A&QT-R 2002 (THETA 13)
International Conference on Automation, Quality and Testing, Robotics
May 23 – 25, 2002, Cluj-Napoca, Romania

Cost-Effective Method for Motion Detection in Building Security Systems

Authors:
Mihai-Dan Steriu
Cristian Marinescu
Andrei Stanescu
Mihai Maiorescu

"Politehnica" University Bucharest
steriu@golana.pub.ro

Abstract:

This paper describes motion analyses and detection, as seen by a static camera. The algorithm isn't proposed to identify the objects nor their trajectory, only to detect if there are one or more moving objects. The system is intended for surveillance systems that can also store images in JPEG format. That is why the algorithm uses the DC coefficients, obtained in JPEG compression.

Keywords: JPEG, motion detection, building communication

1. Introduction

This paper is a step in achieving the goal of creating a building communication system that allows complete control of security, monitors activity and allows voice or data communication between terminals.

In the case of a fire disaster, the building is evacuated, and it's possible for an elevator to cease it functioning, or an area, where a Remote Sensing Device (RSD), is placed, to become inaccessible. Following this event once the fireman crew arrives, they will look at the display in the Control Center and if people are caught in the elevator they will hurry to save them. Also, in a malfunctioning elevator, the intervention crew can look at the same display to see if somebody is caught in.

Another possibility of operation is surveillance. The security man can periodically look at the pictures taken in the points where the RSDs are placed. Maybe he will see a burglary in progress, or he will grant access to someone. A voice or data connection can be anytime established, to listen to what is happening in the places where RSDs are located, to give information or guidance to the people near the RSD.

In some applications it's required that a RSD shoud figure by itself if something is moving, besides acquiring and sending pictures. In this case, it informs the Control Center who rings the alarm.

It's possible to add additional features, such as smoke detectors, temperature meters, if this seems appropriate.

2. System description

The system consists in one Control Center and several RSDs, placed in the most sensitive parts of a building (doors, elevators, etc.), depending on customers demands.

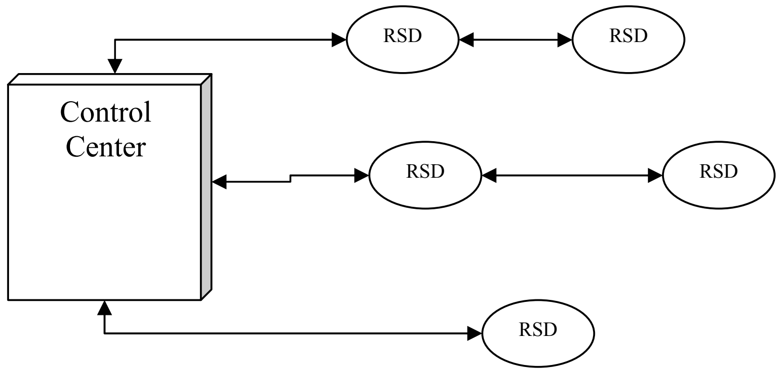

Figure 1

The Control Center is requested to control the correct functioning of the RSD, to gather information, to send commands. Here it is placed a display that reports information about RSDs and shows pictures taken from RSDs. A small keyboard is also present for commands issuing. Events and RSDs' status are indicated by leds.

The RSD is requested to monitor the activity and depending on the mode of operation, to report to the Control Center. Its main components are a video camera and several technical inputs/outputs (smoke detectors, thermometers, door open/close). Motion detection is achieved through image processing in RSDs.

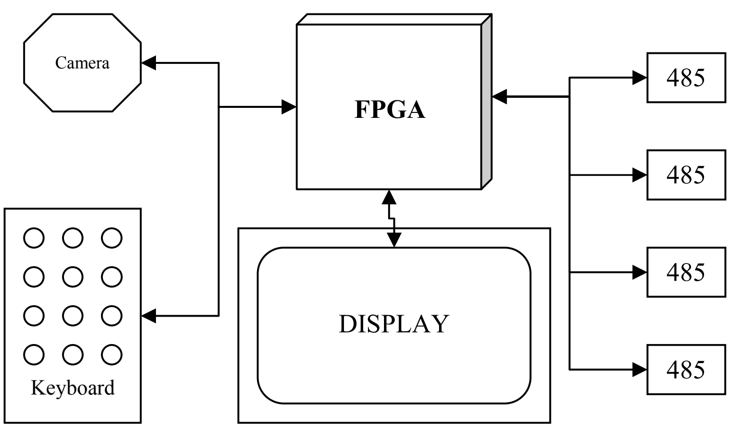

Figure 2

2.1. Control Center

The main physical components of a Control Center are a small LCD, a small keyboard and the FPGA device. The optional features are an audio system and a RS232 line interface (to be used with a modem or with other communication devices).

A red led shows that a RSD is malfunctioning or the communication line is broken, a green led indicates a normal functioning, no led indicates that a RSD is not present (installed). Additional components are a SRAM module, and 4 RS485 line interfaces.

The software of a Control Center (embedded components) is formed of: a display controller, a JPEG decoder, several serial transceivers, an A600 (RISC processor), a memory controller.

The RISC processor is used for coordinating communication (implementation of protocol layer 2) and overseeing the system status. It is also responsible with the commands interpretation (configuration console).

The hardware JPEG decompression (in FPGA), instead of a software approach in A600, is used to increase reliability and to minimize the decompression time.

2.2.Remote Sensing Device

An advanced RSD is constituted by a camera, a FPGA and a SRAM. It is able to take snapshots, and create a JPEG file.

The motion detection is based on the statistical calculation of DC coefficients (from JPEG) for two successive greyscale pictures.

If the picture is a matrix $P_{(m\times n)}$, then the following formula calculate the Discrete Cosine Transform (DCT) for every (i, j, u, v) with $i \in \left[0, \frac{m}{8}\right]$ and $j \in \left[0, \frac{n}{8}\right]$, and u, v from 0 to 7.

$$S_{i,j,u,v} = \frac{1}{4} \cdot C_u \cdot C_v \cdot \sum_{x=0}^{7} \sum_{y=0}^{7} P_{x+i\cdot 8, y+j\cdot 8} \cdot \cos\left[\frac{(2x+1)u\pi}{16}\right] \cdot \cos\left[\frac{(2y+1)v\pi}{16}\right] \qquad (1)$$

For u and v equal to 0, $S_{i,j,0,0}$ are the DC coefficients.

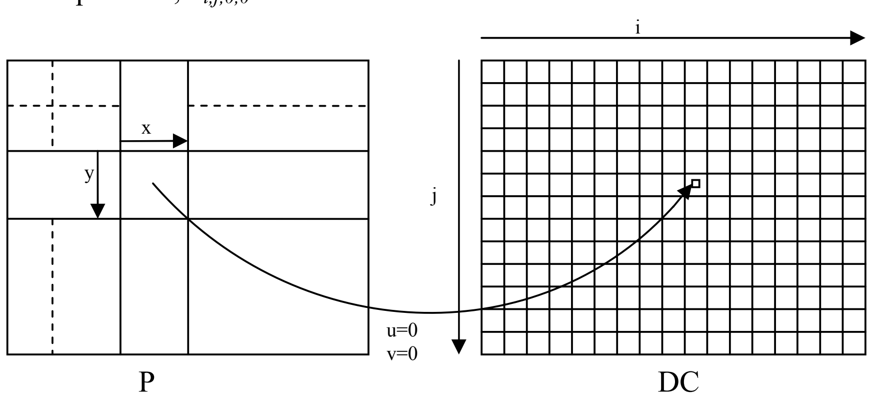

Figure 3

$$DC_{i,j} = S_{i,j,0,0} = \frac{1}{8} \sum_{x=0}^{7} \sum_{y=0}^{7} P_{x+i\cdot 8, y+j\cdot 8} \qquad (2)$$

Then, the mean luminance is computed, and later, it is subtracted from its DC matrix. Those calculations are performed for two successive pictures.

$$\overline{DC}1,2 = \frac{\sum_{k=0}^{\frac{m}{8}-1}\sum_{l=0}^{\frac{n}{8}-1} DC1,2_{k,l}}{\frac{m}{8}\cdot\frac{n}{8}} \qquad (3)$$

The difference between the mean luminances indicates the nature of the difference between the two pictures. A value greater than a few hundreds means that there is a major difference between pictures and a dispersed luminance variation, like light turning on/off.

Further, we need to see more about the nature of the difference between the pictures. We make the notation:

$$\overline{\overline{DC1}} = DC1 - \overline{DC1} \qquad (4)$$

$$\overline{\overline{DC2}} = DC2 - \overline{DC2} \qquad (5)$$

For 2 pictures we calculate the difference, by subtracting the two matrixes obtained:

$$D = \overline{\overline{DC1}} - \overline{\overline{DC2}} \qquad (6)$$

Then, we calculate the sum of positive elements, and the sum of negative ones, with a threshold value, so as to eliminate the noise:

$$negative = \sum_{k=0}^{\frac{m}{8}-1}\sum_{l=0}^{\frac{n}{8}-1} \Phi\left(-threshold - D_{k,l}\right)\cdot D_{k,l} \qquad (7)$$

$$positive = \sum_{k=0}^{\frac{m}{8}-1}\sum_{l=0}^{\frac{n}{8}-1} \Phi\left(D_{k,l} - threshold\right)\cdot D_{k,l} \qquad (8)$$

$\Phi(x)$ stands for the Heavyside function. A threshold value of 20 is appropriate.

When an object is moving through the pictures, the difference in terms of DC coefficients is that in the areas of current and past position, the coefficients' variations are in opposite directions. In other words, the negative footprint of a moving object is found in his last position. One moving object and its negative footprint are represented below:

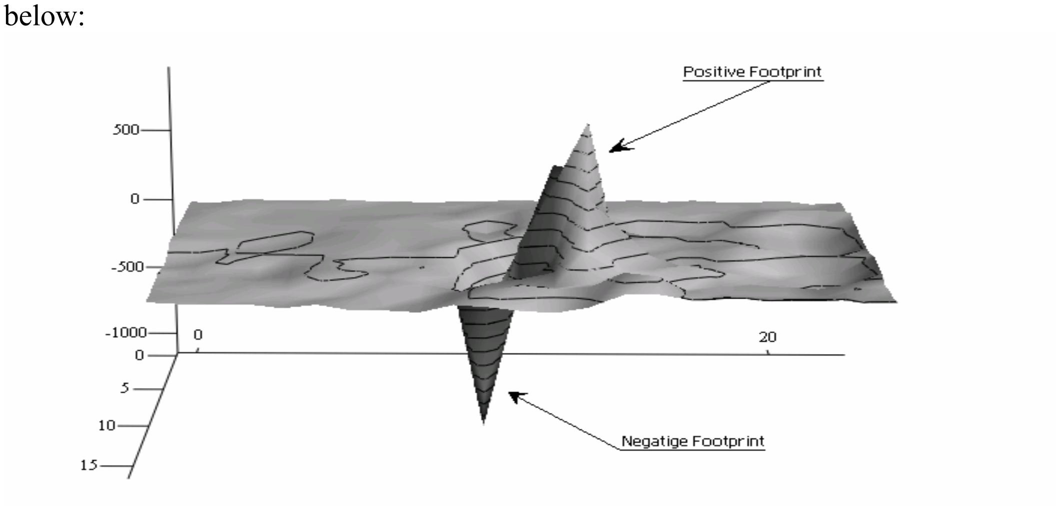

Figure 4

So, if there is an object moving, the variables *negative* and *positive* are close to each other. To that purpose we calculate the ratio:

$$\frac{\max(negative, positive)}{\min(negative, positive)} \qquad (9)$$

If this rapport is close to 1, there is movement in the picture. If it is bigger than a threshold value, the difference between the pictures is not coming from a moving object. For identical pictures, *positive* and *negative* are around 0.

Also, the movement parameter is computed with the formula:

$$move = \frac{|positive| + |negative|}{\sum_{k=0}^{\frac{m}{8}-1}\sum_{l=0}^{\frac{n}{8}-1}\Phi(-threshold - DC_{k,l}) + \sum_{k=0}^{\frac{m}{8}-1}\sum_{l=0}^{\frac{n}{8}-1}\Phi(DC_{k,l} - threshold)} \qquad (10)$$

For identical pictures the movement parameter has a value around 0; for pictures with movement, its value is greater.

The table below illustrates the calculations for a few pairs of pictures with the movement parameter and *positive/negative* values already computed:

Picture	**Move**	**pos.**	**neg.**	**pn**	**s1-s2**
Figure 5 and Figure 6	77	8305	-9249	1.14	81.8
Figure 6 and Figure 7	96.7	16126	16283	1.01	111

Below there are illustrated some surveillance pictures, and the D matrixes for the pairs (Figure 5, Figure 6) and (Figure 6, Figure 7).

Figure 5

Figure 7

Figure 6

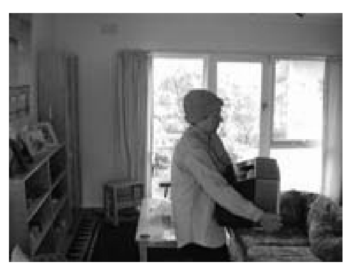

Figure 8

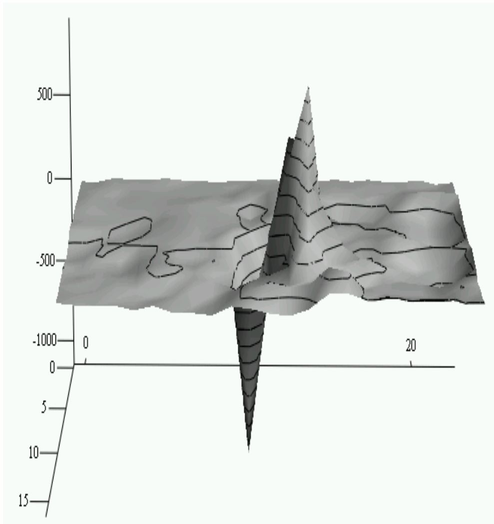

Figure 9

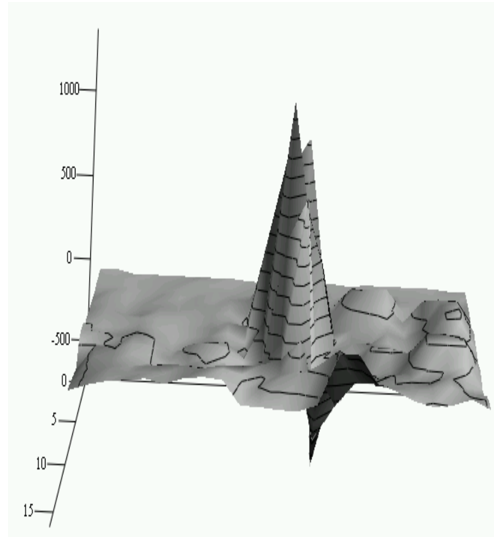

Figure 10

3. Communications

The standard version is based on a peer-to-peer connection between terminals in tree architecture, using a RS485 physical layer. The Control Center features up to 4 physical buses and up to 15 RSDs per bus (limited software). The tree can't have a branch longer than 4000 feet. The baud rate is between 115Kb/s and 921Kb/s.

An option to the 485 physical layer is Ethernet (10(coaxial or TP) or 100 Mb/s (TP)), but the cost is higher.

Basically, the communication is master slave oriented. The Control Center is acquiring pictures from RSDs, or it's informed about their status. Anyway, the communication between RSDs is also supported. It is also possible to use in the system different type of Remote Devices, identified by an ID number.

4. Conclusions

The use of JPEG is an elegant technique because it allows the compression of images, and a relative easy movement detection. The form recognition and movement direction detection are possible developments in the future.

5. References:

1) ITU-T - "Information Technology – Digital Compression and Coding Of Continuous-Tone Still Images" (T81) [1992]

2) http://www.dindima.com/swise/html/__camera_images.html

3) Ross Cutler and Larry Davis. "Robust Real-Time Periodic Motion Detection, Analysis, and Applications." [2000];