4th WSEAS International Conference on ELECTRONICS, CONTROL and SIGNAL PROCESSING, Miami, Florida, USA, 17-19 November, 2005 (pp.205-210)

A wireless sensor network for visual detection and classification of intrusions

ANDRZEJ SLUZEK[1,3], PALANIAPPAN ANNAMALAI[2], MD SAIFUL ISLAM[1]
[1]School of Computer Engineering, [2]IntelliSys Centre
Nanyang Technological University, Blk N4 Nanyang Avenue
SINGAPORE
[3]SWPS, ul.Chodakowska 19/31, Warszawa
POLAND

Abstract: - The paper reports works on a wireless sensor network supporting a human operator monitoring a wide area of an unstructured environment. The main objective is to develop mechanisms allowing intelligent preliminary analysis of data (captured by a large number of sensing devices, including cameras) so that the human intervention is needed only for cases identified as "potentially dangerous intrusions". The developed feasibility-study platform contains two types of nodes: relatively simple warning nodes with a variety of sensors detecting a potential presence of intruders (proximity, vibration, magnetic activity, noise, etc.) and more sophisticated nodes equipped with cameras and capable to perform complex image analysis tasks so that the intrusion can be visually verified and (possibly) classified before involving a human operator. The nodes can wirelessly communicate using a standard protocol. Several mechanisms of visual detection have been proposed and implemented. Thus, the actual details of a network for a given application can be optimized, taking into account the current conditions, hardware constraints and user-defined requirements. Generally, minimization of the wirelessly transmitted data is of high priority. The paper briefly overviews the developed platform and discusses methodologies of the algorithms used in the system. Some implementation details are not disclosed.

Key-Words: - **visual surveillance, sensor network, FPGA, intrusion detection and classification**

1 Introduction

Although visual surveillance in considered the most effective method of monitoring complex environments, systems that could perform such a task fully autonomously and reliably are still very difficult to build. In less ambitious (but more realistic) scenario, a human operator can be supported by a visual surveillance system that can automatically handle typical situations. The human intervention is needed only in special situation (or situations involving decisions which are reserved for humans).

In this paper we report development of wireless sensor network that is based on the abovementioned concept. The network is supposed to monitor a relatively large area (and to detect intruders in this area) using sensor nodes (usually deployed randomly or partially randomly) of two types. The first-level nodes (which would be more numerous) contain relatively simple sensors detecting non-visual signatures of intruders. When a potential presence of an intruder is detected, the second-level nodes are wirelessly activated. The second-level nodes are equipped with cameras that (upon activation of the node) capture images of the area where the intrusion is suspected. The images are subsequently processed by the digital hardware (FPGA) of the node in order to classify and possibly identify the intruder. If the node is unable to categorize the intrusion, the intruder's image can be wirelessly transmitted to a higher decision level, possibly involving a human operator.

In Section 2 of the paper, we briefly present the developed hardware platform. Section 3 discusses image processing algorithms used in the second-level nodes to classify intruders. More advanced techniques of intrusion identification (performed by the host computer using data transmitted by the second-level nodes) are presented in Section 4. Section 5 summarizes the paper and highlights directions for the future works.

2 Description of the Network Components

The developed network consists of an unspecified number of nodes – no predefined structure of the network is envisaged. The first-level nodes (Fig. 1)

incorporate a wireless communication chip, a low-cost microcontroller (performing sensor data acquisition, generatingf wirelessly-transmitted messages and interfacing the wireless chip) and basic non-vision sensors. Battery power supply is provided. Currently, only infrared proximity sensors and vibrations sensors are used, but other additional options are planned (and the corresponding sensors are available).

Fig. 1. The first-level node of the network.

The sensors of first-level nodes monitor conditions in the protected area and act as preliminary detectors of possible intrusions. A possible intrusion is defined as a particular coincidence of sensor readings. Various definitions of possible intrusions can be used, either for different applications or within the same application. Whenever a possible intrusion is sensed, the node wirelessly transmits a message containing the node identifier and (if several definitions of possible intrusions are simultaneously used) the type of sensed intrusion. To provide a higher level of reliability, the message may be repeated several times. A standard communication protocol for low-power wireless networks is used.

The second-level nodes area built around an FPGA module that can control up to four cameras. Additionally, the nodes are equipped with the same wireless communication components as the first-level nodes. Typically, one second-level node is associated with several first-level nodes, but the network structure is not permanent. As an illustration, a typical fragment of the network is shown in Fig. 2.

Low energy consumption in the network nodes is a crucial issue since very long operational lifetime is expected. In the first-level nodes, this is achieved by using high performance batteries and low-power components. Thus, the nodes can be permanently active during their whole operational lifetime.

In the second-level nodes, power consumption by both FPGA's and cameras is relatively high. Therefore, only the kernel of a second-level node would be permanently active, while the other parts (e.g. the cameras) are activated only when a warning message from the first level is received. Two option are possible: (1) a second-level node can be activated by any first-level node within the wireless range or (2) a second-level level node can be activated by selected first-level nodes (i.e. the warning messages containing identifiers of other nodes are ignored).

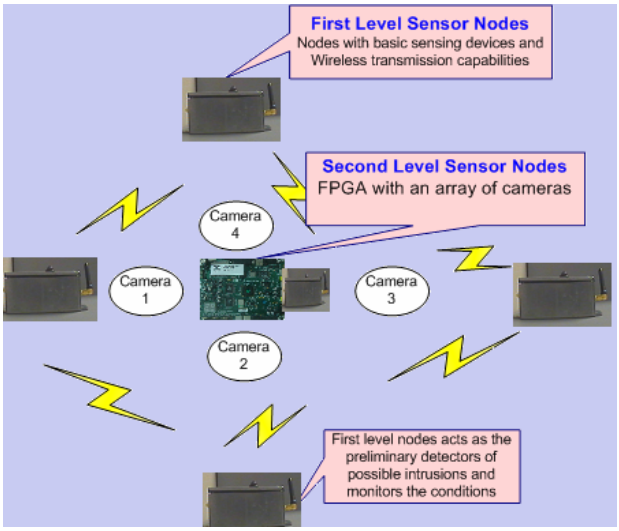

Fig. 2. Fragment of the network.

Upon activation, a second-level node camera captures a short sequence of images (typically two or three) that are subsequently processed using dedicated algorithms implemented in the node's FPGA. In general, the purpose of image processing is to extract the possible intruders from a captured image and to classify/identify them. After the task is completed, selected fragments of camera-captured images and/or other results produced by the algorithms may be wirelessly transmitted to higher-level components of a decision system (possibly including a human operator). Additionally, the second-level nodes can be periodically activated in order to update the background image (see Section 3).

In general, the peak energy consumption in a second-level node may be high, but the average power requirements could be low enough to provide long operational lifetime. In the prototype platform, power is supplied by a DC adaptor.

The prototype second-level nodes are built around a commercially available FPGA development board (see Fig. 3). However, only selected resources available in the board are used in the prototype (as shown in Fig. 3). In the final product, a less powerful FPGA will be possibly used.

Details of image processing algorithms implemented in the second-level nodes are discussed in Section 3.

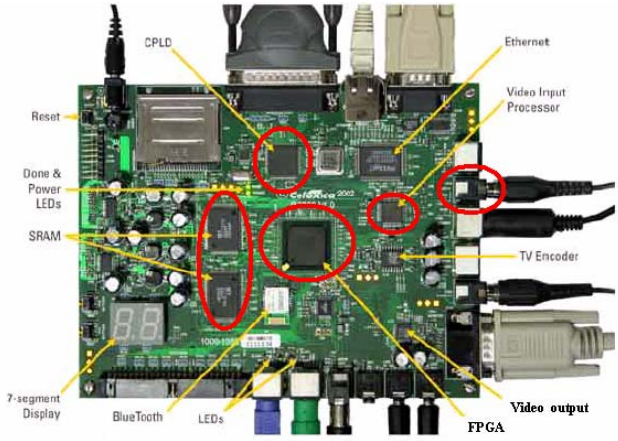

Fig. 3. Components of the FPGA development board used in the second-level node.

3 Image Processing in Intrusion Detection

3.1 Extraction of Intruder's Silhouettes

When a camera attached to a second level node captures an image of the suspected intrusion area, the first image processing operation is extraction of fragments not present in the background image currently memorized in the node. Although such an operation can be (hypothetically) implemented as a simple image subtraction, a more sophisticated method is needed for scenarios of realistic complexity. In general, the scene background may be affected by both illumination variations and minor configuration changes (e.g. swaying trees, grass, vibrations of the network nodes, etc.).

Fig. 4. An exemplary background image captured by the camera.

Fig. 5. An exemplary scene containing an intruder.

Fig. 6. Intruder's silhouettes extracted by the proposed algorithms (both versions are shown).

The first problem can be solved by occasional background updates (see the later part of this section) but for the second one, a more sophisticated approach has been proposed. In the developed system, we apply a combination of subtractions in RGB space (with a threshold defining the acceptable differences) followed by a sequence of selected

morphological operations. The purpose of the morphological operations is to filter out differences caused by minor configuration changes.

The extracted silhouettes of intruders are available in two different forms: either as a binary blob of the intruder's shape or as a full image within the extracted silhouette. Examples of the original scene, the scene with the intruder, and two variants of the intruder's silhouette are shown in Figs 4, 5 and 6.

The intruder's silhouettes can be subsequently used for either visual identification or for further analysis and classification of the intrusion.

In the actual implementation of the method, two (or possibly more) images of the intruded area are captured, and the silhouettes are extracted from both images.

The presented system has been designed under the general assumption that the intruders can be associated with other physical phenomena, so that relatively simple sensors (proximity, vibration, magnetic, etc.) can be used as warning devices indicating a potential presence of intruders. Thus, if there is a change in the camera-captured images without the presence of the corresponding sensor warnings, it should be considered a background change rather than intrusion.

Therefore, the algorithm used for intruder extraction can also be used for periodical background update. It is particularly useful if the background change is due to rapid illumination changes (additional shadows, etc.). For slow changes of the background a simple correlation method has been implemented based on the following formula:

$$r = \frac{\sum_m \sum_n (A_{mn} - \overline{A})(B_{mn} - \overline{B})}{\sqrt{(\sum_m \sum_n (A_m - \overline{A}))^2 (\sum_m \sum_n (B_m - \overline{B}))^2}} \quad (3.1)$$

Where ***A*** and ***B*** are the current background and a new image, respectively. The background is updated when the value of ***r*** reaches the predefined threshold.

3.2 Further Analysis of Silhouettes

Detected intruders can be further characterized by using selected features of their binary silhouettes. The proposed features are moments of low order, from which several useful descriptors of the intruders can be derived. In particular, the type of intruder's mobility can be broadly categorized by comparing the low order moments of the silhouettes extracted from a pair of subsequently captured images.

The example in Fig. 7 shows a pair of (overlapping) silhouettes of a human. A simple comparison of the low order moments calculated from both silhouettes, would indicate that a vertical intruder of irregular shape is moving toward the node and turning to the left. The speed of motion can be estimated based on the size changes and displacements of the gravity centre. Such a characteristic could be a sufficient evidence to recognize the intruder as a human.

In case of some man-made objects, the silhouettes extracted from the images would have more consistent shapes so that a broad classification of such objects can be done within the second-level node by using moment invariants (eg. [2], [3]).

Generally, intrusions that can be satisfactorily identified at the second-level node are not sent for visual verification by a human operator. Unknown cases, however, have to be inspected. The images of such intruders would be wirelessly transmitted to the next level (which would usually incorporate a human operator). In order to save the bandwidth, actually only the full-image silhouettes (or their rectangular outlines) are transmitted.

Another prospective application of the full-image silhouettes is discussed in the next section.

Fig. 7. Intruder's silhouettes extracted from two consecutive images (the green area is shared by both silhouettes).

4…Advanced Method of Intruder Identification

The shape-based methods discussed in Section 4

would fail in most cases where intruders are only partially visible. Therefore, selected methods that have been originally intended for vision-based robotic navigation are being adopted for the developed system.

The proposed approach is based on detection and matching interest points in relative scale (e.g. [4], [5]). Interest points (sometimes referred to as corner points) are easily perceivable small areas where the *corner response* (based on the matrix of 2D partial derivatives of the image intensities - see [X]) reaches its local maximum. Examples of interest points automatically found in two images of real objects are shown in Fig. 8.

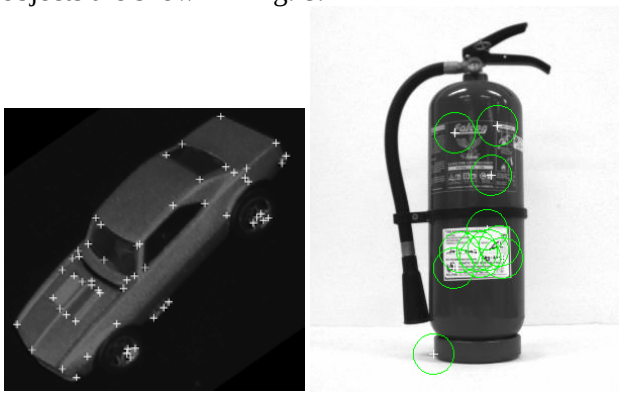

Fig. 8. Two examples of object images with automatically extracted interest points.

If an object of interest is represented by a dataset of its interest points (including geometric relations between the points) an image of such an object can be hypothetically identified matching image-extracted interest points to the dataset.

The major practical difficulties, i.e. sensitivity to illumination variation, scale of the objects and perspective distortions of 3D objects, have been successfully solved in the algorithm presented in [7] and [8]. Exemplary results of matching interest point of a model object and a camera-captured image of a partially visible object (under different illumination conditions and in a different scale) are presented in Fig. 9.

With a sufficient number of interest points matched, the presence of an object of interest can be positively verified in the camera-captured image.

Therefore, the identification of intruders by using their silhouettes can be done using the same principle. Interest points are detected within the intruder's silhouettes (full-image) and their parameters would be wirelessly transmitted to a higher level system that contains a database of know objects. If the intruder can be verified based on the interest points, the human assistance is not required.

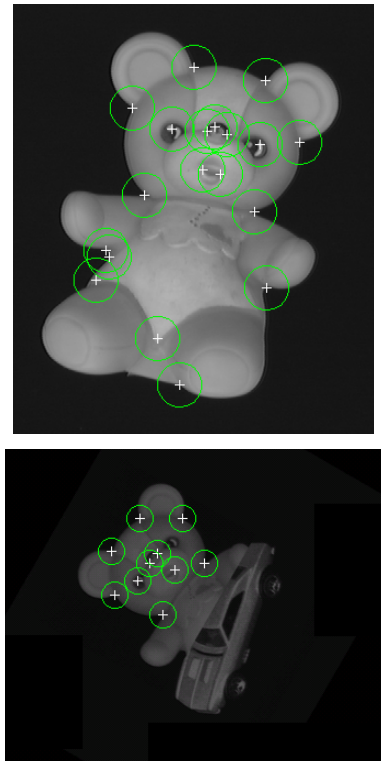

Fig. 9. Interest points in a model object and their correspondences is geometrically and photometrically distorted image of the object.

4 Conclusions

The system reported in this paper has been implemented using commercially available development boards (RC200 FPGA development boards and Ember microcontroller boards). The final prototype model, however, will have a dedicated architecture designed based on conducted tests and further analysis and researches. Similar concepts of an FPGA-based wireless sensor network are discussed in details in [1].

Although the paper highlights mostly the image processing aspects of the developed systems, other components also play important roles in the overall functionality of the network. It can be briefly mentioned that mechanisms for reliable wireless transmission have been implemented in the system (including on-chip implementation of AES-128 encryption algorithm). Additional tests and experiments are needed, however, to identify the most suitable wireless protocols. Development of a proprietary protocol is seriously taken as a feasible alternative.

In the future, further development is envisaged that may lead to a commercially available product.

4 Acknowledgment

The authors gratefully acknowledge the support and funding from IntelliSys Research Centre, a partnership between Nanyang Technological University and Singapore Technologies Engineering.

References:

[1] Bellis, S.J. et al. *Development of field programmable modular wireless sensor network nodes for ambient systems*. Computer Communications, to appear, 2005.

[2] Hu M.K. *Visual pattern recognition by moment invariants*. IRE Trans.Inf.Theory vol.8, pp 179-187, 1962.

[3] Maitra, S. *Moment invariants*, Proc. of IEEE, vol. 67(4), pp. 697–699, 1979.

[4] Mikolajczyk, K. and Schmid, C. *Scale & affine invariant interest point detectors*, International Journal of Computer Vision, vol. 60(1), pp. 63-86, 2004.

[5] Schmid, C., Mohr, R. and Bauckhage, C. *Evaluation of interest point detectors*, International Journal of Computer Vision, vol. 37(2), pp. 151-172, 2000.

[6] Harris, C. and Stephens, M. *A combined corner and edge detector*, 4th Alvey Vision Conference, Manchester, pp 147-151, 1988.

[7] Islam, M.S. et al. *Towards invariant interest point detection in an object*, Conf. WSCG'2005, Plzen, pp. 101-104, 2005.

[8] Islam, M.S., Sluzek, A. and Zhu, L. *Representing and Matching the Local Shape of an Object*, Conf. MIRAGE 2005, Rocquencourt, pp 9-16, 2005.