Volume 59, Number 3, 2018

ACTA TECHNICA NAPOCENSIS
Electronics and Telecommunications

RISKCAM: A SMART VIDEO SURVEILLANCE APPLICATION DESIGN CONCEPT

Elena Roxana BUHUS[1)], Lacrimioara GRAMA[1)], Alexandra DOBRE[2)]
1) Signal Processing Group, Basis of Electronics Department,
Faculty of Electronics, Telecommunications and Information Technology, Technical University of Cluj-Napoca
Cluj-Napoca, 400114, Romania
I. {Roxana.Buhus, Lacrimioara.Grama}@bel.utcluj.ro
2) Faculty of Automation and Computer Science, Technical University of Cluj-Napoca
Cluj-Napoca, 400027, Romania
Alexandra.Dobre@student.utcluj.ro

Abstract: **Fire outbreak, theft, logging, and poaching in the wildlife areas or illegal trespassing in the high security outdoor protected properties are one of the topics that should to be controlled, every day. We present a design concept for a smart application that could be capable of monitoring remote, several events according with the desired configuration and report the detected risk via email alert. This approach can successfully become the next level in surveillance applications based on smart video wireless cameras, since it provides a configurable approach of the application using just one type of wireless camera rather than a dedicated smart camera for each event detection, decreasing the costs. We named our target application RiskCam. Several features were considered for it that imply fire detection, abnormal movement detection (e.g. theft, trespassing), facial authentication and notification via email for the detected alarm (e.g. "Fire present in zone 1!"). We defined and tested the concept in MATLAB, using a modular approach, except the facial recognition module that was developed in C++.**

Keywords: *video surveilance, face recognition, fire detection, smoke detection, smart application.*

I. INTRODUCTION

These days one area of demand consists of high end security and surveillance systems that can prevent or decrease the impact of unexpected events such as fires, theft, illegal trespassing, logging and poaching in the wildlife protected areas. We can see that these problems are strongly presented by the online media via organizations interested in environmental protection (e.g. AgentGreen, GreenPeace) and not only. One direct impact of unexpected fires or illegal logging of the forest is the climate change, that can lead to the beginning of the endangered species phenomenon in that certain area. This means that an imbalance can produce in that ecosystem.

The advance of electronic technologies increases the possibility to perform advanced video analysis and extract useful information [1], motion detection and characterization [2], remote sensing and pattern recognition among others [3],[4]. In many applications of pattern recognition, the recognition capability is based on a corpus of data. This data can be treated either in a holistic manner or it can be partitioned by application of prior knowledge such as texture, shape, direction, magnitude, velocity and so on from different type of objects near the area of interest [5]. Regardless of the chosen approach, applications that target a video surveillance system are faster and responsive to the society need, and can combat the mentioned weaknesses, and help the environment.

The work is structured into five sections. We expose in Section II an overview upon the defined concept for the RiskCam application, while Section III provides the theoretical background of the used methods for each feature of the application. We make the subject of Section IV, for the experimental setup, and present the obtained results in Section V. In Section VI we provide an overview of our conclusions.

II. OVERVIEW OF THE CONCEPT RISKCAM APPLICATION

Although, custom dedicated smart video devices can successfully detect remote a certain situation from a video, when it comes to significant large areas that need surveillance, the costs of such project are quite high. We assume that we have such a situation, and we would like to propose a solution that can decrease the costs in a significant way. In order to make our point a bit clearer, together with our motivation, we developed a case study in the following subchapter.

A. Case study and motivation

Let's assume that we have to develop a solution, that implies the surveillance of the property with increased security. The area of the property is made of more than twenty hectares of forest area. The area also contains fauna with engendered species, shelters for the animals and also several material deposits. The owner provided us a map of the property, that highlights the most important objectives that need surveillance, as in Fig. 1.

Manuscript received July 31, 2018; revised September 26, 2018

Figure 1. Orientative map of the objectives that need surveillance.

The owner would like a video solution that implements the below features:

- The access on the property should be done via a facial authentication, and the entrance zone
- Fire breakout should be detected at any moment
- Intruder detection based on abnormal movements should be detected in the areas of shelters and deposits (possible theft)
- Email notification with the identified alarm event (e.g. "Fire present in zone C1!")

After doing an analysis of the costs, that would imply to purchase the necessary devices to implement the request from the customer, we obtained Table I.

TABLE I. ESTIMATED COSTS FOR WIFI IP SMART CAMERAS

Camera type	**Price euro**	**No. cameras**	**Total Costs**	**Custom feature**
Wifi IP Light Blub camera	122	4	488	Facial authentication
SG7200WF Extreme life Wifi night vision detector	324	25	8100	Fire detection with smoke detection capability
Aeon OT-1080p POE wifi camera	122	10	1220	Abnormal movement detection

For each feature that the video solution should have, we identified custom IP smart wifi cameras, that can address successfully each task. Even though there are plenty of producers in this industry, we tried to select the best trade between price and our target and selected the providers [6][7] and [8]. According with Table I, a final sum of the costs would reach 9808 euro. If we scale this sum for a project that needs even more cameras, we can imagine that the costs will increase greatly. A simulation on how costs would increase related with the number of cameras, is presented in Fig. 2

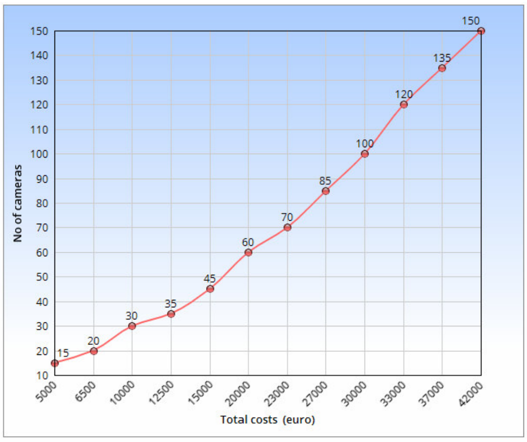

Figure 2. Costs simulation.

Considering the number of different dedicated cameras that should be used to implement a video solution that can address the customers requirements, a question raises: why not use just one type of wifi camera and develop an application that can have user configurable video detection algorithm for different taks? If we use our case study and assume that we choose the Aeon OT – 1080p POE Wi-Fi camera for surveillance of each defined objective we can redo the cost estimates. In this sense, we obtain a total cost of 3538 euro, which means a cost reduction of almost 64%.One can notice that the cost reduction was greatly reduced, which highly motivated us to develop a new approach for an configurable integrated smart video application, as we will present it in the following subchapter.

B. Concept and development methodology

In the first phase the application (App) should be developed as a prototype in MATLAB, to make sure a proper selection of the algorithms was done. The target implementation is for smart devices, but prior to that a PC version of the application should be done and fully tested. The design flow is presented in Fig. 3.

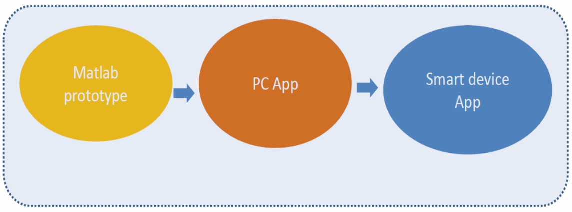

Figure 3. Application design flow.

Since we are talking about a complex project, the development process was estimated for almost two years, within a team of minimum three program engineers. This setup should help obtaining a final reliable and competitive product for the industry. Within Fig.4, we expose the recommended development process. The term module, represents the block that implements a certain functionality, defined with a certain scope (e.g. facial authentication). One can see the development cycle of the module, from requirement definition unti the testing phase.

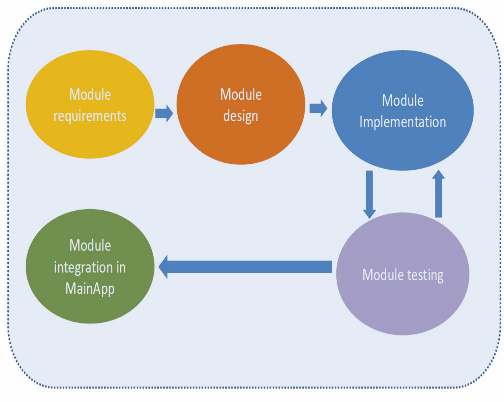

Figure 4. Development process for the application.

The idea of the application, implies the existence of an intuitive graphical interface, that allows the user to configure the folowing options:

- the number of cameras connected to the system (identified via their IP),
- the algorithms that are used for the surveilance system,
- the mapping between what algorithm is applied to the images from what camera, to perform one or more tasks,
- alarm definition for each detected event by camera N, other than the default text,
- email notification with the triggered alarm, for remote awarness.

If we consider the mentioned case study, a posible alarm definition can be the one listed in Table II.

TABLE II. ALARMS CONFIGURATION EXAMPLE

Feature	**No. cameras**	**Alarm**
Facial authentication	4	"Authentication failed @Cam_i!", i=1,4
Fire detection	25	"Fire present @Cam_i!", i=1,25
Abnormal movement detection	10	"Intruder detected@Cam_i! ", i=1,25

We have named the application: RiskCam. A graphic representation of the concept of the application was implemented in Fig. 5.

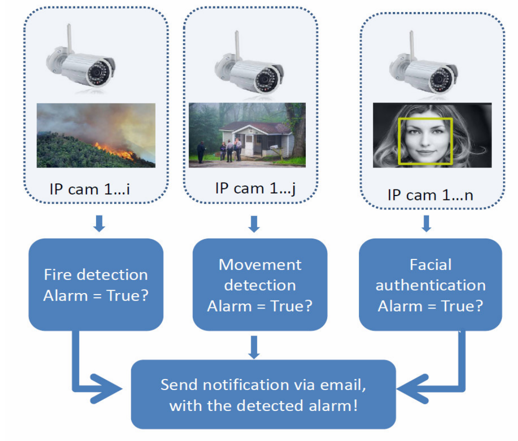

Figure 5. Concept application RiskCam.

The proposed concept is novel, and we thing that is worth investing time, funds and effort is implementing a prototype to prove the concept. An Indiegogo [8] campaign might be our option to make this target realistic.

III. THEORETICAL BACKGROUND

For each module of the application, we researched and tested algorithms dedicated to resolve each task for the event detection from video. In the below subchapters we present the theory developed behind each algorithm used by each module.

A. *Fire detection*

The detection of fire from an image was done by using a fusion of two techniques. In this sense the fire pixels detection was done using a color model. We blended RGB (red-green-blue), YCbCr (luma-blue relative to green-red relative to green) and HSV (hue-saturation-value) for reasons of segmentation the fire pixels. The smoke pixels are detected using Fuzzy Rules [12].

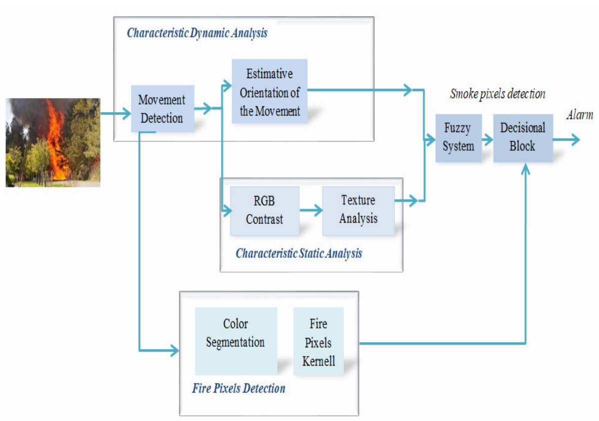

Figure 6. Solution of smoke/fire detection by blending the color model and Fuzzy rules.

1) *Fire color model definition*

Fire detection can be done by using an implementation of the color models and the technique of moving objects segmentation. In order to create a color model to segment the fire pixels, there can be taken into account one or more color spaces for which a set of characteristic rules are applied on the fire pixels. This being said, for this proposed implementation a fusion between RGB (red-green-blue), YCbCr (luma-blue relative to green-red relative to green) and HSV (hue-saturation-value) was used to segment the fire pixels. The implementation was done by using some step values determined experimentally, as they are presented in Table I.

TABLE III. FIRE PIXEL CHARACTERIZATION

Color Space	Rules for classifing fire pixels
RGB	*1. R > G > B* *2. If R > 200 ∩ G > 130 ∩ B < 120*
HSV	*1. H €[0.13, 0.46)* *2. S €[0.1, 0.34]* *3. V €[0.96, 1]*
YCbCr	*1. Y (x, y) ≥ Cb (x, y)* *2. Cr (x, y) ≥ CB (x, y)* *3. Y (x, y) ≥ 180 ∩ Y (x, y) □ 210* *4. Cb (x, y) ≥ 80 ∩ Cb (x, y) □ 120* *5. Cr (x, y) ≥ 80 ∩ Cr (x, y) □ 139*

These values can slightly suffer from some changes depending on the monitored area and the performance of the camera. The final kernel is obtained from intersecting the results of the rules for each color space.

2) *Smoke detection*

The solution given for smoke detection is a video detector, using Fuzzy inference together with dynamic and static components of the smoke, similar with the one from [13]. The method suggests two stages of analysis: one on the dynamics characteristics and the other one on the static characteristics. After these steps, the result is sent to a fuzzy inference system. The system sends an input to a decisional block for setting the fire alarm.

For examining the dynamic of the smoke, the movement of the smoke is considered one of the most important characteristics, exploited in three steps, as explained below.

Movement detection: for movement detection in a video, the subtraction between each two successive frames is performed and the difference is compared with T1, a predefined threshold. The difference is computed by the following formula:

$$d(x,y,t+\Delta t) = \begin{cases} 1, if\ |I(x,y,t+\Delta t) - I(x,y,t)| > T1 \\ 0, otherwise \end{cases}$$
(1)

Movement direction: one of the most important characteristics of smoke is the ascending movement due to hot air flow. A transform is applied for each block of motion, using the following formula:

$$argmin_{b'_{ij}} \left|\sum_{x,y \in b_{ij}} I(x,y,t) - \sum_{x,y \in b'_{ij}} I(x,y,t+\Delta t)\right|$$
(2)

where b_{ij} is a candidate of the motion block in the frame t, and b_{ij} is the search block in the frame t + Δt.

The search block includes eight directions: *0, 45, 90, 135, 180, 225, 270, 315 degrees* [y2] around the motion block candidate, in the frame t + Δt with the distance dis = 1, 2, 3 pixels to the center; this being said, b_{ij} contains a total of 8*3 = 24 candidates as motion blocks.

Accumulation of movement: for avoiding mistakes caused by direction of the movement analysis, a waiting time Wt was considered for each of the motion block in the frame t. The waiting time is the time in which the previous step's estimated movement direction accumulates and the direction with the maximum number of occurrences is reported as the direction of interest in that motion block.

The static analysis, consists of two steps as it follows.

RGB Contrast Estimation: usually, the smoke is grey and the values for the three components R, G, B are similar. For a neighborhood of 2*2 pixels, an RGB-contrast image is computed after the following formula:

$$RGCc(i,j) = \begin{cases} \min(R,G,B)\ (i+j)\ mod\ 3 = 2 \\ median(R,G,B)(i+j)\ mod\ 3 = 1 \\ \max(R,G,B)\ (i+j)\ mod\ 3 = 0 \end{cases} \quad (3)$$

Texture analysis: the result of evaluating an RGB-contrast characteristic is a texture image different from a non-greyscale image. The texture property is extracted using the following formula:

$$R = \left(1 - \frac{1}{1+\sigma^2}\right) \quad (4)$$

where σ has small values for similar R, G, B values, so that R has small values for colors similar to the ones of smoke.

Fuzzy inference system (FIS) is realized for modeling by means of quality the detection of an event, in this case a fire, without quantity details.

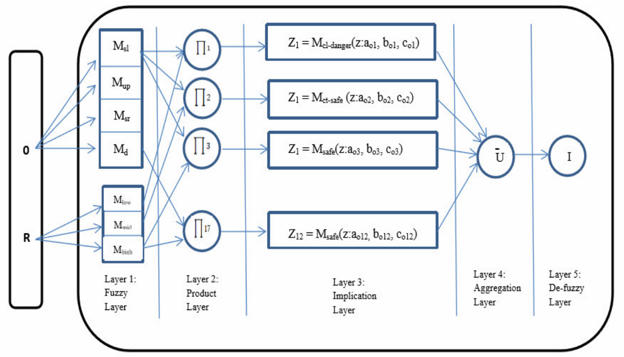

Figure 7. Fuzzy Inference System [y2].

The structure of Mamdani FIS model is realized from a set of inputs which represent the direction of the motion of the motion block and the computed texture for each block of motion. FIS is applied on each motion block (MB) and for each MB a score based on merging the inputs in a Fuzzy manner is attached. The FIS output is p, an accommodation value which represents the probability that a MB is a smoke block (SB). This procedure is applied for all the MBs in the image.

In general, if an image contains N blocks from which Ns = a * N blocks are SB, then the Fuzzy system can produce an ideal value of "1" for each SB which would theoretically provide a threshold T = a * N as the threshold for fire detection. In reality, concerning the characteristics of SBs, a value between 0 and 1 is computed. In conclusion, the applicable threshold is T = (a * N) / b with a fire threshold where (a / b) << 1.

B. Movement detection from video

Movement detection of abnormal activities consists in constantly monitoring the difference between two consecutive frames and comparing it with a step value. If the value of the difference is bigger than the step value, then an abnormal activity is considered to exist into the monitored scene. What is more, the capture camera is considered to be static. The step value was experimentally determined and can be adapted to the sensibility needed for detecting the activities. In Fig.8, we present the implementation of the technique.

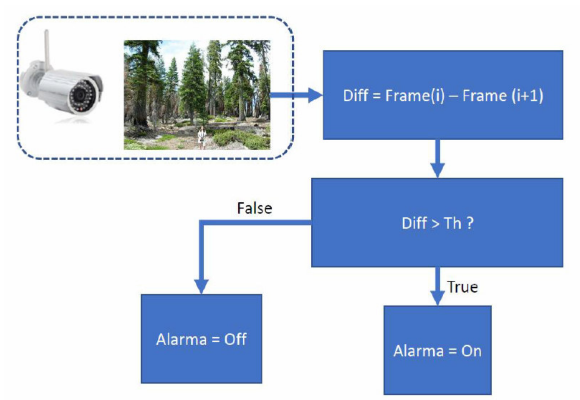

Figure 8. Detection of abnormal activity from video.

C. Facial authentication

The facial authentication was implemented by using a new technique as presented in [10], [9] based on local binary patterns and incremental supervised learning. A consistent description of the theoretical aspects can be found in [10] and [9]. We find no reason to reiterate here these theoretical aspects, rather just mention the reference of the used implementation.

IV. EXPERIMENTAL SETUP

The resources and tools used for the experimental setup, are:

- Intel Core i3 CPU 2.40 GHz @2.40 GHz, 4GB RAM
- Windows 7 Pro 64 bits
- OpenCV library version 2.4.6
- Visual Studio 10 Premium
- MATLAB Student Edition

We defined a MATLAB user interface to test the fire detection and abnormal motion detection, as presented in Fig. 9. One can see that it has the possibility to test the fire detection from images, but also from a loaded video. The motion detection was tested from the webcam of the PC. We have selected several videos, and images (60) with fire over the internet and used them for tests.

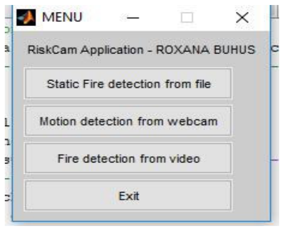

Figure 9. Matlab interfate to test the fire detection and abnormal movement.

The facial authentication was developed and tested in C++, as mentioned in [10]. As mentioned in [10], we developed a .exe, that gets for input five arguments. For a detailed description check work [10]. What we see in Fig. 10, is an example on how the application was called and the obtained results for a Yale data set, with 16 persons and 70% data used for training. For this scenario we obtain a recognition rate of 86%.

Figure 10. Comand defined to run the facial authentication application.

We have defined a collection of images to test the application, as presented in Fig. 11.

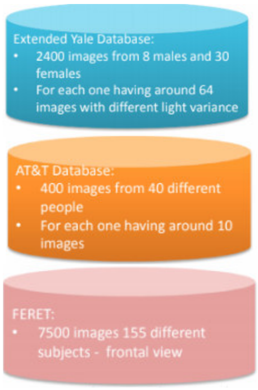

Figure 11. Description for the testing images collection.

We provide the results in the following chapter.

V. EXPERIMENTAL RESULTS

All the results obtained for the facial authentication method are present in [10] and [9].

The abnormal detection module was tested via the webcam. We have evaluated the recognition rate, by observing the Matlab logs, on consecutive scenes. We have defined two types of test: negative and positive tests. In the negative tests (Ni, i=1…15) the object stays still, therefore no movements is detected, while in the positive tests (Pi, i=1…15) the object is moving. All tests passed, with a response of two seconds. Am example of test is presented in Fig. 12.

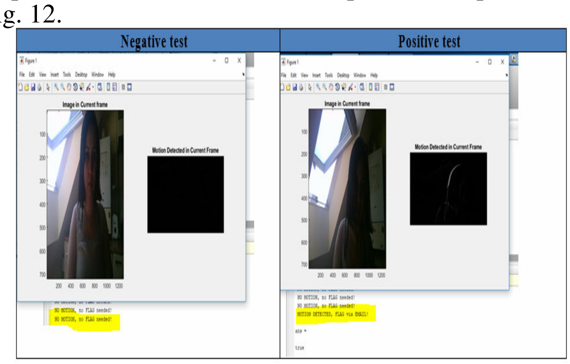

Figure 12. Motion detection tests example.

For fire detection we have performed a static analysis of the performances, but also a dynamic one from the video. For the static analyses we have used a set of 39 images (Fi, i=1…39), and defined a region of interest (ROI) for fire pixels. If the ROI is present we used the notation ROI-P, while the absence was noted with ROI-A. Our validation process implies the usage of the terminology positive truth (TP) and negative truth (TN) for a correct decision for the test. A negative false (FP) is when the algorithm provided a false classification, while a negative false (FN) is when the algorithm identified a ROI though is absent from the image. The accuracy was measured by the ability of the algorithm to detect the ROI, using the formula: Accuracy = TP/(TP+TN) *100%. Based on the observed results for the static analyses, we have obtained an accuracy of 89%. Results examples are provided in Fig. 13.

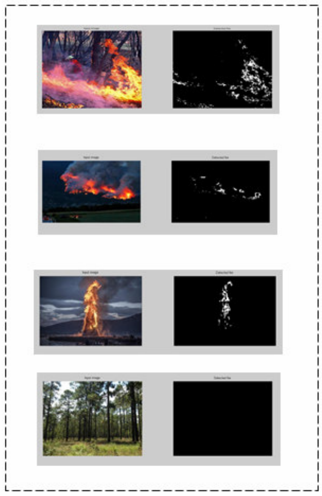

Figure 13. Static analysis fire pixel ROI results

For the dynamic testing we have used three sections of video that contain fire and smoke pixels. In the first scenario we have used 145 frames from the video, while in the second scenario we have used 625 frames, and in the third 185 frames. When fire or smoke is detected on a frame the message "Danger Smoke – Detected!" is visible. Example of the obtained results, in both scenarios are presented in Fig. 14.

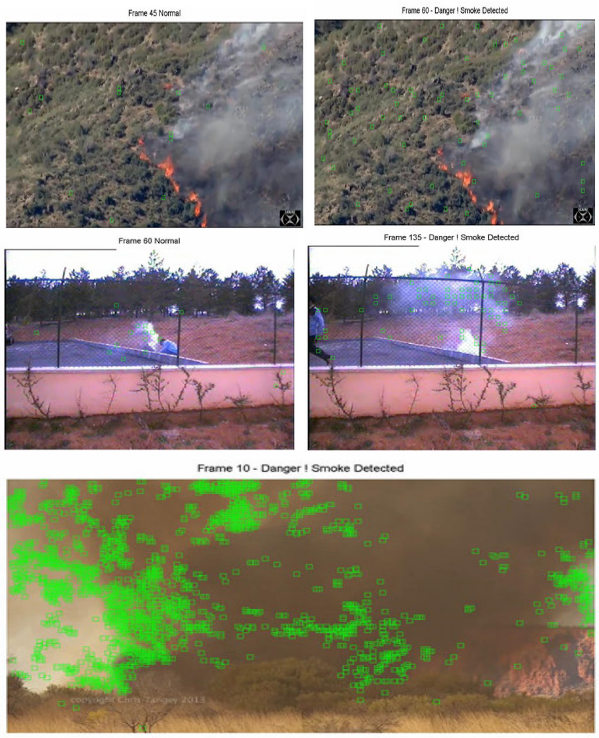

Figure 14. Dynamic analysis fire and smoke detection example results.

The obtained accuracy for the fire and smoke detection in the video is 92%.

VI. CONCLUSIONS

In this paper, we present a design concept for a smart surveillance application. The obtained results for each module prototype are promising.

This approach can successfully become the next level in surveillance applications based on smart video wireless cameras. This is possible due to its configurable approach of the application using just one type of wireless camera rather than a dedicated smart camera for each event detection, decreasing the costs as highlighted in Fig. 2.

The next step would be implementing the prof of concept as a PC application, under Windows.

REFERENCES

[1] Jones, B.T.: Low-Cost Outdoor Video Motion and Non-Motion Detection. Processing of Security Technology, 376–380 (1995)

[2] Paladino, V.: Introduction of Video Compression Under the Standard MPEG-2. The Institute of Electronic Engineer,Spain, 3–24 (2005)

[3] Zhang, Z.: Mining Surveillance Video for Independent Motion Detection. In: IEEE Internacional Conference on Data Mining, pp. 741–744 (2005)

[4] Hariharakrishnan, K., Schonfeld, D., Raffy, P., Yassa, F.: Video Tracking Using Block Matching. In: IEEE, International Conference on Image Processing, pp. 945–948 (2003)

[5] Avidan, S.: Support Vector Tracking. IEEE 26(8), 1064–1071 (2004)

[6] http://www.spytecinc.com/720p-hd-wi-fi-ip-light-bulb-camera.html [last accessed 19.06.2018]

[7] http://www.spytecinc.com/xtremelife-quad-smoke-detector.html [last accessed 19.06.2018]

[8] http://www.spytecinc.com/aeon-ot-poe-hd-wi-fi-wireless-video-monitoring-camera-outdoors.html [last accessed 19.06.2018]

[9] E. R. Buhuş, L. Grama and C. Şerbu,“ Histograms and Supervised Learning for Facial Recognition Applications”,Novice insights in Electronics and Telecomunications, Student Symposium on Electronics and Telecomunications 2018, pp 83-84

[10] E. R. Buhuş, L. Grama and C. Şerbu, "A facial recognition application based on incremental supervised learning," 2017 13th IEEE International Conference on Intelligent Computer Communication and Processing (ICCP), Cluj-Napoca, 2017, pp. 279-286 doi: 10.1109/ICCP.2017.8117017)

[11] K. Yesu, H. J. Chakravorty, P. Bhuyan, R. Hussain and K. Bhattacharyya, "Hybrid features based face recognition method using Artificial Neural Network," 2012 2nd National Conference on Computational Intelligence and Signal Processing (CISP), Guwahati, Assam, 2012, pp. 40-46.doi: 10.1109/NCCISP.2012.6189705

[12] Yashar Deldjoo, Fatemef Nazary, Ali M. Fotoului, “A Novel Fuzzy – Based Smoke Detection System Using Dynamic and Static Smoke Features”, Iranian Conference on Electrical Engineering, Volume 23 – 2015

[13] F. Yuan, “A fast accumulative motion orientation model based on integral image for video smoke detection”, Pattern Recognition Letters, Volume 29, No. 7, pp.925 -932 – 2008

[14] Rui Huang, V. Pavlovic and D. N. Metaxas, "A hybrid face recognition method using Markov random fields," Proceedings of the 17th International Conference on Pattern Recognition, 2004. ICPR 2004., 2004, pp. 157-160 Vol.3.doi: 10.1109/ICPR.2004.1334492

[15] S. Mandal and B. C. Dhara, "A Hybrid Face Recognition Method Based on Structural and Holistic Features," 2009 Seventh International Conference on Advances in Pattern Recognition, Kolkata, 2009, pp. 441-444.doi: 10.1109/ICAPR.2009.60