. The positioning of the sensors is illustrated in Figure 1. Note that all three entrances to the surveillance area are guarded by passage detectors.

First, four professional video cameras (K1–4) were installed on the buildings around the block. Some of the cameras changed field of view once during the experiment, but stayed fixed otherwise. One thermal infrared camera (K5) was used in the same way, and on a high tower nearby, a combined visual and thermal camera (K6) was placed approimately 60 meters above ground.

Additionally, several mobile cameras (K10–16) were used as well, mostly for documenting the experiment. Those cameras were handheld

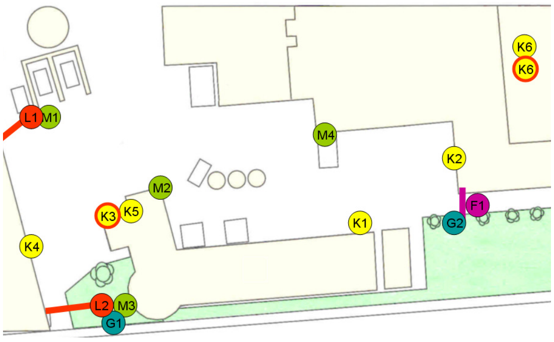

Figure 1. Positioning of the sensors.

digital still image cameras with the possiblity of recording low-end video and one video camera (K7) operated by one of the policemen participating in the experiment.

Around the main action area an acoustic sensor network was installed. The acoustic sensors (M1–4) each consisted of two or three microphones, each sensor being able to give the direction to a sound source. The entire sensor network can thus be regarded as a sensor that can locate sound sources.

Passage detection sensors of various types were placed at the entrances to the test site (F1, L1–2, G1–2).

3.4 Scenario

Two different scenarios were planned; one military and one civilian.

The military scenario involved two groups of own troops moving through the area, being assaulted by a small enemy force, and finally conquering the area. The actors were professionals (i.e., military, not actors).

Table 1: Summary of sensors

Sensor no.	Description
K1–4	Video camera
K5	Thermal video camera
K6	Visual and thermal video camera
K7–9	Video camera
K10–16	Handheld digital cameras
G1–2	Geophone
L1–2	Laser passage detector
F1	Pressure-sensitive fibre-optic cable
M1	Microphone disc
M2–4	Microphone duo

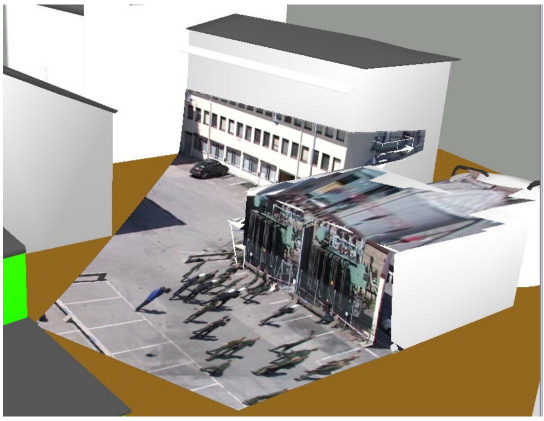

Figure 2. Projecting video onto a 3D model.

In the civilian scenario a group of shouting people (the green team) moved through the area, simulating, for example, a sports audience after winning a game or excited but non-violent demonstrators. At the main action area they were attacked by a smaller group (the red team), at first verbally, but followed-up by stone throwing and brawling. Finally, someone in the red team raised a firearm and fired several shots in the air, thus making the green team run away. Shortly afterwards, police (the blue team) entered the scene, capturing the the red team, and, as it turned out, also pacifying the by then not so non-violent green team.

The red and green teams were populated by military and Home Guard personell, providing superb and very affectionate acting. The blue team were professionals, i.e., police officers, and acted, as expected, professionally.

4 Results

Basically there were three kinds of processing; visual data, acoustic data, and passage detector data. The processing is not described in detail here, only results are reported.

- **Optical:** Software for projecting video on a 3D model was developed and used two visualize the recorded video from the two scenarios. An example is given in Figure 2.
- **Acoustics:** In the military scenario, the acoustic sensors were of less use due to flooding during most of the scenario – there were several automatic weapons fired close to the microphones. However, in the civil-

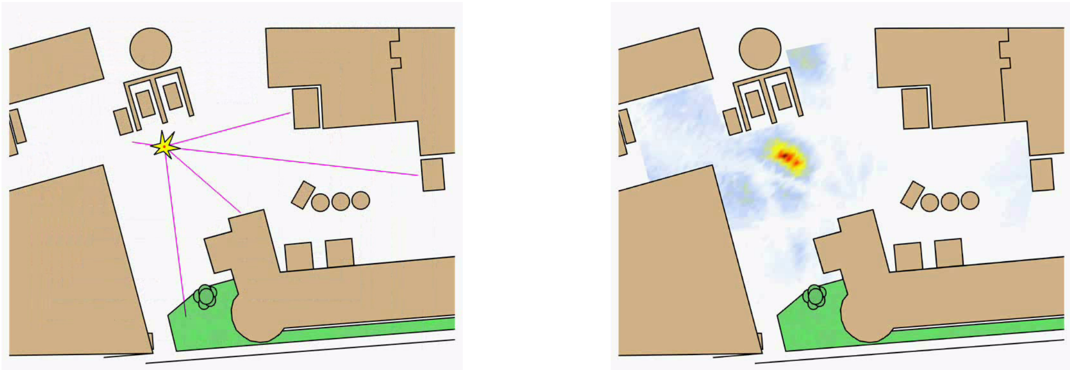

Figure 3. Localization using acoustic sensors. Left: Gunfire. Right: A shouting mob.

ian scenario, the acoustic sensors were able to track the group of shouting people through the surveillance area, and also to localize the fired shots with a precision of a few decimeters, as illustrated in Figure 3.

- **Passage detection:** In both scenarios, the passage detectors easily picked up the entry of the different groups.

5 Conclusion

We have presented a concept for surveillance of urban areas combining 3D models and multiple heterogeneous sensors to achive superior sitation awareness, and successfully applied the system to military and civilian scenarios. We have no quantitative results to prove our point, however, the system clearly demonstrates the advantages by combining the different sensors and also proves the feasibility.

There are several applications of this system, both for civilian and military use. The most prominent are tactical decision support, documentation, forensics, and simulation/training. Future enhancements will be able to provide warnings for abnormal or criminal behaviour, integrity-preserving surveillance and automized sensor location.

6 Acknowledgement

The project described in this paper involved a large number of people. The authors would like to thank Magnus Elmqvist for mastering the field trial; Håkan Larsson, Göran Carlsson, and Roland Lindell for handling the optical sensors; Hans Habberstad, Fredrik Kullander, Claes Vahlberg, and David Better for all the other sensors; Per Carleberg for software development; our 3D artist Joakim Johansson; and, finally, Tomas Chevalier and Pierre Andersson for being involved in most of the above.

Additionally, the authors want to express strong gratitude to Sydkraft for letting us use their site in Norrköping; the Army Ground Combat School at Kvarn, the Home Guard, and the police forces in Norrköping for acting and assistance; and FLIR Systems for lending us their Sentry thermal/visual surveillance pod.

References

[1] S. Ahlberg, M. Elmqvist, Å. Persson and U. Söderman, “Three-dimensional environment models from airborne laser radar data,” *Proc. SPIE Vol. 5412, Conf. Laser Radar Technology and Applications VII*, 2004.