



Video fire detection – Review



A. Enis Çetin^a, Kosmas Dimitropoulos^b, Benedict Gouverneur^c, Nikos Grammalidis^b,
Osman Günay^a, Y. Hakan Habiboğlu^a, B. Uğur Töreyn^{d,*}, Steven Verstockt^e

^a Department of Electrical and Electronics Engineering, Bilkent University, Ankara, Turkey

^b Information Technologies Institute, Centre of Research and Technology Hellas, 1st km Thermi-Panorama Rd, 57001 Thermi-Thessaloniki, Greece

^c Xenics Infrared Solution, Ambachtenlaan 44, Leuven, Belgium

^d Department of Electronic and Communication Engineering, Cankaya University, Ankara, Turkey

^e Multimedia Lab, ELIS Department, Ghent University, iMinds, Gaston Crommenlaan 8, bus 201, Ledeborg-Ghent, Belgium

ARTICLE INFO

Article history:

Available online 19 July 2013

Keywords:

Video based fire detection
Computer vision
Smoke detection
Wavelets
Covariance matrices
Decision fusion

ABSTRACT

This is a review article describing the recent developments in Video based Fire Detection (VFD). Video surveillance cameras and computer vision methods are widely used in many security applications. It is also possible to use security cameras and special purpose infrared surveillance cameras for fire detection. This requires intelligent video processing techniques for detection and analysis of uncontrolled fire behavior. VFD may help reduce the detection time compared to the currently available sensors in both indoors and outdoors because cameras can monitor “volumes” and do not have transport delay that the traditional “point” sensors suffer from. It is possible to cover an area of 100 km² using a single pan-tilt-zoom camera placed on a hilltop for wildfire detection. Another benefit of the VFD systems is that they can provide crucial information about the size and growth of the fire, direction of smoke propagation.

© 2013 Elsevier Inc. All rights reserved.

1. Introduction

Video surveillance cameras are widely used in security applications. Millions of cameras are installed all over the world in recent years. But it is practically impossible for surveillance operators to keep a constant eye on every single camera. Identifying and distilling the relevant information is the greatest challenge currently facing the video security and monitoring system operators. To quote New Scientist magazine: “There are too many cameras and too few pairs of eyes to keep track of them” [1]. There is a real need for intelligent video content analysis to support the operators for undesired behavior and unusual activity detection before they occur. In spite of the significant amount of computer vision research commercial applications for real-time automated video analysis are limited to perimeter security systems, traffic applications and monitoring systems, people counting and moving object tracking systems. This is mainly due to the fact that it would be very difficult to replicate general human intelligence.

Fire is one of the leading hazards affecting everyday life around the world. Intelligent video processing techniques for the detection and analysis of fire are relatively new. To avoid large scale fire and smoke damage, timely and accurate fire detection is crucial. The sooner the fire is detected, the better the chances are for survival. Furthermore, it is also crucial to have a clear understanding of the

fire development and the location. Initial fire location, size of the fire, the direction of smoke propagation, growth rate of the fire are important parameters which play a significant role in safety analysis and fire fighting/mitigation, and are essential in assessing the risk of escalation. Nevertheless, the majority of the detectors that are currently in use are “point detectors” and simply issue an alarm [2]. They are of very little use to estimate fire evolution and they do not provide any information about the fire circumstances.

In this article, a review of video flame and smoke detection research is presented. Recently proposed Video Fire Detection (VFD) techniques are viable alternatives or complements to the existing fire detection techniques and have shown to be useful to solve several problems related to the traditional sensors. Conventional sensors are generally limited to indoors and are not applicable in large open spaces such as shopping centers, airports, car parks and forests. They require a close proximity to the fire and most of them cannot provide additional information about fire location, dimension, etc. One of the main limitations of commercially available fire alarm systems is that it may take a long time for carbon particles and smoke to reach the “point” detector. This is called the transport delay. It is our belief that video analysis can be applied in conditions in which conventional methods fail. VFD has the potential to detect the fire from a distance in large open spaces, because cameras can monitor “volumes”. As a result, VFD does not have the transport and threshold delay that the traditional “point” sensors suffer from. As soon as smoke or flames occur in one of the camera views, it is possible to detect fire immediately. We all know that human beings can detect an uncontrolled fire using their eyes and

* Corresponding author.

E-mail address: toreyin@cankaya.edu.tr (B.U. Töreyn).

vision systems but as pointed out above it is not easy to replicate human intelligence.

The research in this domain was started in the late nineties. Most of the VFD articles available in the literature are influenced by the notion of 'weak' Artificial Intelligence (AI) framework which was first introduced by Hubert L. Dreyfus in his critique of the so-called 'generalized' AI [3,4]. Dreyfus presents solid philosophical and scientific arguments on why the search for 'generalized' AI is futile [5]. Therefore, each specific problem including VFD fire should be addressed as an individual engineering problem which has its own characteristics [6]. It is possible to approximately model the fire behavior in video using various signal and image processing methods and automatically detect fire based on the information extracted from video. However, the current systems suffer from false alarms because of modeling and training inaccuracies.

Currently available VFD algorithms mainly focus on the detection and analysis of smoke and flames in consecutive video images. In early articles, mainly flame detection was investigated. Recently, smoke detection problem is also considered. The reason for this can be found in the fact that smoke spreads faster and in most cases will occur much faster in the field of view of the cameras. In wildfire applications, it may not be even possible to observe flames for a long time. The majority of the state-of-the-art detection techniques focuses on the color and shape characteristics together with the temporal behavior of smoke and flames. However, due to the variability of shape, motion, transparency, colors, and patterns of smoke and flames, many of the existing VFD approaches are still vulnerable to false alarms. Due to noise, shadows, illumination changes and other visual artifacts in recorded video sequences, developing a reliable detection system is a challenge to the image processing and computer vision community.

With today's technology, it is not possible to have a fully reliable VFD system without a human operator. However, current systems are invaluable tools for surveillance operators. It is also our strong belief that combining multi-modal video information using both visible and infrared (IR) technology will lead to higher detection accuracy. Each sensor type has its own specific limitations, which can be compensated by other types of sensors. Although it would be desirable to develop a fire detection system which could operate on the existing closed circuit television (CCTV) equipment without introducing any additional cost. However, the cost of using multiple video sensors does not outweigh the benefit of multi-modal fire analysis. The fact that IR manufacturers also ensure a decrease in the sensor cost in the near future, fully opens the door to multi-modal video analysis. VFD cameras can also be used to extract useful related information, such as the presence of people caught in the fire, fire size, fire growth, smoke direction, etc.

Video fire detection systems can be classified into various sub-categories according to

- (i) the spectral range of the camera used,
- (ii) the purpose (flame or smoke detection),
- (iii) the range of the system.

There are overlaps between the categories above. In this article, video fire detection methods in visible/visual spectral range are presented in Section 2. Infrared camera based systems are presented in Section 3. Flame and smoke detection methods using regular and infrared cameras are also reviewed in Sections 2 and 3, respectively. In Sections 4 and 5, wildfire detection methods using visible and IR cameras are reviewed. Finally, conclusions are drawn in the last section.

2. Video fire detection in visible/visual spectral range

Over the last years, the number of papers about visual fire detection in the computer vision literature is growing exponentially [2]. As is, this relatively new subject in vision research is in full progress and has already produced promising results. However, this is not a completely solved problem as in most computer vision problems. Behavior of smoke and flames of an uncontrolled fire differs with distance and illumination. Furthermore, cameras are not color and/or spectral measurement devices. They have different sensors and color and illumination balancing algorithms. They may produce different images and video for the same scene because of their internal settings and algorithms.

In this section, a chronological overview of the state-of-the-art, i.e., a collection of frequently referenced papers on short range (<100 m) fire detection methods, is presented in Tables 1, 2 and 3. For each of these papers we investigated the underlying algorithms and checked the appropriate techniques. In the following, we discuss each of these detection techniques and analyze their use in the listed papers.

2.1. Color detection

Color detection was one of the first detection techniques used in VFD and is still used in almost all detection methods. The majority of the color-based approaches in VFD make use of RGB color space, sometimes in combination with HSI/HSV saturation [10,24,27,28]. The main reason for using RGB is that almost all visible range cameras have sensors detecting video in RGB format and there is the obvious spectral content associated with this color space. It is reported that RGB values of flame pixels are in red-yellow color range indicated by the rule ($R > G > B$) as shown in Fig. 1. Similarly, in smoke pixels, R , G and B values are very close to each other. More complex systems use rule-based techniques such as Gaussian smoothed color histograms [7], statistically generated color models [15], and blending functions [20]. It is obvious that color cannot be used by itself to detect fire because of the variability in color, density, lighting, and background. However, the color information can be used as a part of a more sophisticated system. For example, chrominance decrease is used in smoke detection schemes of [14] and [2]. Luminance value of smoke regions should be high for most smoke sources. On the other hand, the chrominance values should be very low.

The conditions in YUV color space are as follows:

Condition 1: $Y > T_Y$,

Condition 2: $|U - 128| < T_U$ and $|V - 128| < T_V$,

where Y , U and V are the luminance and chrominance values of a particular pixel, respectively. The luminance component Y takes values in the range [0,255] in an 8-bit quantized image and the mean values of chrominance channels, U and V are increased to 128 so that they also take values between 0 and 255. The thresholds T_Y , T_U and T_V are experimentally determined [37].

2.2. Moving object detection

Moving object detection is also widely used in VFD, because flames and smoke are moving objects. To determine if the motion is due to smoke or an ordinary moving object, further analysis of moving regions in video is necessary.

Well-known moving object detection algorithms are background (BG) subtraction methods [16,21,18,14,13,17,20,22,27,28,30,34], temporal differencing [19], and optical flow analysis [9,8,29]. They can all be used as part of a VFD system.

Table 1

State-of-the-art: underlying techniques (PART 1: 2002–2007).

Paper	Color detection	Moving object detection	Flicker/energy (wavelet) analysis	Spatial difference analysis	Dynamic texture/pattern analysis	Disorder analysis	Subblocking	Training (models, NN, SVM, ...)	Clean-up post-processing	Localization/analysis	Flame detection	Smoke detection
Phillips [7], 2002	RGB		X	X				X	X		X	
Gomez-Rodriguez [8], 2002		X	X			X						X
Gomez-Rodriguez [9], 2003		X	X			X						X
Chen [10], 2004	RGB/HSI	X				X					X	X
Liu [11], 2004	HSV		X			X					X	
Marbach [12], 2006	YUV		X			X					X	
Toreyin [13], 2006	RGB	X	X	X							X	
Toreyin [14], 2006	YUV	X	X			X						X
Celik [15], 2007	YCbCr/RGB										X	X
Xu [16], 2007		X	X			X					X	X

Table 2

State-of-the-art: underlying techniques (PART 2: 2007–2009).

Paper	Color detection	Moving object detection	Flicker/energy (wavelet) analysis	Spatial difference analysis	Dynamic texture/pattern analysis	Disorder analysis	Subblocking	Training (models, NN, SVM, ...)	Clean-up post-processing	Localization/analysis	Flame detection	Smoke detection
Celik [17], 2007	RGB	X				X		X	X		X	
Xiong [18], 2007		X	X			X						X
Lee [19], 2007	RGB	X						X	X		X	X
Calderara [20], 2008	RGB	X	X				X	X				X
Piccinini [21], 2008	RGB	X	X					X				X
Yuan [22], 2008	RGB	X				X	X					X
Borges [23], 2008	RGB					X					X	
Qi [24], 2009	RGB/HSV		X	X					X		X	
Yasmin [25], 2009	RGB/HSI	X				X	X			X		X
Gubbi [26], 2009			X				X	X				X

Table 3
State-of-the-art: underlying techniques (PART 3: 2010–2011).

Paper	Color detection	Moving object detection	Flicker/energy (wavelet) analysis	Spatial difference analysis	Dynamic texture/pattern analysis	Disorder analysis	Subblocking	Training (models, NN, SVM, ...)	Clean-up post-processing	Localization/analysis	Flame detection	Smoke detection
Chen [27], 2010	RGB/HSI	X	X					X			X	
Gunay [28], 2010	RGB/HSI	X	X	X		X		X			X	
Kolesov [29], 2010		X			X			X			X	
Ko [30], 2010	RGB	X	X			X		X			X	
Gonzalez-Gonzalez [31], 2010			X									X
Borges [32], 2010	RGB			X		X		X			X	
Van Hamme [33], 2010	HSV				X		X	X			X	
Celik [34], 2010	CIE L*a*b*	X				X		X	X		X	
Yuan [35], 2011					X			X		X		
Rossi [36], 2011	YUV/RGB							X	X	X		

In background subtraction methods, it is assumed that the camera is stationary. In Fig. 2, a background subtraction based motion detection example is shown using the dynamic background model proposed by Collins et al. [38]. This Gaussian Mixture Model based approach model was used in many of the articles listed in Tables 1, 2 and 3.

Some of the early VFD articles simply classified fire-colored moving objects as fire but this approach leads to many false alarms, because falling leaves in autumn or fire-colored ordinary objects, etc., may all be incorrectly classified as fire. Further analysis of motion in video is needed to achieve more accurate systems.

2.3. Motion and flicker analysis using Fourier and wavelet transforms

As it is well known, flames flicker in uncontrolled fires, therefore flicker detection [24,18,12,13,27,28,30] in video and wavelet-domain signal energy analysis [21,14,20,26,31,39] can be used to distinguish ordinary objects from fire. These methods focus on the temporal behavior of flames and smoke. As a result, flame colored pixels appear and disappear at edges of turbulent flames. The research in [16,18] shows experimentally that the flicker frequency of turbulent flames is around 10 Hz and that it is not greatly affected by the burning material and the burner. As a result, it is proposed to use frequency analysis to differentiate flames from other moving objects. However, an uncontrolled fire in its early stage exhibits a transition to chaos due to the fact that combustion process consists of nonlinear instabilities which result in transition to chaotic behavior via intermittency [40–43]. Consequently, turbulent flames can be characterized as a chaotic wide band frequency activity. Therefore, it is not possible to observe a single flickering frequency in the light spectrum due to an uncontrolled fire. This phenomenon was observed by independent researchers working on video fire detection and methods were proposed accordingly [14,44,27]. Similarly, it is not possible to talk about a specific flicker frequency for smoke but we clearly observe a time-varying meandering behavior in uncontrolled fires. Therefore, smoke flicker detection does not seem to be a very reliable technique but it can be used as part of a multi-feature algorithm fusing various vision clues for smoke detection. Temporal Fourier analysis can still be used to detect flickering flames, but we believe that there is no need to detect specifically 10 Hz. An increase in Fourier domain energy in 5 to 10 Hz is an indicator of flames.

The temporal behavior of smoke can be exploited by wavelet domain energy analysis. As smoke gradually softens the edges in an image, Toreyin et al. [14] found the energy variation between background and current image as a clue to detect the presence of smoke. In order to detect the energy decrease in edges of the image, they use the Discrete Wavelet Transform (DWT). The DWT is a multi-resolution signal decomposition method obtained by convolving the intensity image with filter banks. A standard halfband filterbank produces four wavelet subimages: the so-called low–low version of the original image C_t , and the horizontal, vertical and diagonal high frequency band images H_t , V_t , and D_t . The high-band energy from subimages H_t , V_t , and D_t is evaluated by dividing the image I_t in blocks b_k of arbitrary size as follows:

$$E(I_t, b_k) = \sum_{i,j \in b_k} H_t^2(i, j) + V_t^2(i, j) + D_t^2(i, j). \quad (1)$$

Since contribution of edges are more significant in high-band wavelet images compared to flat areas of the image, it is possible to detect smoke using the decrease in $E(I_t, b_k)$. As the energy value of a specific block varies significantly over time in the presence of smoke, temporal analysis of the ratio between the current input frame wavelet energy and the background image wavelet energy is used to detect the smoke as shown in Fig. 3.

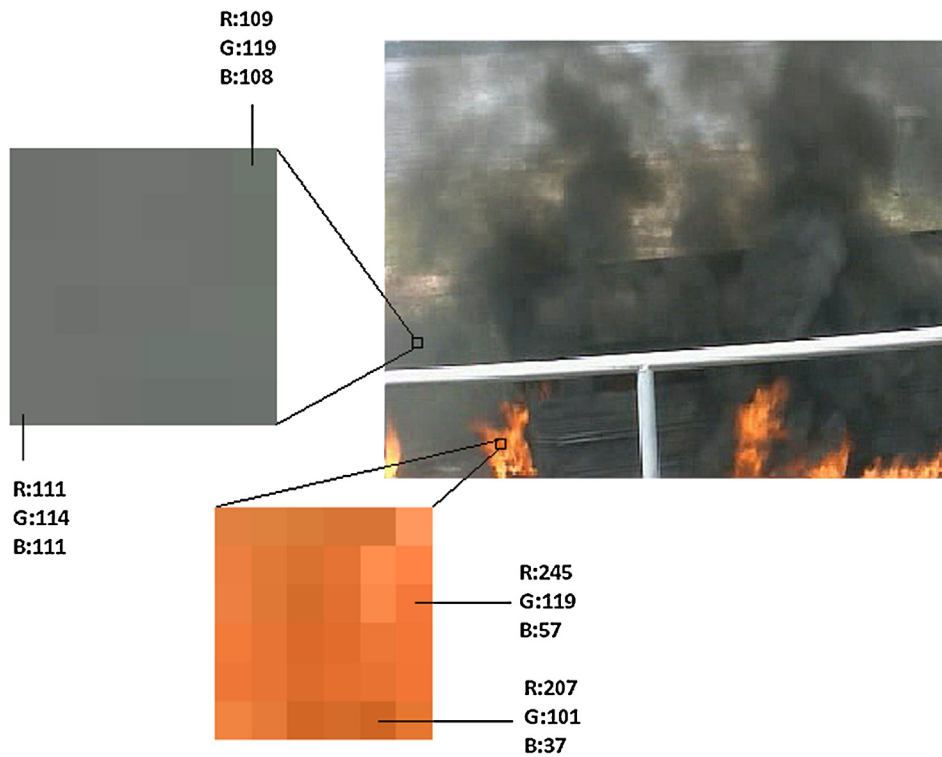


Fig. 1. Color detection: smoke region pixels have color values that are close to each other. Pixels of flame regions lie in the red–yellow range of RGB color space with $R > G > B$. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

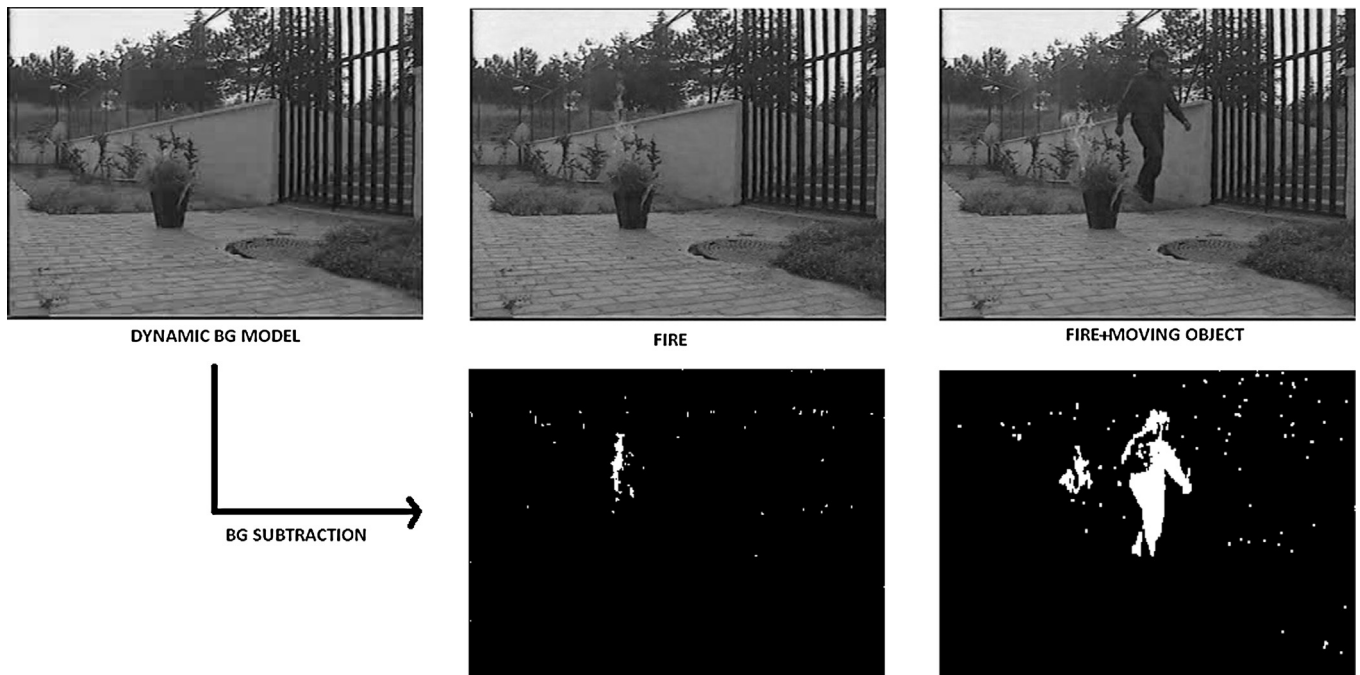


Fig. 2. Moving object detection: background subtraction using dynamic background model.

2.4. Spatial wavelet color variation and analysis

Flames of an uncontrolled fire have varying colors even within a small area. Spatial color difference analysis [24,13,28,32] focuses on this characteristic. Using range filters [24], variance/histogram analysis [32], or spatial wavelet analysis [13,28], the spatial color variations in pixel values are analyzed to distinguish ordinary fire-colored objects from uncontrolled fires. In Fig. 4 the concept of spatial difference analysis is further explained by means of a

histogram-based approach, which focuses on the standard deviation of the green color band. It was observed by Qi and Ebert [24] that this color band is the most discriminative band for recognizing the spatial color variation of flames. This can also be seen by analyzing the histograms. Green pixel values vary more than red and blue values. If the standard deviation of the green color band exceeds $t_\sigma = 50$ (\sim Borges [32]) in a typical color video the region is labeled as a candidate region for a flame. For smoke detection, on the other hand, experiments revealed that these techniques are

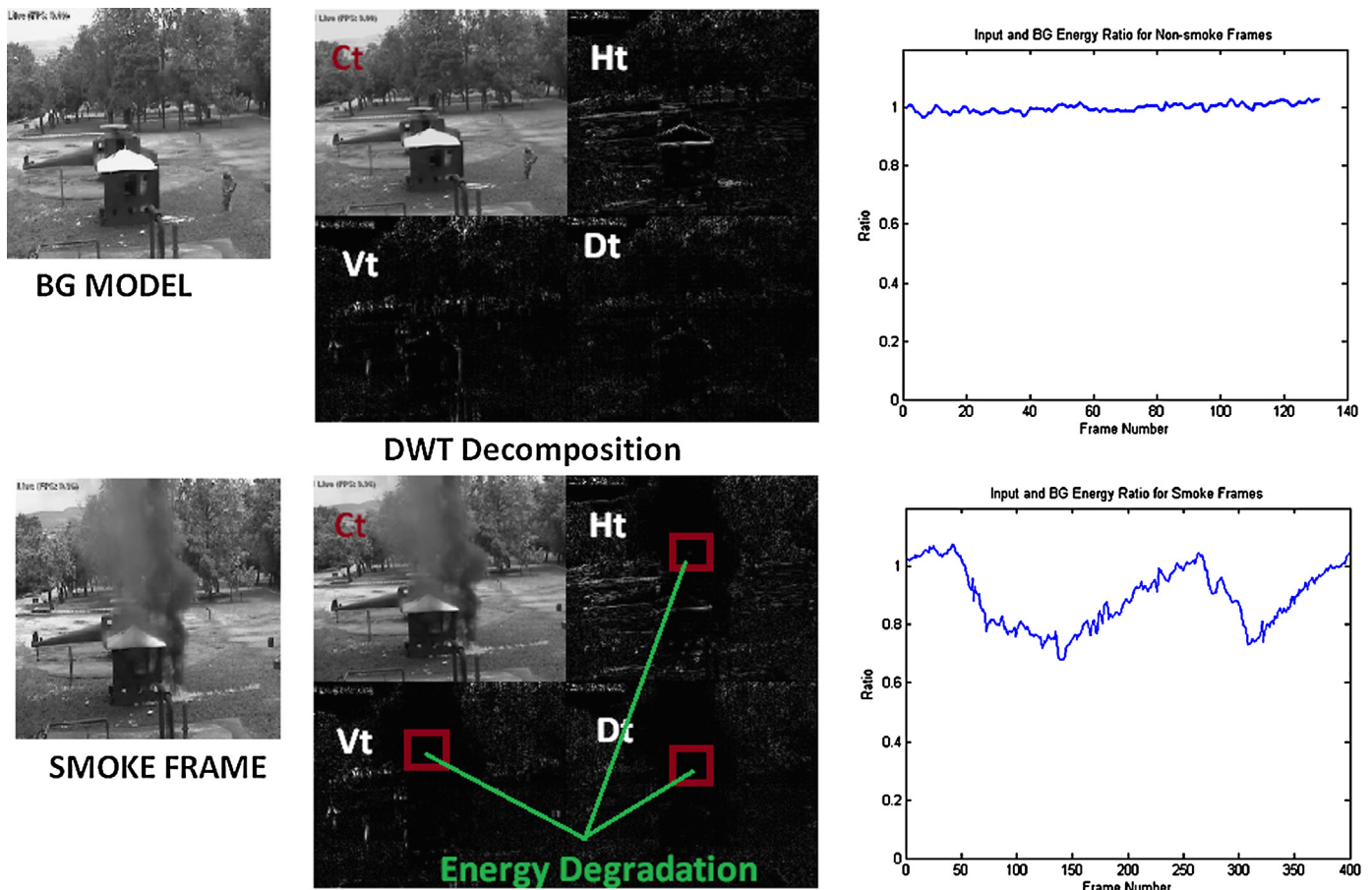


Fig. 3. DWT based video smoke detection: when there is smoke, the ratio between the input frame wavelet energy and the BG wavelet energy decreases and shows a high degree of disorder.

not always applicable, because smoke regions often do not show as high spatial color variation as flame regions. Furthermore, textured smoke-colored moving objects are difficult to distinguish from smoke and can cause false detections. In general, smoke in an uncontrolled fire is gray and it reduces the color variation in the background. Therefore, in YUV color space we expect to have reduction in the dynamic range of chrominance color components U and V after the appearance of smoke in the viewing range of camera.

2.5. Dynamic texture and pattern analysis

A dynamic texture or pattern in video, such as smoke, flames, water and leaves in the wind can be simply defined as a texture with motion [45,46], i.e., a spatially and time-varying visual pattern that forms an image sequence or part of an image sequence with a certain temporal stationarity [47]. Although dynamic textures are easily observed by human eyes, they are difficult to discern using computer vision methods as the spatial location and extent of dynamic textures can vary with time and they can be partially transparent. Some dynamic texture and pattern analysis methods in video [29,33,35] are closely related to spatial difference analysis. Recently, these techniques are also applied to the flame and smoke detection problem [46]. Currently, a wide variety of methods including geometric, model-based, statistical and motion based techniques are used for dynamic texture detection [48–50].

In Fig. 5, dynamic texture detection and segmentation examples are shown, which use video clips from the DynTex dynamic texture and Bilkent databases [51,52,50,47]. Contours of dynamic texture regions, e.g., fire, water and steam, are shown in this fig-

ure. Dynamic regions in video are seemed to be segmented very well. However, due to the high computational cost, these general techniques are not used in practical fire detection algorithms which should run on low-cost computers, FPGAs or digital signal processors. If future developments in computers and graphics accelerators could lower the computational cost, dynamic texture detection methods can be incorporated into the currently available video fire detection systems to achieve more reliable systems.

Ordinary moving objects in video, such as walking people, have a pretty stable or almost periodic boundary over time. On the other hand, uncontrolled flame and smoke regions exhibit chaotic boundary contours. Therefore disorder analysis of boundary contours of a moving object is useful for fire detection. Some examples of frequently used metrics are randomness of area size [23, 32], boundary roughness [14,11,28,32], and boundary area disorder [18]. Although those metrics differ in definition, the outcome of each of them is almost identical. In the smoke detector developed by Verstock [2], disorder analysis of the Boundary Area Roughness (BAR) is used, which is determined by relating the perimeter of the region to the square root of the area (Fig. 6). Another technique is the histogram-based orientation accumulation by Yuan [22]. This technique also produces good disorder detection results, but it is computationally more complex than the former methods. Related to the disorder analysis is the growing of smoke and flame regions in the early stage of a fire. In [31,34], the growth rate of the region-of-interest is used as a feature parameter for fire detection. Compared to disorder metrics, however, growth analysis is less effective in detecting the smoke especially in wildfire detection. This is because smoke region appears to grow very slowly in

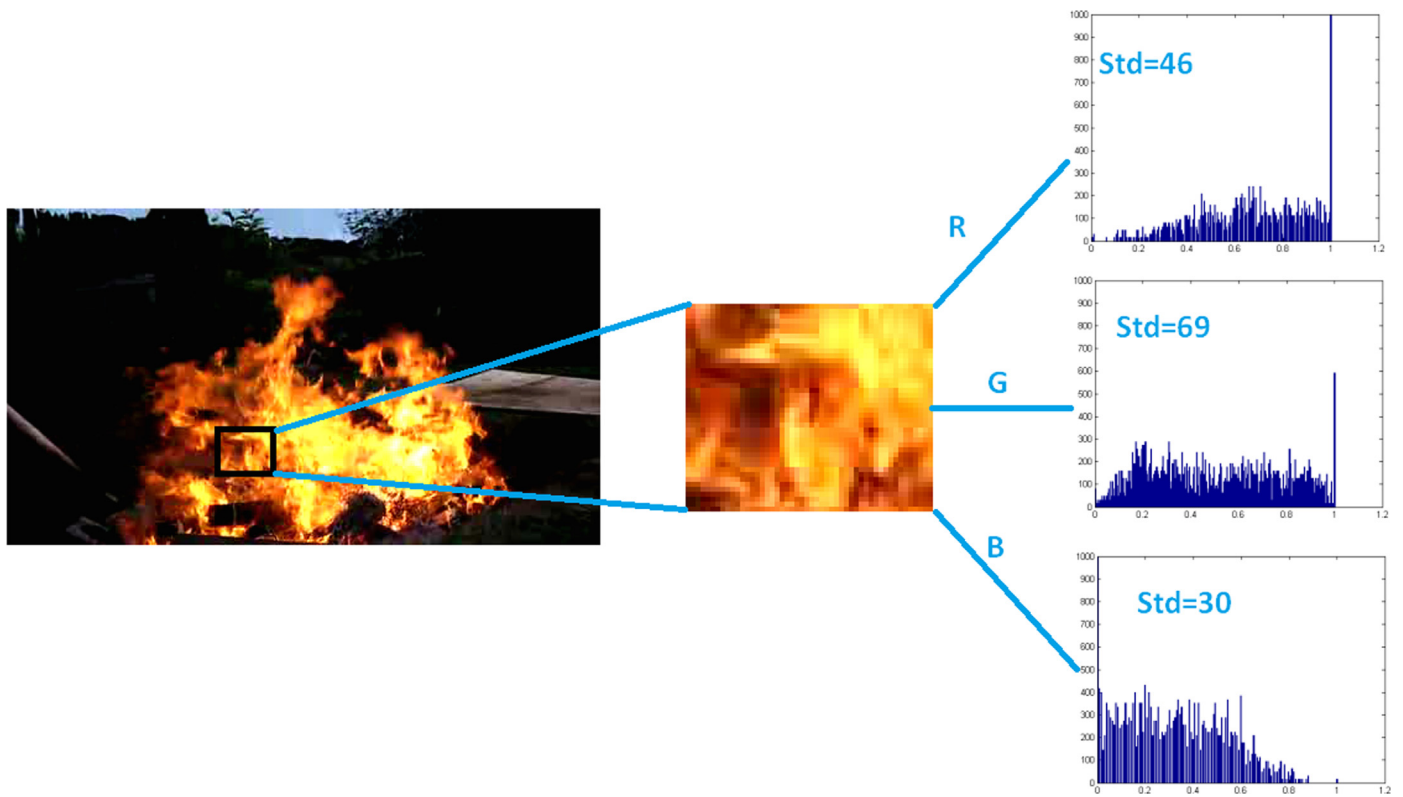


Fig. 4. Spatial difference analysis: in case of flames the standard deviation σ_G of the green color band of the flame region exceeds $t_\sigma = 50$ (\sim Borges [32]).

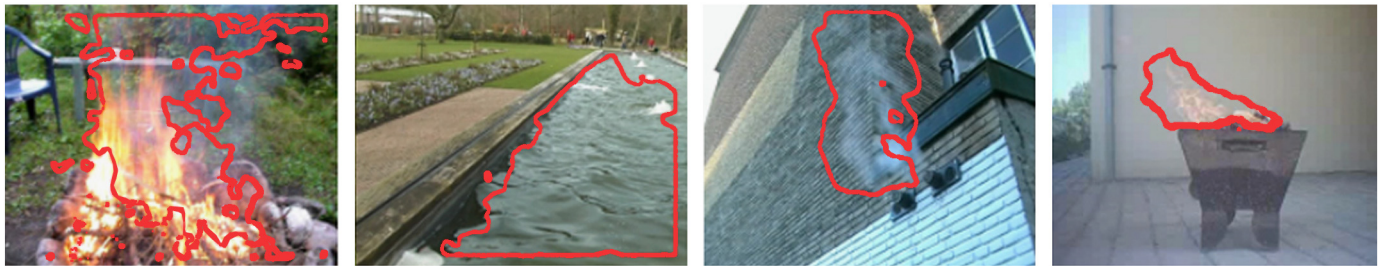


Fig. 5. Dynamic texture detection: contours of detected dynamic texture regions are shown in the figure (results from DYNTEX and Bilkent databases [51,53]).



Fig. 6. Boundary area roughness of consecutive flame regions.

wildfires when they are viewed from long distances. Furthermore, an ordinary object may be approaching to the camera.

2.6. Spatio-temporal normalized covariance descriptors

A recent approach which combines color and spatio-temporal information by region covariance descriptors is used in European Commission funded FP-7 FIRESENSE project [54–56]. The method is based on analyzing the spatio-temporal blocks. The blocks are obtained by dividing the fire and smoke-colored regions into 3D regions that overlap in time. Classification of the features is performed only at the temporal boundaries of blocks instead of performing it at each frame. This reduces the computational complexity of the method.

Covariance descriptors are proposed by Tuzel, Porikli and Meer to be used in object detection and texture classification problems [54,55]. In [57] temporally extended normalized covariance descriptors to extract features from video sequences are proposed.

Temporally extended normalized covariance descriptors are designed to describe spatio-temporal video blocks. Let $I(i, j, n)$ be the intensity of (i, j) th pixel of the n th image frame of a spatio-temporal block in video. The property parameters defined in equations below are used to form a covariance matrix representing spatial information. In addition to spatial parameters, temporal derivatives, I_t and I_{tt} are introduced, which are the first and second derivatives of intensity with respect to time, respectively. By adding these two features to the previous property set, normalized covariance descriptors can be used to define spatio-temporal blocks in video. (See Fig. 7.)

For flame detection:

$$R_{i,j,n} = \text{Red}(i, j, n), \quad (2)$$

$$G_{i,j,n} = \text{Green}(i, j, n), \quad (3)$$

$$B_{i,j,n} = \text{Blue}(i, j, n), \quad (4)$$

$$I_{i,j,n} = \text{Intensity}(i, j, n), \quad (5)$$

$$I_{xi,j,n} = \left| \frac{\partial \text{Intensity}(i, j, n)}{\partial i} \right|, \quad (6)$$

$$I_{yj,j,n} = \left| \frac{\partial \text{Intensity}(i, j, n)}{\partial j} \right|, \quad (7)$$

$$I_{xxi,j,n} = \left| \frac{\partial^2 \text{Intensity}(i, j, n)}{\partial i^2} \right|, \quad (8)$$

$$I_{yyi,j,n} = \left| \frac{\partial^2 \text{Intensity}(i, j, n)}{\partial j^2} \right|, \quad (9)$$

$$I_{ti,j,n} = \left| \frac{\partial \text{Intensity}(i, j, n)}{\partial n} \right|, \quad \text{and} \quad (10)$$

$$I_{tti,j,n} = \left| \frac{\partial^2 \text{Intensity}(i, j, n)}{\partial n^2} \right|. \quad (11)$$

For smoke detection:

$$Y_{i,j,n} = \text{Luminance}(i, j, n), \quad (12)$$

$$U_{i,j,n} = \text{Chrominance } U(i, j, n), \quad (13)$$

$$V_{i,j,n} = \text{Chrominance } V(i, j, n), \quad (14)$$

$$I_{i,j,n} = \text{Intensity}(i, j, n), \quad (15)$$

$$I_{xi,j,n} = \left| \frac{\partial \text{Intensity}(i, j, n)}{\partial i} \right|, \quad (16)$$

$$I_{yj,j,n} = \left| \frac{\partial \text{Intensity}(i, j, n)}{\partial j} \right|, \quad (17)$$

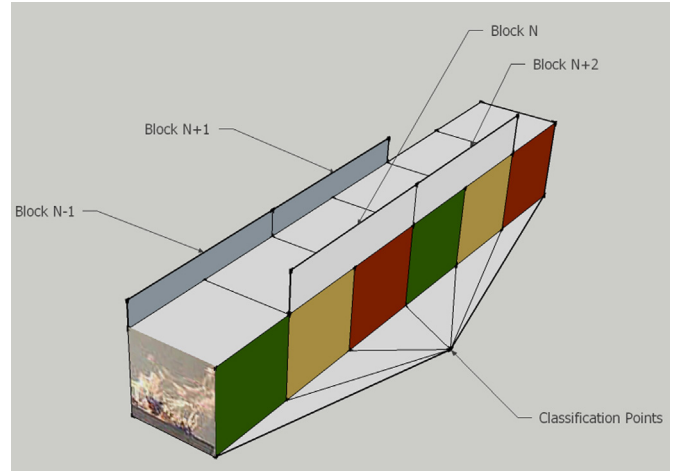


Fig. 7. An example for spatio-temporal block extraction and classification.

$$I_{xxi,j,n} = \left| \frac{\partial^2 \text{Intensity}(i, j, n)}{\partial i^2} \right|, \quad (18)$$

$$I_{yyi,j,n} = \left| \frac{\partial^2 \text{Intensity}(i, j, n)}{\partial j^2} \right|, \quad (19)$$

$$I_{ti,j,n} = \left| \frac{\partial \text{Intensity}(i, j, n)}{\partial n} \right|, \quad (20)$$

$$I_{tti,j,n} = \left| \frac{\partial^2 \text{Intensity}(i, j, n)}{\partial n^2} \right|. \quad (21)$$

Computation of normalized covariance values in spatio-temporal blocks.

The video is divided into blocks of size $10 \times 10 \times F_{rate}$ where F_{rate} is the frame rate of the video. Computing the normalized covariance parameters for each block of the video would be computationally inefficient. Therefore, only pixels corresponding to the non-zero values of the following mask are used in the selection of blocks. The mask is defined by the following function:

$$\Psi(i, j, n) = \begin{cases} 1 & \text{if } M(i, j, n) = 1, \\ 0 & \text{otherwise} \end{cases} \quad (22)$$

where $M(i, j, n)$ is the binary mask obtained from color detection and moving object detection algorithms. A total of 10 property parameters are used for each pixel satisfying the color condition (RGB version of the formula is used for flame detection). If we use all 10 property parameters we obtain $\frac{10 \times 11}{2} = 55$ correlation values. This means that the feature vector for each spatio-temporal block has 55 elements. To further reduce the computational cost, the normalized covariance values of the pixel property vectors

$$\Phi_{color}(i, j, n) = [Y(i, j, n) \quad U(i, j, n) \quad V(i, j, n)]^T \quad (23)$$

and

$$\Phi_{ST}(i, j, n) = \begin{bmatrix} I(i, j, n) \\ I_x(i, j, n) \\ I_y(i, j, n) \\ I_{xx}(i, j, n) \\ I_{yy}(i, j, n) \\ I_t(i, j, n) \\ I_{tt}(i, j, n) \end{bmatrix} \quad (24)$$

are computed separately. Therefore, the property vector $\Phi_{color}(i, j, n)$ produces $\frac{3 \times 4}{2} = 6$ and the property vector $\Phi_{ST}(i, j, n)$ produces $\frac{7 \times 8}{2} = 28$ correlation values, respectively and 34 correlation parameters are used in training and testing of the Support Vector Machine (SVM) instead of 55 parameters.

During the implementation of the correlation method, the first derivative of the image is computed by filtering the image with $[-1 \ 0 \ 1]$ and second derivative is found by filtering the image with $[1 \ -2 \ 1]$ filters, respectively. The lower or upper triangular parts of the matrix $\hat{C}(a, b)$ is obtained by normalizing the covariance matrix $\hat{S}(a, b)$ form the feature vector of a given image region:

$$\hat{S}(a, b) = \frac{1}{N-1} \left(\sum_i \sum_j \Phi_{i,j}(a) \Phi_{i,j}(b) - C_N \right) \quad (25)$$

where

$$C_N = \frac{1}{N} \left(\sum_i \sum_j \Phi_{i,j}(a) \right) \left(\sum_i \sum_j \Phi_{i,j}(b) \right), \quad (26)$$

$$\hat{C}(a, b) = \begin{cases} \sqrt{\hat{S}(a, b)} & \text{if } a = b, \\ \frac{\hat{S}(a, b)}{\sqrt{\hat{S}(a, a)} \sqrt{\hat{S}(b, b)}} & \text{otherwise.} \end{cases} \quad (27)$$

Entries of $\hat{C}(a, b)$ matrix are processed by a Support Vector Machine which had been previously trained with fire and smoke video clips.

In order to improve the detection performance, the majority of the articles in the literature use a combination of the fire feature extraction methods described above. Depending on the fire/environmental characteristics, one combination of features will outperform the other and vice versa. In Section 4, we describe an adaptive fusion method combining the results of various fire detection methods in an online manner.

It should be pointed out that articles in the literature and those which are referenced in this state-of-the-art review indicate that ordinary visible range camera based detection systems promise good fire detection results. However, they still suffer from a significant amount of missed detections and false alarms in practical situations as in other computer vision problems [5,6]. The main cause of these problems is the fact that visual detection is often subject to constraints regarding the scene under investigation, e.g., changing environmental conditions, different camera parameters and color settings and illumination. It is also impossible to compare the articles with each other and determine the best one. This is because they use different training and data sets.

A data set of fire and non-fire videos is available to the research community in European Commission funded FIRESENSE project web-page [56]. These test videos were used for training and testing purposes of the smoke and flame detection algorithms developed within the FIRESENSE project. Thus, a fair comparison of the algorithms developed by individual partners could be conducted. The test database includes 27 test and 29 training sequences of visible spectrum recordings of flame scenes, 15 test and 27 training sequences of visible spectrum recordings of smoke scenes, and 22 test and 27 training sequences of visible spectrum recordings of forest smoke scenes. This database is currently available to registered users of the FIRESENSE website [Reference: FIRESENSE project File Repository, <http://www.firesense.eu>, 2012].

2.7. Classification techniques

A popular approach for the classification of the multi-dimensional feature vectors obtained from each candidate flame or smoke blob is SVM classification, typically with Radial Basis Function (RBF) kernels. A large number of frames of fire and non-fire video sequences need to be used for training these SVM classifiers, otherwise the number of false alarms (false positives or true negatives) may be significantly increased.

Other classification methods include the AdaBoost method [22], neural networks [29,35], Bayesian classifiers [30,32], Markov models [28,33] and rule-based classification [58].

As in any video processing method, morphological operations, subblocking and clean-up post-processing such as median-filtering are used as an integral part of any VFD system [21,22,25,20,26,33,36,59].

2.8. Evaluation of visible range video fire detection methods

An evaluation of different visible range video fire detection methods is presented in Table 4. Table 4 summarizes comparative detection results for the smoke and flame detection algorithm by Verstockt [2] (Method 1), a combination of the flame detection method by Celik et al. [60] and the smoke detection by Toreyin et al. [14] (Method 2) and a combination of the feature-based flame detection method by Borges et al. [23] and the smoke detection method by Xiong et al. [18] (Method 3). Among various algorithms, Verstockt's method is a relatively recent one whereas flame detection methods by Celik and Borges and the smoke detection methods by Toreyin and Xiong are commonly referenced methods in the literature.

Test sequences used for performance evaluation are captured in different environments under various conditions. Snapshots from test videos are presented in Fig. 8. In order to objectively evaluate the detection results of different methods, the 'detection rate' metric [61,2] is used which is comparable to the evaluation methods used by Celik et al. [60] and Toreyin et al. [13]. The detection rate equals the ratio of the number of correctly detected frames as fire, i.e., the detected frames as fire minus the number of falsely detected frames, to the number of frames with fire in the manually created ground truth frames. As results indicate, the detection performances of different methods are comparable with each other.

3. Video fire detection in infrared (IR) spectral range

When there is no or very little visible light or the color of the object to be detected is similar to the background, IR imaging systems provide solutions [62–68]. Although there is an increasing trend in IR-camera based intelligent video analysis, the number of papers in the area of IR-based fire detection is few [64–68]. This is mainly due to the high cost of IR imaging systems compared to ordinary cameras. Manufacturers predict that IR camera prices will go down in the near future. Therefore, we expect that the number of IR imaging applications will increase significantly [63]. Long-Wave Infrared (8–12 micron range) cameras are the most widely available cameras in the market. Long-Wave Infrared (LWIR) light goes through smoke therefore it is easy to detect smoke using LWIR imaging systems. Nevertheless, results from existing work already ensure the feasibility of IR cameras for flame detection.

Owrutsky et al. [64] worked in the near infrared (NIR) spectral range and compared the global luminosity L , which is the sum of the pixel intensities of the current frame, to a reference luminosity L_b and a threshold L_{th} . If there are a number of consecutive frames where L exceeds the persistence criterion $L_b + L_{th}$, the system goes into an alarm stage. Although this fairly simple algorithm seems to produce good results in the reported experiments, its limited constraints do raise questions about its applicability in large and open uncontrolled public places and it will probably produce many false alarms to hot moving objects such as cars and human beings. Although the cost of NIR cameras are not high, their imaging ranges are shorter compared to visible range cameras and other IR cameras.

Toreyin et al. [65] detect flames in LWIR by searching for bright-looking moving objects with rapid time-varying contours. A wavelet domain analysis of the 1D-curve representation of the contours is used to detect the high frequency nature of the boundary of a fire region. In addition, the temporal behavior of the region

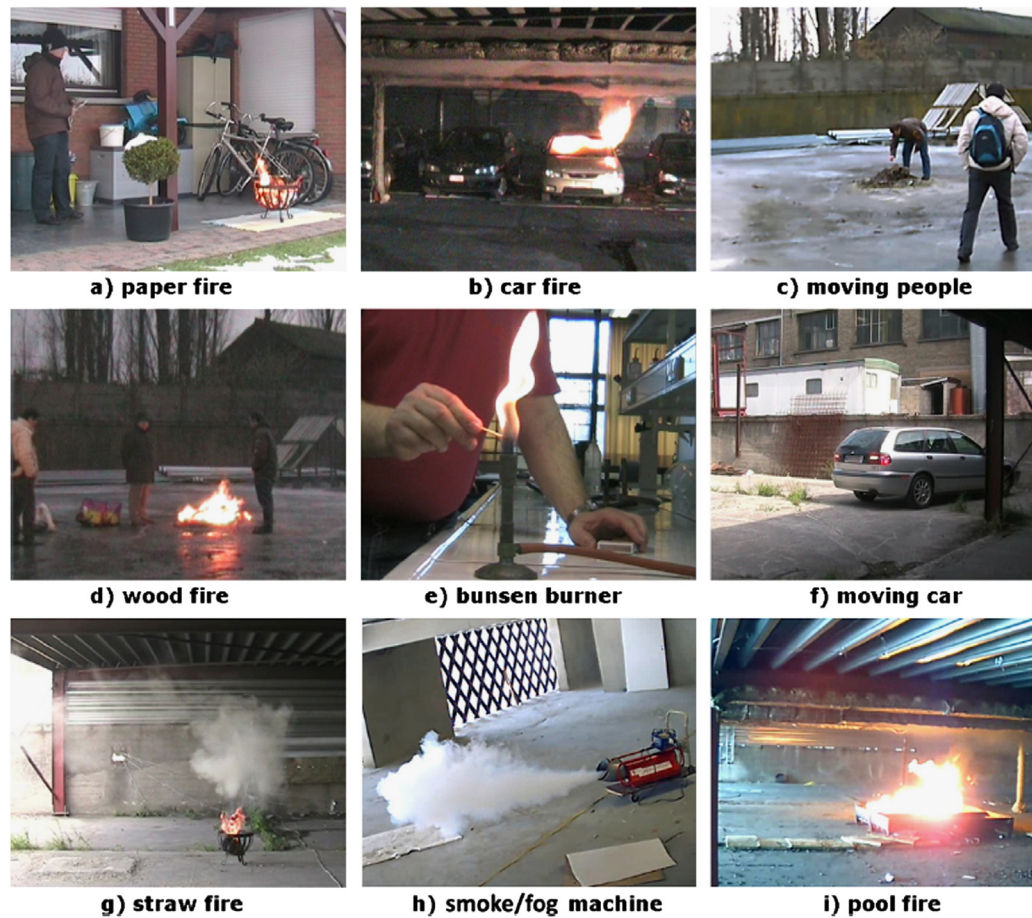


Fig. 8. Snapshots from test sequences with and without fire.

Table 4
Comparison of the smoke and flame detection method by Verstockt [2] (Method 1), the combined method based on the flame detector by Celik et al. [60] and the smoke detector described in Toreyin et al. [14] (Method 2), and combination of the feature-based flame detection method by Borges et al. [23] and the smoke detection method by Xiong et al. [18] (Method 3).

Video sequence (# frames)	# Fire frames ground truth	# Detected fire frames Method			# False positive frames Method			Detection rate*		
		1	2	3	1	2	3	1	2	3
Paper fire (1550)	956	897	922	874	9	17	22	0.93	0.95	0.89
Car fire (2043)	1415	1293	1224	1037	3	8	13	0.91	0.86	0.73
Moving people (886)	0	5	0	28	5	0	28	–	–	–
Wood fire (592)	522	510	489	504	17	9	16	0.94	0.92	0.93
Bunsen burner (115)	98	59	53	32	0	0	0	0.60	0.54	0.34
Moving car (332)	0	0	13	11	0	13	11	–	–	–
Straw fire (938)	721	679	698	673	16	21	12	0.92	0.93	0.92
Smoke/fog machine (1733)	923	834	654	789	9	34	52	0.89	0.67	0.80
Pool fire (2260)	1844	1665	1634	1618	0	0	0	0.90	0.89	0.88

* Detection rate = (# detected fire frames – # false alarms) / # fire frames.

is analyzed using a Hidden Markov Model (HMM). The combination of both spatial and temporal clues seems more appropriate than the luminosity approach and, according to the authors, their approach greatly reduces false alarms caused by ordinary bright moving objects. A similar combination of temporal and spatial features is also used by Bosch et al. [66]. Hotspots, i.e., candidate flame regions, are detected by automatic histogram-based image thresholding. By analyzing the intensity, signature, and orientation of these resulting hot objects' regions, discrimination between flames and other objects is made. Verstockt [2] also proposed an IR-based fire detector which mainly follows the latter feature-based strategy, but contrary to Bosch et al.'s work [66] a dynamic background subtraction method is used, which aims at coping with the time-varying characteristics of dynamic scenes.

To sum up, it should be pointed out that it is not straightforward to detect fires using IR cameras. Not every bright object in IR video is a source of wildfire. It is important to mention that IR imaging has its own specific limitations, such as thermal reflections, IR blocking and thermal-distance problems. In some situations, IR-based detection will perform better than visible VFD, but under other circumstances, visible VFD can improve IR flame detection. This is due to the fact that, smoke appears earlier and becomes visible from long distances in a typical uncontrolled fire. Flames and burning objects may not be in the viewing range of the IR camera. As such, higher detection accuracies with lower false alarm rates can be achieved by combining multi-spectrum video information. Various image fusion methods may be employed for this purpose [69,70]. Clearly, each sensor type has its own specific



Fig. 9. Snapshot of typical wildfire smoke captured by a forest watch tower which is 5 km away from the fire (rising smoke is marked with an arrow).

limitations, which can only be compensated by other types of sensors.

4. Wildfire smoke detection using visible range cameras

As pointed out in the previous section, smoke is clearly visible from long distances in wildfires and forest fires. In most cases flames are hindered by trees. Therefore, IR imaging systems may not provide solutions for early fire detection in wildfires but ordinary visible range cameras can detect smoke from long distances. (See Fig. 9.)

Smoke at far distances (>100 m to the camera) exhibits different spatio-temporal characteristics than nearby smoke and fire [71, 59,13]. This demands specific methods explicitly developed for smoke detection at far distances rather than using nearby smoke detection methods described in [72]. Cetin et al. proposed wildfire smoke detection algorithms consisting of five main sub-algorithms: (i) slow moving object detection in video, (ii) smoke-colored region detection, (iii) wavelet transform based region smoothness detection, (iv) shadow detection and elimination, (v) covariance matrix based classification, with individual decision functions, $D_1(x, n)$, $D_2(x, n)$, $D_3(x, n)$, $D_4(x, n)$ and $D_5(x, n)$, respectively, for each pixel at location x of every incoming image frame at time step n . Decision results of individual algorithms are fused to obtain a reliable wildfire detection system in [67,37].

The video based wildfire detection system described in this section has been deployed in more than 100 forest look out towers in the world including Turkey, Italy and the US. The system is not fully automatic because forestal scenes vary over time due to weather conditions and changes in illumination. The system is developed to help security guards in look out towers. It is not feasible to develop one strong fusion model with fixed weights in forestal setting which has a time-varying (drifting) nature. An ideal online active learning mechanism should keep track of drifts in video and adapt itself accordingly. Therefore in Cetin et al.'s system, decision functions are combined in a linear manner and the weights are determined according to the weight update mechanism described in the next subsection.

Decision functions D_i , $i = 1, \dots, M$, of sub-algorithms do not produce binary values 1 (correct) or -1 (false), but they produce real numbers centered around zero for each incoming sample x . Output values of decision functions express the confidence level of

each sub-algorithm. The higher the value, the more confident the algorithm.

Morphological operations are applied to the detected pixels to mark the smoke regions. The number of connected smoke pixels should be larger than a threshold to issue an alarm for the region. If a false alarm is issued during the training phase, the oracle gives feedback to the algorithm by declaring a no-smoke decision value ($y = -1$) for the false alarm region. Initially, equal weights are assigned to each sub-algorithm. There may be large variations between forestal areas and substantial temporal changes may occur within the same forestal region. As a result, weights of individual sub-algorithms will evolve in a dynamic manner over time. In Fig. 10, the flowchart of the weight update algorithm is given for one image frame.

4.1. Adaptive Decision Fusion (ADF) framework

Let the compound algorithm be composed of M -many detection sub-algorithms: D_1, \dots, D_M . Upon receiving a sample input x at time step n , each sub-algorithm yields a decision value $D_i(x, n) \in \mathbb{R}$ centered around zero. If $D_i(x, n) > 0$, it means that the event is detected by the i th sub-algorithm.

Let $\mathbf{D}(x, n) = [D_1(x, n), \dots, D_M(x, n)]^T$ be the vector of decision values of the sub-algorithms for the pixel at location x of input image frame at time step n , and $\mathbf{w}(x, n) = [w_1(x, n), \dots, w_M(x, n)]^T$ be the current weight vector.

4.1.1. Entropic projection (e-projection) based weight update algorithm

In this subsection, we review the entropic projection based weight update scheme [73,37,67]. The e-projection onto a closed and convex set is a generalized version of the metric projection mapping onto a convex set [74]. Let $\mathbf{w}(n)$ denote the weight vector for the n th sample. Its' e-projection \mathbf{w}^* onto a closed convex set C with respect to a cost functional $g(\mathbf{w})$ is defined as follows:

$$\mathbf{w}^* = \arg \min_{\mathbf{w} \in C} L(\mathbf{w}, \mathbf{w}(n)) \quad (28)$$

where

$$L(\mathbf{w}, \mathbf{w}(n)) = g(\mathbf{w}) - g(\mathbf{w}(n)) - \langle \nabla g(\mathbf{w}), \mathbf{w} - \mathbf{w}(n) \rangle \quad (29)$$

and $\langle \cdot, \cdot \rangle$ represents the inner product.

In the adaptive learning problem, we have a hyperplane $H(x, n)$: $\mathbf{D}^T(x, n) \cdot \mathbf{w}(n+1) = y(x, n)$ for each sample x . For each hyperplane $H(x, n)$, the e-projection (28) is equivalent to

$$\nabla g(\mathbf{w}(n+1)) = \nabla g(\mathbf{w}(n)) + \lambda \mathbf{D}(x, n), \quad (30)$$

$$\mathbf{D}^T(x, n) \cdot \mathbf{w}(n+1) = y(x, n) \quad (31)$$

where λ is the Lagrange multiplier. As pointed out above, the e-projection is a generalization of the metric projection mapping.

When the cost functional is the entropy functional $g(\mathbf{w}) = \sum_i w_i(n) \log(w_i(n))$, the e-projection onto the hyperplane $H(x, n)$ leads to the following update equations:

$$w_i(n+1) = w_i(n) e^{\lambda D_i(x, n)}, \quad i = 1, 2, \dots, M, \quad (32)$$

where the Lagrange multiplier λ is obtained by inserting (32) into the hyperplane equation:

$$\mathbf{D}^T(x, n) \mathbf{w}(n+1) = y(x, n) \quad (33)$$

because the e-projection $\mathbf{w}(n+1)$ must be on the hyperplane $H(x, n)$ in Eq. (31). When there are three hyperplanes, one cycle of the projection algorithm is depicted in Fig. 11. If the projections are continued in a cyclic manner the weights will converge to the intersection of the hyperplanes, w_c .

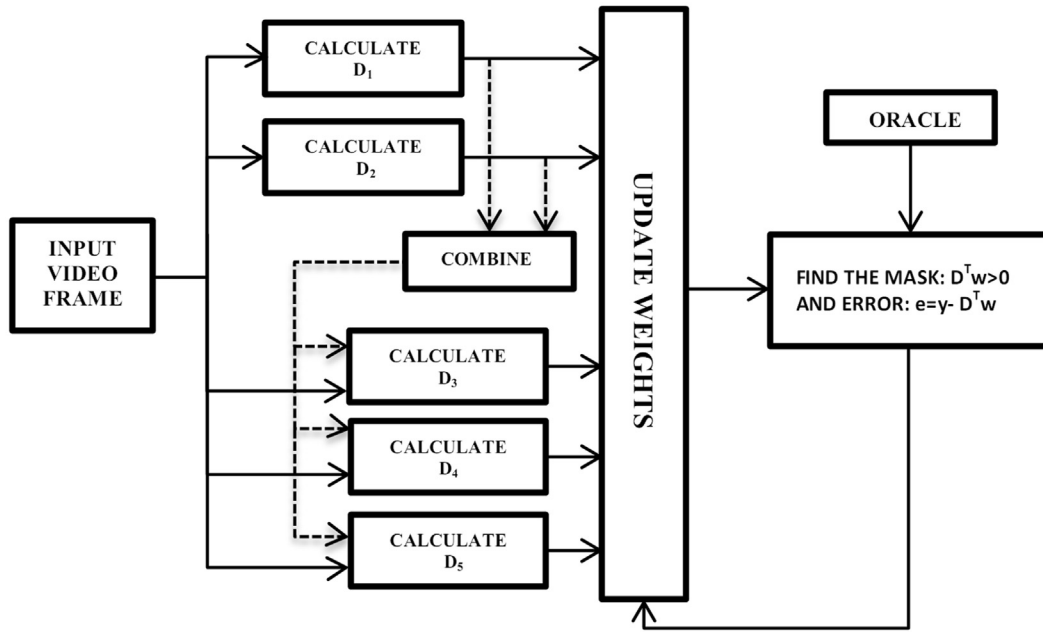


Fig. 10. Flowchart of the weight update algorithm for one image frame.

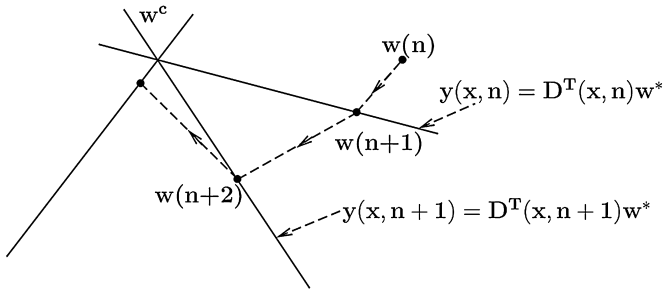


Fig. 11. Geometric interpretation of the entropic-projection method: weight vectors corresponding to decision functions at each frame are updated to satisfy the hyperplane equations defined by the oracle's decision $y(x, n)$ and the decision vector $D(x, n)$.

It is desirable that each sub-algorithm should contribute to the compound algorithm because each characterizes a feature of wildfire smoke for the wildfire detection problem. Therefore weights of algorithms can be set between 0 and 1 representing the contribution of each feature. We want to penalize extreme weight values 0 and 1 more compared to values in between them, because each sub-algorithm is considered to be “weak” compared to the final algorithm. The entropy functional achieves this. On the other hand, the commonly used Euclidean norm penalizes high weight values more compared to zero weight.

In real-time operating mode, the PTZ cameras are in continuous scan mode visiting predefined preset locations. In this mode, constant monitoring from the oracle can be relaxed by adjusting the weights for each preset once, and then use the same weights for successive classifications. Since the main issue is to reduce false alarms, the weights can be updated when there is no smoke in the viewing range of each preset and after that, the system becomes autonomous. The cameras stop at each preset and run the detection algorithm for some time before moving to the next preset.

4.2. Fire and smoke detection criteria

In VFD, cameras are used for fire detection. In many cases there will be a large distance between the PTZ camera and the wildfire. Therefore, it is important to define when the wildfire is visible by

the camera. For this purpose, we propose Johnson's criteria used in the infrared camera literature [75].

Johnson's criteria are about “seeing a target” in an infrared camera. The first criterion defines detection: In order to detect an object its critical dimension needs to be covered by 1.5 or more pixels in the captured image. Therefore the wildfire is detectable when it occupies more than 1.5 pixels in video image. This is the ultimate limit. One or two pixels can be easily confused with noise. In Fig. 12, minimum smoke size versus detection range is shown for wildfire smoke using a visible range camera.

Curvature of the earth also affects the detection range. For ranges above 20 km, smoke should rise even higher to compensate for the earth's curvature. A sketch depicting a wildfire smoke detection scenario from a camera placed on top of a 40-m-mast is presented in Fig. 13. At a distance of 40 km, a 40 m × 40 m smoke has to rise an additional 20 m to be detected by the camera on top of the mast.

The second criterion defines recognition which means that it is possible to make the distinction between a person, a car, a truck or wildfire. In order to recognize an object it needs to occupy at least 6 pixels across its critical dimension in a video image. The third criterion defines identification. This term relates to the military terminology. The critical dimension of the object should be at least 12 pixels so that the object is identified as “friend or foe”. We can use the Johnson's identification criterion for wildfire identification because the white smoke may be due to a dirt cloud from an off-road vehicle or may be a cloud or it may be fog rising above the trees.

These criteria applied to IR-camera based wildfire flame detection are summarized in the following figure. In Fig. 14, minimum flame sizes versus varying line-of-sight ranges are shown for detection, recognition and identification using an MWIR InP camera with a spectral range of 3–5 μm . Note that, a minimum flame size of 1 m^2 is enough to identify a wildfire at a range of 1.8 km and it is enough to recognize it at a range of 2.7 km. However, at a distance of 11 km, one can only detect the same fire.

5. Wildfire smoke detection using IR camera

The smoke of a wildfire can be detected using a visible range camera as explained in the previous section (cf. Fig. 15). On the

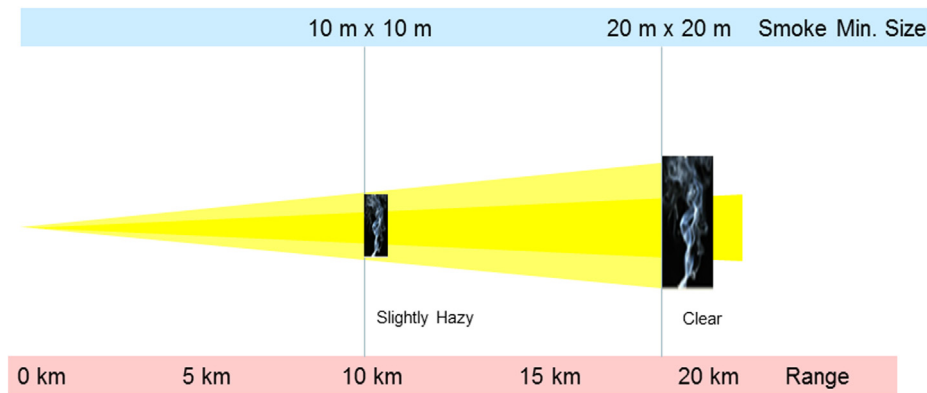


Fig. 12. Minimum smoke size versus detection distance with a visible range camera.

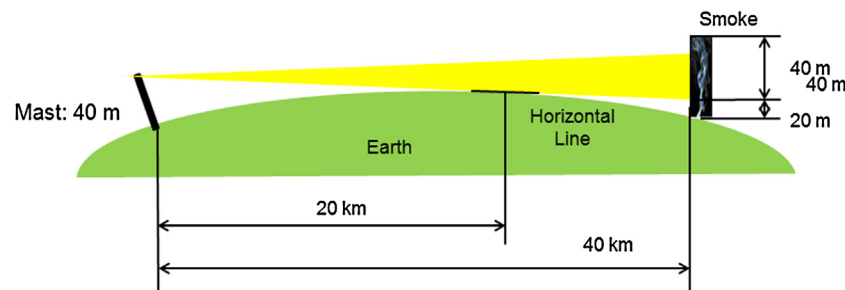


Fig. 13. Effect of the earth's curvature on wildfire smoke detection range.

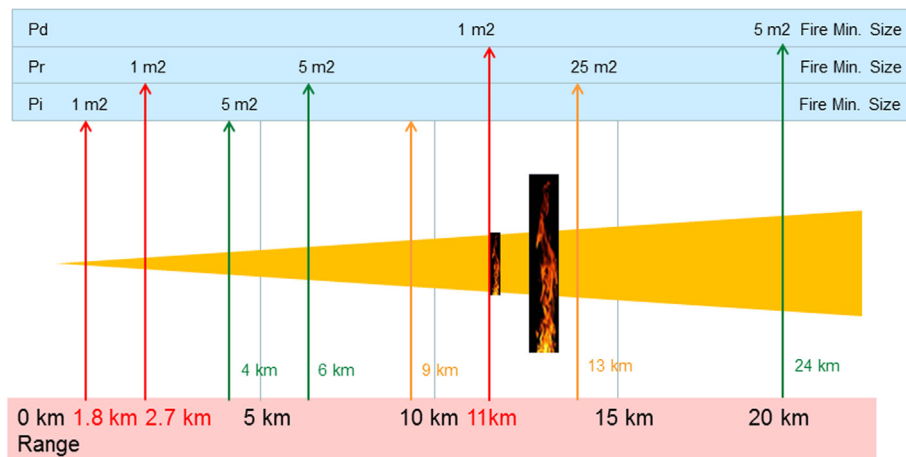


Fig. 14. Minimum fire sizes versus detection, recognition and identification ranges.

other hand, wildfire smoke detection using an ordinary LWIR camera with spectral range 8–12 μm is very difficult as smoke is invisible (cf. Fig. 16). This is evident from the snapshots below corresponding to tests using both LWIR and visible range cameras.

Wildfire flame detection is possible using an IR camera (cf. Fig. 16). However, in most wildfires, smoke appears first. Therefore, it is not necessary to employ an expensive IR camera to spot flames which may be occluded by tree trunks.

On the other hand, LWIR cameras are extremely useful to pinpoint hotspots and flames from a fire extinguishing helicopter because they have the capability to see through smoke. Therefore, they are invaluable tools during wildfire fighting.

Smoke detection can be made possible by joint analysis of visible and IR range camera outputs, as well. A method for smoke detection exploiting both modalities is proposed in [2]. The two-step method, called LWIR-visual smoke detection, comprises the silhou-

ette coverage analysis step and disorder analysis of visual silhouette step. The LWIR-visual smoke detector analyzes the silhouette coverage of moving objects in visible range and long-wave infrared images. Moreover, the flicker in the silhouette of visual smoke regions is taken into account as an additional clue for smoke presence within the viewing range of the camera [14].

A comparison of performances of the LWIR-visual smoke detector of Verstockt [2] and visible range camera based smoke detectors of Xiong et al. [18], Calderara et al. [20] and Toreyin et al. [14] is presented in Table 5. Among these methods, Xiong et al. [18] proposed a visible range smoke detection algorithm based on background subtraction and flicker/disorder analysis. Calderara et al.'s method is based on mixture of Gaussians (MoG) of discrete wavelet transform energy variations and color blending [20]. On the other hand, smoke detection method developed by Toreyin et al. [14] is based on block-based spatial wavelet analysis and HMM.

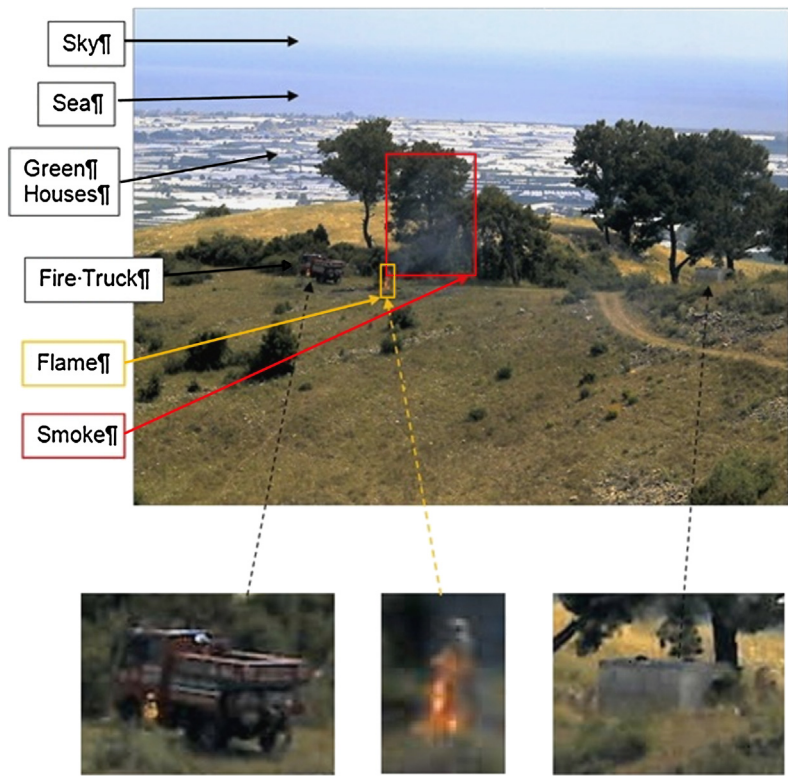


Fig. 15. Wildfire smoke and fire captured using a visible range camera.

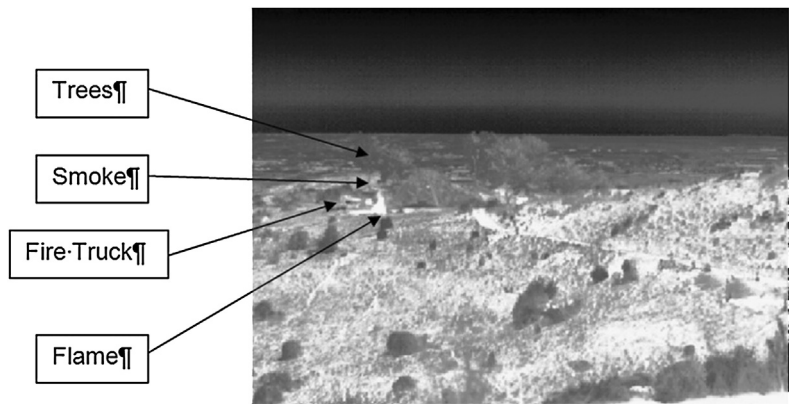


Fig. 16. Wildfire smoke and fire in Fig. 15 captured using an LWIR camera.

Five different setups are formed for test purposes: car fire, straw fire, moving people, moving car, and paper fire. For each of these fire and non-fire test setups, several video sequences are generated. The test set contains 18 multi-modal fire and 13 non-fire video sequences under different lighting conditions. For each of these sequences, ground truth fire decisions are generated manually. Comparative performance results in terms of alarm percentages with respect to ground truth are presented in Table 5.

The method which exploits both visible and IR range data, namely the LWIR-visual smoke detector, yields best detection/false alarm ratio among other methods which only analyze data from the visible range of the spectrum. The detection performance of visible range smoke detectors degrades substantially for the sequences where the scene is not well lit yielding a poor visibility for smoke. On the contrary, the multi-modal method takes advantage of silhouette coverage analysis step which compares the silhouettes covered in both LWIR and visible range images. This

Table 5
A comparison of performances of the LWIR-visual smoke detector of Verstockt [2] and visible range camera based smoke detectors of Xiong et al. [18], Calderara et al. [20] and Toreyin et al. [14].

Method	Fire alarm	
	Fire tests (%)	Non-fire tests (%)
Xiong et al.	68	7
Calderara et al.	83	4
Toreyin et al.	86	4
Verstockt	98	2

analysis step helps to result in a better detection performance for the multi-modal method.

6. Conclusion

The concept of artificial intelligence and artificial systems capable of perceiving their environment and taking necessary actions

was introduced 68 years ago in 1955 by John McCarthy. There has been significant progress in some applications such as restricted speech recognition, character recognition, chess and game playing, etc. On the other hand, there is very little progress even in some simple recognition problems. Humans can easily recognize uncontrolled fire whenever they see it even from long distances. Computers cannot imitate human intelligence whose operation principles are not clearly understood as of today. It is possible to achieve significant progress when an intelligent task is described in terms of mathematical and physical terms. This is usually a slow and difficult process but it has produced results in the past as in speech recognition. We believe that VFD falls into this category of problems that can be reduced to an engineering description.

During the last two decades video fire detection research led to commercial and experimental VFD systems. They are especially used in high risk buildings and areas. Hundreds of video smoke detection systems are installed in lookout towers and poles for wildfire detection. Although they are not fully automated systems they are invaluable tools for security personnel. Whenever the VFD system produces an alarm, the operator can check if it is a real fire or a false alarm. In this way the operator can handle multiple cameras.

We believe that further research will improve the detection performance of the VFD systems while reducing the false alarm rates. We also believe that multi-modal systems employing IR and regular cameras will provide fully reliable VFD solutions. Such systems are currently expensive, but they will be feasible with the decline of IR camera costs in the near future.

Acknowledgments

This work was supported in part by FIRESENSE (Fire Detection and Management through a Multi-Sensor Network for the Protection of Cultural Heritage Areas from the Risk of Fire and Extreme Weather Conditions, FP7-ENV-2009-1244088-FIRESENSE) and by The Scientific and Technological Research Council of Turkey (TUBITAK) under grant EEEAG 111E217.

References

- [1] J. Hogan, Your every move will be analysed, *New Scientist* 2403 (July 2003) 4–5.
- [2] S. Verstockt, Multi-modal video analysis for early fire detection, PhD thesis, Universiteit Gent, 2011.
- [3] H.L. Dreyfus, *What Computers Can't Do*, MIT Press, 1972.
- [4] H.L. Dreyfus, *What Computers Still Can't Do*, MIT Press, 1992.
- [5] T. Pavlidis, Computers versus humans, <http://www.theopavlidis.com/comphumans/comphuman.htm>.
- [6] T. Pavlidis, Why machine intelligence is so hard – A discussion of the difference between human and computer intelligence, <http://www.theopavlidis.com/technology/MachineIntel/index.htm>.
- [7] W. Phillips, M. Shah, N. da Vitoria Lobo, Flame recognition in video, *Pattern Recogn. Lett.* 23 (1–3) (January 2002) 319–327.
- [8] F. Gomez-Rodriguez, S. Pascual-Pena, B. Arrue, A. Ollero, Smoke detection using image processing, in: *Proceedings of 4th International Conference on Forest Fire Research & Wildland Fire Safety*, November 2002, pp. 1–8.
- [9] F. Gomez-Rodriguez, B.C. Arrue, A. Ollero, Smoke monitoring and measurement using image processing – Application to forest fires, in: *Proceedings of SPIE AeroSense 2003: XIII Automatic Target Recognition*, April 2003, pp. 404–411.
- [10] T.-H. Chen, P.-H. Wu, Y.-C. Chiou, An early fire-detection method based on image processing, in: *Proceedings of IEEE International Conference on Image Processing (ICIP)*, vol. 3, October 2004, pp. 1707–1710.
- [11] C.B. Liu, N. Ahuja, Vision based fire detection, in: *Proceedings of 17th International Conference on Pattern Recognition (ICPR)*, vol. 4, August 2004, pp. 134–137.
- [12] G. Marbach, M. Loepfe, T. Brupbacher, An image processing technique for fire detection in video images, *Fire Saf. J.* 41 (4) (June 2006) 285–289.
- [13] B.U. Toreyin, Y. Dedeoglu, U. Gudukbay, A.E. Cetin, Computer vision based method for real-time fire and flame detection, *Pattern Recogn. Lett.* 27 (1) (January 2006) 49–58.
- [14] B.U. Toreyin, Y. Dedeoglu, A.E. Cetin, Contour based smoke detection in video using wavelets, in: *Proceedings of European Signal Processing Conference (EUSIPCO)*, September 2006.
- [15] T. Celik, H. Ozkaramanli, H. Demirel, Fire and smoke detection without sensors: Image processing based approach, in: *Proceedings of 15th European Signal Processing Conference (EUSIPCO)*, September 2007, pp. 1794–1798.
- [16] Z. Xu, J. Xu, Automatic fire smoke detection based on image visual features, in: *Proceedings of International Conference on Computational Intelligence and Security Workshops*, December 2007, pp. 316–319.
- [17] T. Celik, H. Demirel, H. Ozkaramanli, M. Uyguroglu, Fire detection using statistical color model in video sequences, *J. Vis. Commun. Image Represent.* 18 (2) (January 2007) 176–185.
- [18] Z. Xiong, R. Caballero, H. Wang, A.M. Finn, M.A. Lelic, P.-Y. Peng, Video-based smoke detection: Possibilities, techniques, and challenges, in: *Proceedings of Suppression and Detection Research and Applications (SUPDET) – A Technical Working Conference*, 2007.
- [19] B. Lee, D. Han, Real-time fire detection using camera sequence image in tunnel environment, in: *Proceedings of International Conference on Intelligent Computing*, August 2007, pp. 1209–1220.
- [20] S. Calderara, P. Piccinini, V. Cucchiara, Smoke detection in video surveillance: A mog model in the wavelet domain, in: *Proceedings of 6th International Conference in Computer Vision Systems (ICVS)*, May 2008, pp. 119–128.
- [21] P. Piccinini, S. Calderara, R. Cucchiara, Reliable smoke detection system in the domains of image energy and color, in: *Proceedings of International Conference on Image Processing*, October 2008, pp. 1376–1379.
- [22] F. Yuan, A fast accumulative motion orientation model based on integral image for video smoke detection, *Pattern Recogn. Lett.* 29 (7) (May 2008) 925–932.
- [23] P.V.K. Borges, J. Mayer, E. Izquierdo, Efficient visual fire detection applied for video retrieval, in: *Proceedings of 16th European Signal Processing Conference (EUSIPCO)*, August 2008.
- [24] X. Qi, J. Ebert, A computer vision based method for fire detection in color videos, *Int. J. Imag.* 2 (S09) (Spring 2009) 22–34.
- [25] R. Yasmin, Detection of smoke propagation direction using color video sequences, *Int. J. Soft Comput.* 4 (1) (2009) 45–48.
- [26] J. Gubbi, S. Marusic, M. Palaniswami, Smoke detection in video using wavelets and support vector machines, *Fire Saf. J.* 44 (8) (November 2009) 1110–1115.
- [27] J. Chen, Y. He, J. Wang, Multi-feature fusion based fast video flame detection, *Build. Environ.* 45 (5) (May 2010) 1113–1122.
- [28] O. Gunay, K. Tasdemir, B.U. Toreyin, A.E. Cetin, Fire detection in video using lms based active learning, *Fire Technol.* 46 (3) (2010) 551–577.
- [29] I. Kolesov, P. Karasev, A. Tannenbaum, E. Haber, Fire and smoke detection in video with optimal mass transport based optical flow and neural networks, in: *Proceedings of IEEE International Conference on Image Processing (ICIP)*, September 2010, pp. 761–764.
- [30] B.C. Ko, K.H. Cheong, J.Y. Nam, Early fire detection algorithm based on irregular patterns of flames and hierarchical Bayesian networks, *Fire Saf. J.* 45 (4) (June 2010) 262–270.
- [31] R.A. Gonzalez-Gonzalez, V. Alarcon-Aquino, O. Starostenko, R. Rosas-Romero, J.M. Ramirez-Cortes, J. Rodriguez-Asomoza, Wavelet-based smoke detection in outdoor video sequences, in: *Proceedings of the 53rd IEEE Midwest Symposium on Circuits and Systems (MWSCAS)*, August 2010, pp. 383–387.
- [32] P.V.K. Borges, E. Izquierdo, A probabilistic approach for vision-based fire detection in videos, *IEEE Trans. Circuits Syst. Video Technol.* 20 (5) (May 2010) 721–731.
- [33] D. Van Hamme, P. Veelaert, W. Philips, K. Teelen, Fire detection in color images using Markov random fields, in: *Proceedings of Advanced Concepts for Intelligent Vision Systems (ACIVS)*, vol. 2, December 2010, pp. 88–97.
- [34] T. Celik, Fast and efficient method for fire detection using image processing, *ETRI J.* 32 (6) (December 2010) 881–890.
- [35] F. Yuan, Video-based smoke detection with histogram sequence of lbp and lbpyramids, *Fire Saf. J.* 46 (3) (April 2011) 132–139.
- [36] L. Rossi, M. Akhloufi, Y. Tison, On the use of stereo vision to develop a novel instrumentation system to extract geometric fire fronts characteristics, *Fire Saf. J. (Forest Fires)* 46 (1–2) (January 2011) 9–20.
- [37] B.U. Toreyin, Fire detection algorithms using multimodal signal and image analysis, PhD thesis, Bilkent University, 2009.
- [38] R.T. Collins, A.J. Lipton, T. Kanade, A system for video surveillance and monitoring, in: *Proceedings of American Nuclear Society (ANS) Eighth International Topical Meeting on Robotics and Remote Systems*, 1999.
- [39] S. Calderara, P. Piccinini, R. Cucchiara, Vision based smoke detection system using image energy and color information, *Mach. Vis. Appl.* (May 2010) 1–15.
- [40] M. Wendeker, J. Czarnigowski, G. Litak, K. Szabelski, Chaotic combustion in spark ignition engines, *Chaos Solitons Fractals* 18 (2003) 803–806.
- [41] P. Manneville, *Instabilities, Chaos and Turbulence*, Imperial College Press, 2004.
- [42] C.K. Law, *Combustion Physics*, Cambridge University Press, 2006.
- [43] V. Gubernov, A. Kolobov, A. Polezhaev, H. Sidhu, G. Mercer, Period doubling and chaotic transient in a model of chain-branching combustion wave propagation, *Proc. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci.* 466 (2121) (2010) 2747–2769.
- [44] J. Yang, R.S. Wang, A survey on fire detection and application based on video image analysis, *Video Eng.* 2006 (8) (2006) 92–96.
- [45] G. Doretto, A. Chiou, Y.N. Wu, S. Soatto, Dynamic textures, *Int. J. Comput. Vis.* 51 (2) (2003) 91–109.
- [46] F. Porikli, A.E. Cetin, Special issue on dynamic textures in video, *Mach. Vis. Appl.* 22 (5) (2011) 739–740.

- [47] T. Amiaz, S. Fazekas, D. Chetverikov, N. Kiryati, Detecting regions of dynamic texture, in: Proceedings of International Conference on Scale-Space and Variational Methods in Computer Vision (SSVM), May 2007, pp. 848–859.
- [48] D. Chetverikov, R. Peteri, A brief survey of dynamic texture description and recognition, in: Proceedings of 4th International Conference on Computer Recognition Systems, May 2005, pp. 17–26.
- [49] B.U. Toreyin, Y. Dedeoglu, A.E. Cetin, S. Fazekas, D. Chetverikov, T. Amiaz, N. Kiryati, Dynamic texture detection, segmentation and analysis, in: Proceedings of ACM International Conference on Image and Video Retrieval (CIVR), July 2007, pp. 131–134.
- [50] S. Fazekas, T. Amiaz, D. Chetverikov, N. Kiryati, Dynamic texture detection based on motion analysis, *Int. J. Comput. Vis.* 82 (1) (April 2009) 48–63.
- [51] Renaud Péteri, Sándor Fazekas, Mark J. Huiskes, DynTex: A comprehensive database of dynamic textures, *Pattern Recogn. Lett.* 31 (2010) 1627–1632, <http://dx.doi.org/10.1016/j.patrec.2010.05.009>.
- [52] S. Fazekas, D. Chetverikov, Analysis and performance evaluation of optical flow features for dynamic texture recognition, *Signal Process., Image Commun.* 22 (7–8) (August 2007) 680–691, Special Issue on Content-Based Multimedia Indexing and Retrieval.
- [53] A. Enis Cetin, Computer vision based fire detection software – Sample video clips, <http://signal.ee.bilkent.edu.tr/VisiFire/>.
- [54] O. Tuzel, F. Porikli, P. Meer, Region covariance: A fast descriptor for detection and classification, in: *Computer Vision–ECCV*, 2006, pp. 589–600.
- [55] H. Tuna, I. Onaran, A. Enis Cetin, Image description using a multiplier-less operator, *IEEE Signal Process. Lett.* 16 (9) (September 2009) 751–753.
- [56] FIRESENSE, fire detection and management through a multi-sensor network for the protection of cultural heritage areas from the risk of fire and extreme weather conditions, fp7-env-2009-1-244088-firesense, <http://www.firesense.eu>, 2009.
- [57] Y. Habiboglu, O. Gunay, A.E. Cetin, Flame detection method in video using covariance descriptors, in: 2011 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 2011, pp. 1817–1820.
- [58] Fire Safety Engineering (Bart Merci), Fire safety and explosion safety in car parks.
- [59] Y. Dedeoglu, B.U. Toreyin, U. Gudiokbay, A.E. Çetin, Real-time fire and flame detection in video, in: *International Conference on Acoustics Speech and Signal Processing (ICASSP)*, 2005, pp. 669–672.
- [60] T. Celik, H. Demirel, Fire detection in video sequences using a generic color model, *Fire Saf. J.* 44 (2) (May 2008) 147–158.
- [61] S. Verstockt, A. Vanoosthuysen, S. Van Hoecke, P. Lambert, R. Van de Walle, Multi-sensor fire detection by fusing visual and non-visual flame features, in: *Proceedings of International Conference on Image and Signal Processing*, June 2010, pp. 333–341.
- [62] J. Han, B. Bhanu, Fusion of color and infrared video for moving human detection, *Pattern Recogn.* 40 (6) (June 2007) 1771–1784.
- [63] R. Vandersmissen, Night-vision camera combines thermal and low-light-level images, *Photonik Int.* 2 (August 2008) 2–4.
- [64] J.C. Owrtusky, D.A. Steinhurst, C.P. Minor, S.L. Rose-Pehrsson, F.W. Williams, D.T. Gottuk, Long wavelength video detection of fire in ship compartments, *Fire Saf. J.* 41 (4) (June 2006) 315–320.
- [65] B.U. Toreyin, R.G. Cinbis, Y. Dedeoglu, A.E. Cetin, Fire detection in infrared video using wavelet analysis, *SPIE Opt. Eng.* 46 (6) (June 2007) 1–9.
- [66] I. Bosch, S. Gomez, R. Molina, R. Miralles, Object discrimination by infrared image processing, in: *Proceedings of the 3rd International Work-Conference on the Interplay between Natural and Artificial Computation (IWINAC)*, June 2009, pp. 30–40.
- [67] O. Gunay, K. Tasdemir, B.U. Toreyin, A.E. Cetin, Video based wildfire detection at night, *Fire Saf. J.* 44 (6) (August 2009) 860–868.
- [68] S. Verstockt, R. Dekeerschieter, A. Vanoosthuysen, B. Merci, B. Sette, P. Lambert, R. Van de Walle, Video fire detection using non-visible light, in: *Proceedings of the 6th International Seminar on Fire and Explosion Hazards*, April 2010.
- [69] Vittoria Bruni, Domenico Vitulano, Zhou Wang, Special issue on human vision and information theory, *Signal Image Video Process.* 7 (3) (2013) 389–390.
- [70] Sanjit Roy, Tamanna Howlader, S.M. Mahbubur Rahman, Image fusion technique using multivariate statistical model for wavelet coefficients, *Signal Image Video Process.* 7 (2) (2013) 355–365.
- [71] B.U. Toreyin, Y. Dedeoglu, A.E. Çetin, Flame detection in video using hidden Markov models, in: *International Conference on Image Processing (ICIP)*, 2005, pp. 1230–1233.
- [72] B.U. Toreyin, Y. Dedeoglu, A.E. Çetin, Wavelet based real-time smoke detection in video, in: *European Signal Processing Conference (EUSIPCO)*, 2005.
- [73] O. Gunay, B.U. Toreyin, K. Kose, A.E. Cetin, Entropy-functional-based online adaptive decision fusion framework with application to wildfire detection in video, *IEEE Trans. Image Process.* 21 (5) (May 2012) 2853–2865.
- [74] L.M. Bregman, The relaxation method of finding the common point of convex sets and its application to the solution of problems in convex programming, *USSR Comput. Math. Math. Phys.* 7 (1967) 200–217.
- [75] J. Johnson, Analysis of image forming systems, in: *Image Intensifier Symposium*, Ft. Belvoir, VA, Warfare Electrical Engineering Department, U.S. Army Research and Development Laboratories, 1958, pp. 244–273.

A. Enis Çetin received the Ph.D. degree from the University of Pennsylvania, Philadelphia, in 1987. From 1987 to 1989, he was an Assistant Professor of Electrical Engineering with the University of Toronto, Toronto, ON, Canada. Since 1989, he has been with Bilkent University, Ankara, Turkey. His research interests include signal and image processing, human–computer interaction using vision and speech, and audio-visual multimedia databases. Dr. Çetin was an Associate Editor of the IEEE Transactions on Image Processing, between 1999 and 2003. He is currently the Editor-in-Chief of the Signal, Image and Video Processing journal and a member of the Editorial Boards of Journals Signal Processing and Journal of Advances in Signal Processing and Journal of Machine Vision and Applications.

Kosmas Dimitropoulos received his B.Sc. degree in Electrical and Computer Engineering from Democritus University and his Ph.D. degree in Applied Informatics from Macedonia University of Thessaloniki in 2001 and 2007 respectively. He is a postdoctoral research fellow with the Information Technologies Institute, Centre for Research and Technology Hellas and a visiting lecturer at Macedonia University. His main research interests include computer vision, virtual reality, computer graphics and 3D motion analysis. He has participated in several European and national research projects and he has served as a regular reviewer for a number of international journals and conferences.

Benedict Gouverneur received his Electrical Engineering degree in Electronics in 1984 from ISIB at Brussels, his MScEE degree in System, Signal & Control in 1993 and his MScEE degree in Applied Mathematics in 1995 from UCL at LLN Belgium. Benedict joined Xenics as a System Engineer in Space, Security and Research Projects in 2008 and as an Electro-Optical Characterization Engineer in Infrared in 2012. Before joining Xenics, his R&D activities were in the field of microwave, digital hardware, signal processing and pattern recognition for coherent pulse Doppler and high resolution radar systems. He developed various X-Ray, Imaging, VTS and HLS systems. His research interests include sensors and data processing.

Nikos Grammalidis is a Senior Researcher (Researcher Grade B) at the Information Technologies Institute – Centre of Research and Technology Hellas. He received the B.S. and Ph.D. degrees in Electrical and Computer Engineering from the Aristotle University of Thessaloniki, in 1992 and 2000, respectively. Prior to his current position, he was a researcher in 3D Imaging Laboratory at the Aristotle University of Thessaloniki. His main research interests include computer vision, signal, image and video processing, stereoscopy and multiview image sequence analysis and coding. His involvement with those research areas has led to the co-authoring of more than 25 articles in refereed journals and more than 75 papers in international conferences. Since 1992, he has been actively involved in more than 25 EC and National projects. He has served as a regular reviewer for a number of international journals and conferences.

Osman Günay received his B.Sc. and M.S. degrees in Electrical and Electronics Engineering from Bilkent University, Ankara, Turkey. Since 2009, he has been a Ph.D. Student in the Department of Electrical and Electronics Engineering at Bilkent University, Ankara, Turkey. His research interests include computer vision, video segmentation, and dynamic texture recognition.

Yusuf Hakan Habiboglu received his B.Sc. degree in electrical and electronics engineering in 2008 from the Eskisehir Anadolu University. He received his M.S. degree in electrical and electronics engineering from the Bilkent University, Ankara, Turkey, in 2010. His research interests include computer vision, pattern recognition, feature extraction and image segmentation.

Behcet Uğur Toreyin received the B.S. degree in electrical and electronics engineering from Middle East Technical University, Ankara, Turkey, and the M.S. and Ph.D. degrees in electrical and electronics engineering from Bilkent University, Ankara. Between 2009 and 2011, he was a Postdoctoral Research Associate with the Robotic Sensor Networks Lab, University of Minnesota, Minneapolis, and the Wireless Research Lab, Texas A&M University at Qatar, Doha, Qatar, respectively. He is currently an Assistant Professor with Cankaya University, Ankara.

Steven Verstockt received his Master degree in Informatics from Ghent University in 2003. At the end of 2007 he joined the ELIT Lab of the University College West-Flanders as a researcher. In 2008, he started a PhD on video fire analysis at the Multimedia Lab of the Department of Electron-

ics and Information Systems of Ghent University – IBBT (Belgium). Since 2012 he works as a post-doctoral researcher in this lab. His research interests include video surveillance, computer vision and multi-sensor data fusion.