SmartMonitor: An Approach to Simple, Intelligent and Affordable Visual Surveillance System

Dariusz Frejlichowski[1], Paweł Forczmański[1], Adam Nowosielski[1], Katarzyna Gościewska[2], and Radosław Hofman[2]

[1] West Pomeranian University of Technology, Szczecin
Faculty of Computer Science and Information Technology
Żołnierska 52, 71-210, Szczecin, Poland
{dfrejlichowski,pforczmanski,anowosielski}@wi.zut.edu.pl
[2] SmartMonitor, sp. z o.o.
Niemierzynska 17a, 71-440, Szczecin, Poland
{katarzyna.gosciewska,radekh}@smartmonitor.pl

Abstract. The paper provides fundamental information about the SmartMonitor – an innovative surveillance system based on video content analysis. We present a short introduction to the characteristics of the developed system and a brief review of methods commonly applied in surveillance systems nowadays. The main goal of the paper is to describe planned basic system parameters as well as to explain the reason for creating it. SmartMonitor is being currently developed but some experiments have already been performed and their results are provided as well.

1 Introduction

Nowadays the surveillance systems are becoming more and more popular. Their main functionality is associated with ensuring the safety of monitored people, buildings and areas by detecting suspicious behaviours and situations. The most commonly used monitoring systems require human support that can involve observation of multiple monitors and a need of quick reaction. Moreover, intelligent monitoring systems that utilize video content analysis algorithms are usually applied for monitoring wide areas or public places, such as airports or shops, where the number of moving objects is large. Such systems as well as required infrastructure are very specific, targeted and expensive. For those reasons they are unaffordable for individual clients and home use. Nevertheless, there is still a large number of people who want to ensure the safety of themselves and their assets in small areas. Additionally, a high demand on robust alarm systems with low percentage of false alarms and on systems requiring commonly available hardware exists. According to the lack of solutions designated for personal use, SmartMonitor is being developed to satisfy the mentioned needs.

The SmartMonitor is assumed to be an inexpensive security system designed to protect properties as well as to ensure personal safety. It will utilize common

L. Bolc et al. (Eds.): ICCVG 2012, LNCS 7594, pp. 726–734, 2012.

hardware that is available and affordable for individual users – personal computer connected with digital camera(s) (by USB, Ethernet or Wi-Fi), what will offer good image quality. SmartMonitor will analyse captured video stream in four predefined scenarios and in opposition to traditional monitoring systems it is intended to work without users control. The initial system calibration will be the only operation requiring human interaction. The system would react to every learned situation in pre-specified way. The most important advantage will come from system customizability. Each client will be able to decide about the system reaction, the region that is monitored and particular situations or objects (e.g. people, cars, pets, etc.) that must be detected and tracked. Therefore, it will allow for creating individual safety rules that meet ones actual requirements.

The SmartMonitor is related to the category of small-size surveillance systems (Closed Circuit Television, CCTV) that aim at increasing the safety and systems that analyse image content, usually with the use of background modelling methods (Video Content Analysis, VCA). The algorithms that will be employed belong to the state-of-art category, yet they are adopted and simplified to the characteristics of the system. Additionally, SmartMonitor could be integrated with available infrastructure in public areas and better PCs would manage to run more complex procedures. A simplified scheme of planned system modules is presented in Fig. 1.

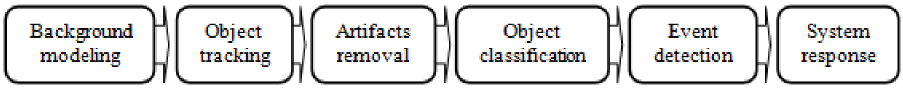

Fig. 1. Simplified scheme of basic SmartMonitor system modules

The rest of the paper is organized as follows. The second section contains a brief review of visual surveillance systems as well as some methods and algorithms that are usually utilized. The third section presents main goals, functions and properties of the developed system. The fourth one provides sample experimental results and the last one concludes the paper.

2 Brief Review of Intelligent Monitoring Systems and Algorithms

Three main modules embedded in visual surveillance system are adaptive background model, object extraction and tracking. Each of them requires specific algorithms that can be utilized in real-time system in respect to computation time. Many current systems focus only on tracking people without the analysis of their behaviour. If the system is assumed to be an intelligent one, it should utilize more sophisticated solutions and during the tracking of particular objects the methods for removing false objects and classification have to be introduced as well. Below we present a set of visual monitoring systems characteristics and algorithms.

Intelligent monitoring system presented in [1] analyses human behaviour on the basis of the location, motion trajectory and velocity, and with the use of Hidden Markov Models (HMM). In order to recognize the type of behaviour, the probability of the similarity between a scene model and a scene actually processed is estimated. Unfortunately, learning process requires the preparation of large database and the involvement of qualified employee ([1]) what is unnecessary in SmartMonitor due to the assumed feature based methods and simple calibration.

In [2] the problem of object tracking was described. Discussed system is based on the generalization of MAPF (Memetic Algorithm Particle Filter) for tracking variable and multiple number of objects. It utilizes simple scenarios only for selected areas, such as enter/exit or restricted ones. That can be found as an important advantage in relation to less computation time ([2]) for SmartMonitor as well.

Another paper ([3]) is focused on problems of automatic monitoring systems with object classification, especially those analysing real video scenes. An approach for classifying objects as well as several features that distinct various classes were proposed. Authors assumed that the background is static and does not change during the video sequence. That is found as an important simplification but also a limitation because of background variability in real videos. Therefore, this approach cannot be introduced to SmartMonitor – the planned system scenarios are associated with very unpredictable scenes, e.g. due to changing weather conditions. Nevertheless, very valuable algorithms were described in [3]. Basic object features were represented in anumerical form as descriptors (e.g. geometric moments). That allowed for the utilization of threshold values during classification ([3]).

Moving object detection is usually based on background subtraction that utilizes algorithms for estimating background models. In reference to the literature, there are three categories of such methods. The first one includes models based on static background. In order to extract foreground image, current video frame is compared with previously prepared background image. That approach is simple but has low efficiency. In the second category the background image is obtained by averaging agroup of initial frames that is compared with currently processed average group of frames. This model is more adaptive but still do not adjust to the changes in lighting. Therefore, the background image can be obtained e.g. every hour or more frequently what gives at least partial adaptability ([4]). The third category includes adaptive background models such as Gaussian Mixture Models (GMM), found in the literature as the most efficient (e.g. [5–8]). In [5] a method for modelling each pixel value as amixture of Gaussians and on-line approximation for model adaptation is discussed. The background image is updated with every processed frame basing on Gaussian distributions – the most probable pixel distribution is considered as abackground. This model adapts to slow changes in lighting conditions and repetitive variations([5]). Despite all advantages, there are also some drawbacks. The first problem is related to the number of Gaussian distributions that represent every pixel. It was proposed in

[6] that it should be equalled to 3 to 5 models or can be modified adaptively. Additionally, false objects (artefacts) such as shadows, reflections or false detections can significantly influence the foreground region. A shadow is most often occurring false object. It moves with the real object and enlarges the detected foreground region. Basic GMM algorithm does not distinguish between moving objects and moving shadows. Hence, in [7] an improved GMM was proposed that is able to detect shadows and to improve final results by accelerating the learning process. Shadow areas can be eliminated by the analysis of HSV and YIQ colour spaces as well. It was proved that shadows change only the image intensity value without influencing the hue. Therefore, the comparison of foreground images estimated for H component of HSV colour space and Ycomponent of YIQ colour space excludes the shadow ([8]). In order to eliminate smaller false detections the morphological operations can be used. Additionally, models based on GMM are very sensitive to sudden changes in lighting. In order to avoid false detections the authors of [8] proposed the use of colour features as well as gradients.

According to variable backgrounds, GMM method proposed in [5] was chosen to be employed in the developed system. It allows for building background model adaptively on the basis of every following frame. In order to eliminate shadow areas the analysis of HSV and YIQ colour spaces described in [9] could be applied together with morphological operations. Moreover, the system could operate only on selected region what was proposed in [2] and object features could be analysed with the use of simple descriptors as it was presented in [3].

3 Parameters and Scenarios of the Developed System

Let us assume that the system will be implemented in several different places and will be aimed at different working conditions. Such combination of intended or predicted situation is called a scenario. Each scenario defines a group of functions and determines the types of algorithms used. Scenarios are system operational modes and are not overlapping each other. Only one mode can be active at the particular moment. Each mode is characterized by particular parameters and functions that are movement detection, object tracking, object classification, region limitation, object size, event detection, weather conditions and work time.

Usually, any detected movement induces the alarm while the lack of activity does not. However, while an ill person is supervised and does not move it can mean afainting. Every object that will be tracked by the system will move along a certain path, called the trajectory. System will follow an object, analyse its behaviour and decide which is a suspicious one. Typical monitoring system turns the alarm on independently from object type. The reaction of the system that includes classification module depends on the type or the size of monitored object. Therefore, it will be possible to specify groups of objects with particular characteristics that can enter observed area. Moreover, system will allow to limit the region that is under analysis or to determine the border (line) that could not be crossed by a moving object. That would reduce the computation time by excluding the regions with lower possibility of movement occurrence and could increase system efficiency in every scenario. Event detection will aim at dynamic analysis of object appearance during following frames. For instance, it will detect changes in object shape that could indicate suspicious behaviour or dangerous situation. Moreover, conditions in which the system will finally work are very important. Changing weather conditions could cause e.g. limited visibility, sudden changes in lighting or shadow occurrence. Those factors will make image analysis more difficult and will influence the final foreground region size. While working at night or in the case of insufficient lighting the use of artificial lighting will be necessary. The cost of infrared cameras is too high.

Each scenario has a specific role that simply results in utilized parameters. For instance, if the movement detection is a key issue, classification methods are redundant. SmartMonitor is planned to work in four independent modes:

– Scenario A: Home/surrounding protection against unauthorized intrusion,
– Scenario B: Supervision over ill person,
– Scenario C: Crime detection,
– Scenario D: Smoke and fire detection.

Scenario A is very similar to the scenarios realized in the traditional monitoring system that induces the alarm in the case of any detected movement. However, in the developed system we will introduce basic classification that will determine whether the moving object is a human or not. Then, object size could be a classification condition. In this mode, the main goal of the image analysis will be to indicate suspicious objects. Weather conditions will have to be taken into consideration as well.

In scenario B the movement and the behaviour will be analysed. Sudden changes in object features will be detected. That will concern e.g. a shape change or a lack of movement. Ill person supervision mode will utilize tracking without classification. Then, image analysis will aim at identifying possible untypical behaviour, e.g. diabetic fainting caused by low blood sugar level. The key issue will be to recognize afall as well.

The third scenario is similar to the previous one – interiors will be monitored as well. Sudden changes in object trajectory and appearance that could indicate suspicious behaviours will be detected, such as raising hands up. Scenario C could be used for monitoring small shops or offices in the attack situations when the offender prevents from activating the alarm. System could recognize a danger and send information as well as camera image to the security services.

The last scenario is an additional one. System could work inside as well as outside the building where the lighting is changeable. The main task will be to detect the movement of specified objects with defined sizes and features (smoke and fire). The classification without tracking will be performed.

Such classification of scenarios makes it possible to select particular algorithms aimed at specific needs, thus it can help in tuning the whole system in order to make it much more precise. It means, that the system aimed at supervising potentially ill persons will not detect crime situations, while a guarding system will not be targeted at detecting fainting.

4 Description of the Experiments and Exemplary Results

A common task for all scenarios will be the analysis of video sequences captured by cameras placed in various locations such as apartments, shops, gardens or different types of buildings and their surroundings. In order to perform the development of the system it is very important that the test database reflects real situations in which the system will finally work. Despite many benchmark solutions (e.g. [10–12]) there is no publicly available database that is universal, free of charge and meets system requirements in full. They usually concern specified applications and situations. Therefore, a set of video sequences containing real scenes that match the system parameters and planned scenarios was prepared.

Some experiments were carried out in order to explore the efficiency of selected algorithms. We present the results of foreground mask extraction that was

Fig. 2. Sample experimental results for scenario A – a person crossing the garden (the description is given in the main text)

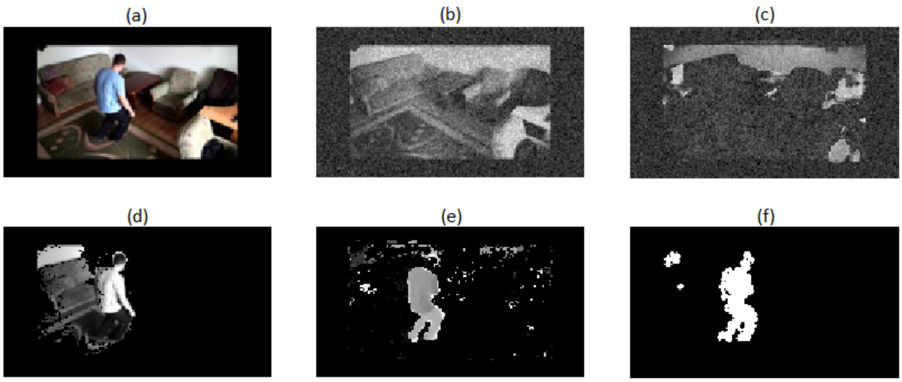

Fig. 3. Sample experimental results for scenario B – a person that is falling down (the description is given in the main text)

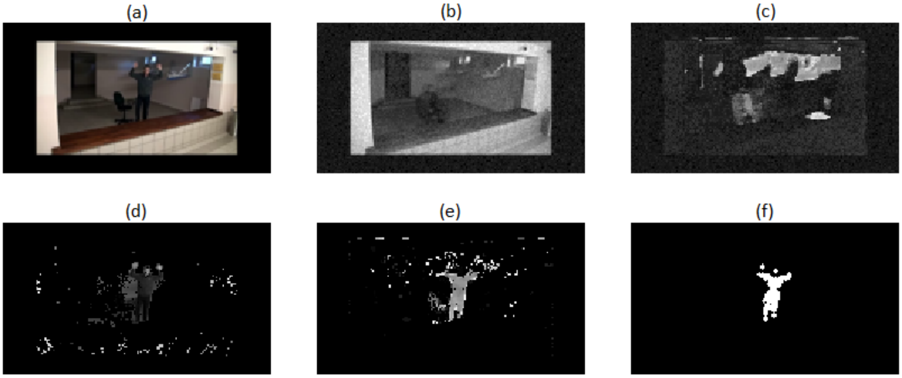

Fig. 4. Sample experimental results for scenario C – an attack situation (the description is given in the main text)

performed with the use of GMM background modelling method and the results of false object removal based on HSV and YIQ colour models analysis and morphological operations. Figures 2, 3 and 4 provide the results for sample frames and for scenario A, B and C respectively. Every following figure contains: (a) original image, (b)background image for Y component (luminance), (c) background image for Hcomponent (hue), (d) foreground image for Y component, (e) foreground image for H component and (f) a final foreground mask after false objects removal.

The experiments performed so far gave very promising results. Moving people were detected and extracted properly. Some additional regions that appeared on foreground masks were not taken into consideration on the basis of threshold values that determine minimal size. Binary form of the final image simplifies the application of shape measurements and basic classification.

5 Conclusions

In this paper the concept of SmartMonitor system was introduced. We provided basic information about planned system parameters and scenarios. Each scenario will determine the way the system will work, its functions and utilized algorithms. A brief review of intelligent monitoring methods as well as specified algorithms that are usually used was included as well. Some of them became very useful for SmartMonitor. Despite the fact that the system is being currently investigated and developed, some experimental results were presented.

Final version of SmartMonitor system will utilize commonly available hardware – personal computer and digital camera. That will give better image quality and faster data transmission. Moreover, some methods for reducing computation time could be introduced as well. That includes e.g. the limitation of the region that is under analysis. These all factors will influence the product availability and affordability for particular users. The most important advantage of the system will be the elimination of human factor – system will work independently, without human involvement. Only the initial system calibration will require it. However, that will be a benefit as well. The calibration will allow for setting individual safety rules that will be adjusted to the situation, monitored people and areas.

Since the system is not finished yet, a lot of future work can be done in order to prepare the final product version. Our aim is to utilize methods as simple as possible but also the most efficient ones. Nevertheless, in some cases more sophisticated algorithms can be taken into consideration, e.g. Histogram of Oriented Gradients (HoG). It detects and extracts objects from static images what is associated with detailed classification. HoG could be introduced to the system with some modifications and on several conditions, e.g. with time interval and on limited region. Additionally, it is important to define a group of methods for describing object features, e.g. using several shape, texture or colour descriptors.

Acknowledgments. The project *Innovative security system based on image analysis — "SmartMonitor" prototype construction* (original title: *Budowa prototypu innowacyjnego systemu bezpieczestwa opartego oanalize obrazu — "SmartMonitor"*) is the project co-founded by European Union (project number PL: UDA-POIG.01.04.0-32-008/10-00, value: 9.996.904 PLN, EU contribution: 5.848.800 PLN, realization period: 07.2011-01.2013). *European Funds – for the development of innovative economy* (*Fundusze Europejskie — dla rozwoju innowacyjnej gospodarki*).

References

1. Robertson, N., Reid, I.: A general method for human activity recognition in video. Computer Vision and Image Understanding 104, 232–248 (2006)
2. Pantrigo, J.J., Hernandez, J., Sanchez, A.: Multiple and variable target visual tracking for video-surveillance applications. Pattern Recognition Letters 31, 1577–1590 (2010)
3. Gurwicz, Y., Yehezkel, R., Lachover, B.: Multiclass object classification for real–time video surveillance systems. Pattern Recognition Letters 32, 805–815 (2011)
4. Frejlichowski, D.: Automatic localisation of moving vehicles in image sequences using morphological opertions. In: Pro. of the 1st IEEE International Conference on Information Technology, Gdansk 2008, pp. 439–442 (2008)
5. Stauffer, C., Grimson, W.E.L.: Adaptive background mixture models for real-time tracking. In: Proc. of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, pp. 2–252 (1999)
6. Zivkovic, Z.: Improved adaptive Gaussian mixture model for background subtraction. In: Proc. of the 17th International Conference on Pattern Recognition, vol. 2, pp. 28–31 (2004)
7. Kaewtrakulpong, P., Bowden, R.: An improved adaptive background mixture model for realtime tracking with shadow detection. In: Video Based Surveillance Systems: Computer Vision and Distributed Processing. Kluwer Academic Publishers (2001)

8. Javed, O., Shafique, K., Shah, M.: A hierarchical approach to robust background subtraction using color and gradient information. In: Proc. of the Workshop on Motion and Video Computing, pp. 22–27 (2002)
9. Forczmański, P., Seweryn, M.: Surveillance Video Stream Analysis Using Adaptive Background Model and Object Recognition. In: Bolc, L., Tadeusiewicz, R., Chmielewski, L.J., Wojciechowski, K. (eds.) ICCVG 2010, Part I. LNCS, vol. 6374, pp. 114–121. Springer, Heidelberg (2010)
10. CAVIAR, Context aware vision using image-based active recognition, http://homepages.inf.ed.ac.uk/rbf/CAVIAR
11. IEEE International Workshop on Performance Evaluation of Tracking and Surveillance (PETS), http://www.cvg.rdg.ac.uk/PETS2010/index.html
12. CLEAR, Classification of events, activities and relationships evaluation campaign and workshop, http://www.clear-evaluation.org