

# Computer vision based method for real-time fire and flame detection <sup>☆</sup>

B. Uğur Töreyn <sup>a,\*</sup>, Yiğithan Dedeoğlu <sup>b</sup>, Uğur Güdükbay <sup>b</sup>, A. Enis Çetin <sup>a</sup>

<sup>a</sup> *Department of Electrical and Electronics Engineering, Bilkent University, 06800 Bilkent, Ankara, Turkey*

<sup>b</sup> *Department of Computer Engineering, Bilkent University, 06800 Bilkent, Ankara, Turkey*

Received 26 June 2004; received in revised form 20 June 2005

Available online 26 August 2005

Communicated by R. Davies

## Abstract

This paper proposes a novel method to detect fire and/or flames in real-time by processing the video data generated by an ordinary camera monitoring a scene. In addition to ordinary motion and color clues, flame and fire flicker is detected by analyzing the video in the wavelet domain. Quasi-periodic behavior in flame boundaries is detected by performing temporal wavelet transform. Color variations in flame regions are detected by computing the spatial wavelet transform of moving fire-colored regions. Another clue used in the fire detection algorithm is the irregularity of the boundary of the fire-colored region. All of the above clues are combined to reach a final decision. Experimental results show that the proposed method is very successful in detecting fire and/or flames. In addition, it drastically reduces the false alarms issued to ordinary fire-colored moving objects as compared to the methods using only motion and color clues.

© 2005 Elsevier B.V. All rights reserved.

*Keywords:* Fire detection in video; Wavelet transform; Gaussian mixture model; Event detection in video

## 1. Introduction

Conventional point smoke and fire detectors are widely used in buildings. They typically detect the presence of certain particles generated by smoke and fire by ionization or photometry. Alarm is not issued unless particles reach the sensors to activate them. Therefore, they cannot be used in open spaces and large covered areas. Video based fire detection systems can be useful to detect fire in large auditoriums, tunnels, atriums, etc. The strength of using video in fire detection makes it possible to serve large and open spaces. In addition, closed circuit television (CCTV) sur-

veillance systems are currently installed in various public places monitoring indoors and outdoors. Such systems may gain an early fire detection capability with the use of a fire detection software processing the outputs of CCTV cameras in real time.

Image and video content understanding and analysis methods have been studied by many researchers including (Davis and Bobick, 1997; Haering et al., 2000; Javed and Shah, 2002; Naphade et al., 1998). Content based understanding methods have to be designed according to the specific application. Fire detection in video is such an application that needs specific methods. There are several video-based fire and flame detection algorithms in the literature (Chen et al., 2004; Fastcom Technology, 2002; Healey et al., 1993; Liu and Ahuja, 2004; Phillips et al., 2002). These methods make use of various visual signatures including color, motion and geometry of fire regions. Healey et al. (1993) use only color clues for flame detection. Phillips et al. (2002) use pixel colors and their temporal variations. Chen et al. (2004) utilize a change detection

<sup>☆</sup> This work is supported in part by European Commission 6th Framework Program with grant number FP6-507752 (MUSCLE Network of Excellence Project).

\* Corresponding author. Tel.: +90 312 290 1286; fax: +90 312 266 4192.

*E-mail addresses:* [bugur@bilkent.edu.tr](mailto:bugur@bilkent.edu.tr) (B.U. Töreyn), [yigithan@cs.bilkent.edu.tr](mailto:yigithan@cs.bilkent.edu.tr) (Y. Dedeoğlu), [gudukbay@cs.bilkent.edu.tr](mailto:gudukbay@cs.bilkent.edu.tr) (U. Güdükbay), [cetin@bilkent.edu.tr](mailto:cetin@bilkent.edu.tr) (A.E. Çetin).

scheme to detect flicker in fire regions. In (Fastcom Technology, 2002) Fast Fourier Transforms (FFT) of temporal object boundary pixels are computed to detect peaks in Fourier domain. Liu and Ahuja (2004) also represent the shapes of fire regions in Fourier domain. An important weakness of Fourier domain methods is that flame flicker is not purely sinusoidal but random. Therefore, it is hard to detect peaks in FFT plots. In addition, Fourier Transform does not carry any time information. In order to make FFTs also carry time information, they have to be computed in windows of data. Hence, temporal window size is very important for detection. If the window size is too long, then one may not observe peakiness in the FFT data. If it is too short, then one may completely miss cycles and therefore no peaks can be observed in the Fourier domain.

We developed a wavelet based fire detection method in video. Our method not only detects fire and flame colored moving regions in video but also analyzes the motion of such regions in wavelet domain for flicker estimation. It is observed that turbulent flames flicker with a characteristic flicker frequency of around 10 Hz independent of the burning material and the burner (Albers and Agrawal, 1999; Chamberlin and Rose, 1965). The appearance of an object whose contours, chrominance or luminosity values oscillate at a frequency greater than 0.5 Hz is an important sign of the possible presence of flames (Fastcom Technology, 2002). Therefore, fire detection schemes can be made more robust to false alarms by detecting periodic high-frequency behavior in flame colored moving pixels.

High-frequency analysis of moving pixels is carried out in wavelet domain in our work. There is an analogy between our motion analysis in wavelet domain and the temporal templates of (Davis and Bobick, 1997) and the motion recurrence images of (Javed and Shah, 2002). Wavelet transform is a time–frequency analysis tool, and one can examine an entire frequency band in the wavelet domain without completely losing the time information (Cetin and Ansari, 1994; Mallat and Zhong, 1992). Since the wavelet transform is computed using a subband decomposition filter bank, it does not require any batch processing. It is ideally suited to determine an increase in high-frequency activity in fire and flame colored moving objects by detecting zero crossings of the wavelet transform coefficients.

Turbulent high-frequency behaviors exist not only on the boundary but also inside a fire region. Another novelty of the proposed method is the analysis of the spatial variations inside fire and flame colored regions. The method described in (Fastcom Technology, 2002), does not take advantage of such color variations. Spatial wavelet analysis makes it possible to detect high-frequency behavior inside fire regions. Variation in energy of wavelet coefficients is an indicator of activity within the region. On the other hand, a fire-colored moving object will not exhibit any change in values of wavelet coefficients because there will not be any variation in fire-colored pixel values.

## 2. Detection algorithm

The proposed video-based fire detection algorithm consists of four steps: (i) moving pixels or regions in the current frame of a video are determined, (ii) colors of moving pixels are checked to see if they match to pre-specified fire-colors, then wavelet analysis in (iii) temporal and (iv) spatial domains is carried out to determine high-frequency activity within these moving regions. In the following subsections, each step of the proposed algorithm is explained in detail.

### 2.1. Moving region detection

Moving pixels and regions in the video are determined by using a hybrid background estimation method developed by Collins et al. (1999). Let  $x_n[k, l]$  represent the intensity (brightness) value at pixel position  $[k, l]$  in the  $n$ th image frame  $x_n$ . Estimated background intensity value at the same pixel position,  $B_{n+1}[k, l]$ , is calculated as follows:

$$B_{n+1}[k, l] = \begin{cases} aB_n[k, l] + (1 - a)x_n[k, l] & \text{if } [k, l] \text{ is non-moving} \\ B_n[k, l] & \text{if } [k, l] \text{ is moving,} \end{cases} \quad (1)$$

where  $B_n[k, l]$  is the previous estimate of the background intensity value at the same pixel position. The update parameter  $a$  is a positive real number close to one. Initially,  $B_0[k, l]$  is set to the first image frame  $x_0[k, l]$ . A pixel positioned at  $[k, l]$  is assumed to be moving if the brightness values corresponding to it in image frame  $x_n$  and image frame  $x_{n-1}$  satisfy the following inequality:

$$|x_n[k, l] - x_{n-1}[k, l]| > T_n[k, l], \quad (2)$$

where  $T_n[k, l]$  is a threshold that is adaptively determined as described in (Collins et al., 1999). It is assumed that the regions significantly different from the background are moving regions. Estimated background image is subtracted from the current image to detect moving regions. In other words, all of the pixels satisfying

$$|x_n[k, l] - B_n[k, l]| > T_n[k, l] \quad (3)$$

are determined. These pixels at  $[k, l]$  locations are grouped into connected regions (blobs) and labeled by using a two-level connected component labeling algorithm (Heijden, 1996). The output of the first step of the algorithm is a binary pixel map  $Blobs[k, l]$  that indicates whether or not the pixel at location  $[k, l]$  in image  $x$  is moving.

Other more sophisticated methods, including the ones developed by Bagci et al. (2002) and Stauffer and Grimson (1999), can also be used for moving pixel estimation. In our application, accurate detection of moving regions is not as critical as in other object tracking and estimation problems; we are mainly concerned with real-time detection of moving regions as an initial step in the fire and flame detection system. We choose to implement the method suggested by Collins et al. (1999), because of its computational efficiency.

2.2. Detection of fire-colored pixels

Color values of moving pixels are compared with a pre-determined color distribution, which represents possible fire-colors in video in RGB color space. The fire-color distribution is obtained from sample images containing fire regions. Possible color values form a three dimensional point cloud in the RGB color space as shown in Fig. 1(a) (For interpretation of colors in the figures the reader is referred to the web version of this article). The cloud is represented by using a mixture of Gaussians in the RGB color space as described in (Reynolds and Rose, 1995). A Gaussian mixture model with ten Gaussian distributions is estimated from past observations as shown in Fig. 1(b).

Let  $x[k, l]$  be a pixel at location  $[k, l]$  with color values  $[r_{[k,l]}, g_{[k,l]}, b_{[k,l]}]$ . We check if the pixel lies within two standard deviations of the centers of the Gaussians to determine its nature. In other words, if a given pixel color value is inside one of the spheres shown in Fig. 1(b), then it is assumed to be a fire-colored pixel. We set a binary mask, called *fire-colored*, which returns whether a given pixel is fire-colored or not. The intersection of this mask with *Blobs* formed in the first step is fed into the next step as a new binary mask called *fire*.

2.3. Temporal wavelet analysis

The third step of our fire detection algorithm is to keep track of the frequency history of pixels in the fire-colored region and analyze the history. In order to detect flicker or oscillations in pixels due to fire in a reliable manner, the video capture rate should be high enough to capture high-frequency flicker in flames. To capture 10 Hz flicker, the video should capture at least 20 frames per second (fps). However, in some surveillance systems, the video capture rate is below 20 Hz. If the video is available at a lower capture rate, aliasing occurs but flicker due to flames can still be observed in the video. For example, 8 Hz sinusoid appears as 2 Hz sinusoid in a 10 fps video.

Each pixel  $x_n[k, l]$  of the binary mask *fire* is fed to a two stage-filter bank as shown in Fig. 2. The signal  $x_n[k, l]$  is a one-dimensional signal representing the temporal variations in color values at location  $[k, l]$  in the  $n$ th frame. Temporal wavelet analysis can be carried out using either the luminance (Y component) in YUV color representation or the red component in RGB color representation. In our implementation the red channel values of the pixels are used. The two-channel subband decomposition filter bank is composed of half-band high-pass and low-pass

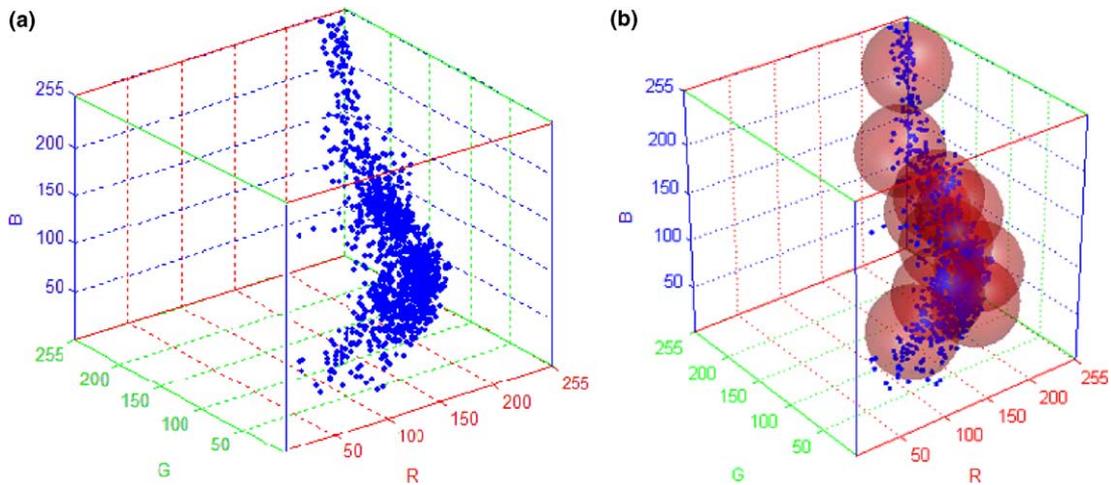


Fig. 1. (a) A sample fire-color cloud in RGB space, and (b) the spheres centered at the means of the Gaussian distributions with radius twice the standard deviation.

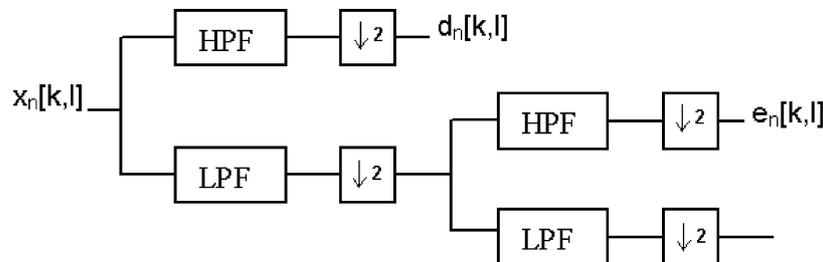


Fig. 2. A two-stage filter bank. HPF and LPF represent half-band high-pass and low-pass filters, with filter coefficients  $\{-0.25, 0.5, -0.25\}$  and  $\{0.25, 0.5, 0.25\}$ , respectively. This filter bank is used for wavelet analysis.

filters with filter coefficients  $\{-0.25, 0.5, -0.25\}$  and  $\{0.25, 0.5, 0.25\}$ , respectively, as shown in Fig. 2. The filter bank produces wavelet subsignals  $d_n[k, l]$  and  $e_n[k, l]$ . If there is high frequency activity at pixel location  $[k, l]$ , high-band subsignals  $d$  and  $e$  get non-zero values. However, in a stationary pixel, the values of these two subsignals should be equal to zero or very close to zero because of high-pass filters used in subband analysis. If the pixel is part of a flame boundary at some time (see Fig. 3), then there will be several spikes in one second due to transitions from background colors to flame colors and vice versa. If there is an ordinary fire-colored moving object going through pixel  $[k, l]$ , then there will be a single spike in one of these wavelet subsignals because of the transition from the background pixel to the object pixel as shown in Fig. 4. The number of zero crossings of the subband signals  $d_n$  and  $e_n$  in a few seconds is used to discriminate between a flame pixel and an ordinary fire-colored object pixel. If this number is above some threshold, then an alarm can be issued for this pixel.

The temporal history of the red channel of a pixel  $x_n[111, 34]$  which is part of a flame, and the corresponding wavelet signals are shown in Fig. 3. A flicker in the red channel values of this flame pixel is obvious from the figure. The pixel is part of a flame for image frames  $I_n$ ,  $n = 1, 2, 3, 19, 23, 24, 41$  and  $50$ . It becomes part of the background for  $n = 12, \dots, 17, 20, 21, 26, 27, 31, \dots, 39,$

$45, 52, \dots$  and  $60$ . Wavelet domain subsignals  $d_n$  and  $e_n$  reveal the fluctuations of the pixel at  $[111, 34]$  with several zero crossings. Due to a down-sampling operation during wavelet computation, the length of wavelet signals are halved after each stage of subband filtering. As a result, the value of a sample in a subband signal corresponds to several samples in the original signal, e.g., the value of  $d_5[111, 34]$  corresponds to the values of  $x_{10}[111, 34]$  and  $x_{11}[111, 34]$ , and the value of  $e_4[111, 34]$  corresponds to the values of  $x_{12}[111, 34]$ ,  $x_{13}[111, 34]$ ,  $x_{14}[111, 34]$  and  $x_{15}[111, 34]$ , in the original signal.

The temporal history of the red channel of a pixel  $x_n[18, 34]$ , which is part of a fire-colored object, and the corresponding wavelet signals are shown in Fig. 4. As shown in this figure, neither the original nor the wavelet signals exhibit oscillatory behavior. The pixel is part of a white-colored background for  $n = 1, 2$  and  $3$ , becomes part of a fire-colored object for  $n = 4, 5, 6, 7$  and  $8$ , then it becomes part of the background again for  $n > 8$ . Corresponding wavelet signals  $d_n$  and  $e_n$  do not exhibit oscillatory behavior as shown in Fig. 4. Small variations due to noise around zero after the 10th frame are ignored by setting up a threshold.

The number of wavelet stages needed for used in flame flicker analysis is determined by the video capture rate. In the first stage of dyadic wavelet decomposition, the low-band subsignal and the high-band wavelet subsignal

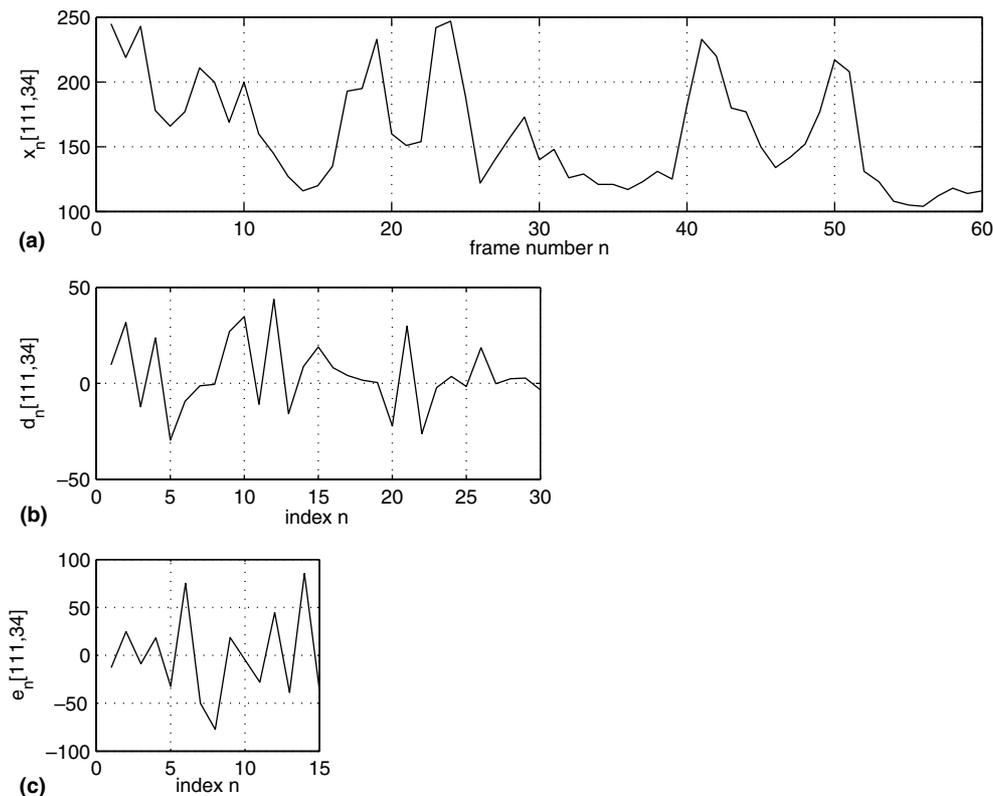


Fig. 3. (a) Temporal variation of image pixels  $x_n[111, 34]$  in time. The pixel at  $[111, 34]$  is part of a flame for image frames  $x_n$ ,  $n = 1, 2, 3, 19, 23, 24, 41$  and  $50$ . It becomes part of the background for  $n = 12, \dots, 17, 20, 21, 26, 27, 31, \dots, 39, 45, 52, \dots$  and  $60$ . Wavelet domain subsignals (b)  $d_n$  and (c)  $e_n$  reveal the fluctuations of the pixel at  $[111, 34]$ .

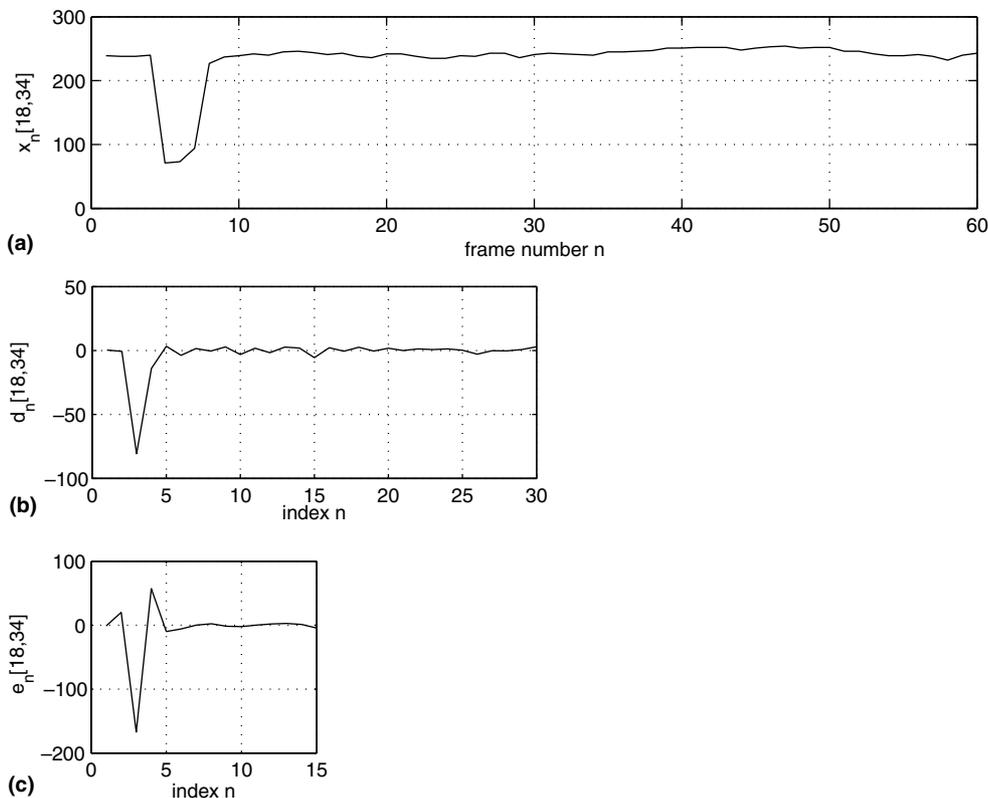


Fig. 4. (a) Temporal history of the pixel [18,34] in time. It is part of a fire-colored object for  $n = 4, 5, 6, 7$  and  $8$ , and it becomes part of the background afterwards. Corresponding subsignals (b)  $d_n$  and (c)  $e_n$  exhibit stationary behavior for  $n > 8$ .

$d_n[k, l]$  of the signal  $x_n[k, l]$  are obtained. The subsignal  $d_n[k, l]$  contains [2.5 Hz, 5 Hz] frequency band information of the original signal  $x_n[k, l]$  in 10 Hz video frame rate. In the second stage the low-band subsignal is processed once again using a dyadic filter bank, and the wavelet subsignal  $e_n[k, l]$  covering the frequency band [1.25 Hz, 2.5 Hz] is obtained. Thus by monitoring wavelet subsignals  $e_n[k, l]$  and  $d_n[k, l]$ , one can detect fluctuations between 1.25 Hz and 5 Hz in the pixel  $[k, l]$ .

#### 2.4. Spatial wavelet analysis

The fourth step of our fire detection algorithm is the spatial wavelet analysis of moving regions containing *fire* mask pixels to capture color variations in pixel values. In an ordinary fire-colored object there will be little spatial variations in the moving region as shown in Fig. 5(a). On the other hand, there will be significant spatial variations in a fire region as shown in Fig. 6(a). The spatial wavelet analysis of a rectangular frame containing the pixels of fire-colored moving regions is performed. The images in Figs. 5(b) and 6(b) are obtained after a single stage two-dimensional wavelet transform that is implemented in a separable manner using the same filters explained in Section 2.3. Absolute values of low–high, high–low and high–high wavelet subimages are added to obtain these images. A decision parameter  $v_4$  is defined for this step, according to the energy of the wavelet subimages:

$$v_4 = \frac{1}{M \times N} \sum_{k,l} |x_{lh}[k, l]|^2 + |x_{hl}[k, l]|^2 + |x_{hh}[k, l]|^2, \quad (4)$$

where  $x_{lh}[k, l]$  is the low–high subimage,  $x_{hl}[k, l]$  is the high–low subimage, and  $x_{hh}[k, l]$  is the high–high subimage of the wavelet transform, respectively, and  $M \times N$  is the number of pixels in the fire-colored moving region. If the decision parameter of the fourth step of the algorithm,  $v_4$ , exceeds a threshold, then it is likely that this moving and fire-colored region under investigation is a fire region.

Both the 1-D temporal wavelet analysis described in Section 2.3 and the 2-D spatial wavelet analysis are computationally efficient schemes because a multiplierless filter bank is used for both 1-D and 2-D wavelet transform computation (Gerek and Cetin, 2000; Kim et al., 1992). Lowpass and highpass filters have weights  $[\frac{1}{4}, \frac{1}{2}, \frac{1}{4}]$  and  $[\frac{-1}{4}, \frac{1}{2}, \frac{-1}{4}]$ , respectively. They can be implemented by register shifts without performing any multiplications.

The wavelet analysis based steps of the algorithm are very important in fire and flame detection because they distinguish ordinary motion in the video from motion due to turbulent flames and fire.

### 3. Decision fusion

In this section, we describe a voting based decision fusion strategy. However, other data fusion methods can be



Fig. 5. (a) A child with a fire-colored t-shirt, and (b) the absolute sum of spatial wavelet transform coefficients,  $|x_{lh}[k, l]| + |x_{hl}[k, l]| + |x_{hh}[k, l]|$ , of the region bounded by the indicated rectangle.



Fig. 6. (a) Fire, and (b) the absolute sum of spatial wavelet transform coefficients,  $|x_{lh}[k, l]| + |x_{hl}[k, l]| + |x_{hh}[k, l]|$ , of the region bounded by the indicated rectangle.

also used to combine the decision of four stages of the flame and fire detection algorithm.

Voting schemes include *unanimity voting*, *majority voting*, and *m-out-of-n voting* in which an output choice is accepted if at least  $m$  votes agree with the decisions of  $n$  sensors (Parhami, 1994). We use a variant of *m-out-of-n voting*, the so-called *T-out-of-v* voting in which the output is accepted if  $H = \sum_i w_i v_i > T$  where the  $w_i$ 's are user-defined weights, the  $v_i$ 's are decisions of the four stages of the algorithm, and  $T$  is a user-defined threshold. Decision parameter  $v_i$  can take binary values 0 and 1, corresponding to normal case and the existence of fire, respectively. The decision parameter  $v_1$  is 1 if the pixel is a moving pixel, and 0 if it is stationary. The decision parameter  $v_2$  is taken

as 1 if the pixel is fire-colored, and 0 otherwise. The decision parameter  $v_3$  is 1 if the number of zero crossings of  $e_n[k, l]$  and/or  $d_n[k, l]$  in a few seconds exceed a threshold value, and 0 otherwise. The decision parameter  $v_4$  is defined in Eq. (4).

In uncontrolled fire, it is expected that the fire region should have a non-convex boundary. To gain a further robustness to false alarms, another step checking the convexity of the fire region is also added to the proposed algorithm. Convexity of regions is verified in a heuristic manner. Boundaries of the regions in the *fire* mask are checked for their convexity along equally spaced five vertical and five horizontal lines using a  $5 \times 5$  grid. The analysis simply consists of checking whether the pixels on each line

belong to the region or not. If at least three consecutive pixels belong to the background, then this region violates the convexity condition. A *fire* mask region which has background pixels on the intersecting vertical and/or horizontal lines, is assumed to have a non-convex boundary. This eliminates false alarms due to match light sources, sun, etc.

#### 4. Experimental results

The proposed method, Method 1, is implemented on a PC with an Intel Pentium 4, 2.40 GHz processor. It is tested for a large variety of conditions in comparison with the method utilizing only the color and temporal variation information, which we call Method 2, described in Phillips et al. (2002). The scheme described in (Chen et al., 2004) is also implemented for comparison and it is called as Method 3 in the rest of the article. The results for some of the test sequences are presented in Table 1.

Method 2 is successful in determining fire and does not recognize stationary fire-colored objects such as the sun as fire. However, it gives false alarms when the fire-colored

ordinary objects start to move, as in the case of a realistic scenario. An example of this is shown in Fig. 7(a). The proposed method does not give any false alarms for this case (Fig. 7(b)). The fire-colored strip on the cargo truck triggers an alarm in Method 2 when the truck starts to move. Similarly, false alarms are issued with Method 2 in Movies 3, 7 and 9, although there are no fires taking place in these videos. The moving arm of a man is detected as fire in Movie 7 (Fig. 8(c)), and a red parking car is marked as fire in Movie 9 with Method 2 (Fig. 8(d)).

Method 3 gives similar detection results for fire. However, it also suffers from inefficient analysis of the motion of fire-colored objects. Fire-colored ordinary moving objects causes Method 3 to give false alarms in Movies 1, 3, 7 and 9. If Method 1 is used, moving fire-colored ordinary objects do not cause an alarm to be raised. This is because the cyclic movement of flames is taken into account in our method, as well as the spatial variation in the color/brightness values of the moving fire-colored regions. Method 1 successfully detects fire in videos covering various scenarios, including partial occlusion of the flame. Sample images showing the detected regions are presented in Fig. 9.

Table 1

Comparison of the proposed method (Method 1), the method based on color and temporal variation clues only (Method 2) described in (Phillips et al., 2002), and the method proposed in (Chen et al., 2004) (Method 3)

Video sequences	Number of frames with fire	Number of frames detected as fire			Number of false positive frames			Description
		Method			Method			
		1	2	3	1	2	3	
Movie 1	0	0	46	13	0	46	13	A fire-colored moving truck
Movie 2	5	5	5	5	0	0	0	Fire in a garden
Movie 3	0	0	7	5	0	7	5	A car leaving a fire-colored parking lot
Movie 4	37	37	44	47	0	7	10	A burning box
Movie 5	64	64	88	84	0	24	20	A burning pile of wood
Movie 6	41	41	56	50	0	15	9	Fire behind a man with a fire-colored shirt
Movie 7	0	0	14	7	0	14	7	Four men walking in a room
Movie 8	18	18	18	18	0	0	0	Fire in a fireplace
Movie 9	0	0	15	5	0	15	5	A crowded parking lot
Movie 10	0	0	0	0	0	0	0	Traffic on a highway
Movie 11	0	9	107	86	9	107	86	Dancing man with fire-colored shirt



Fig. 7. (a) With the method using color and temporal variation only (Method 2) (Phillips et al., 2002), false alarms are issued for the fire colored line on the moving truck and the ground, (b) our method (Method 1) does not produce any false alarms.

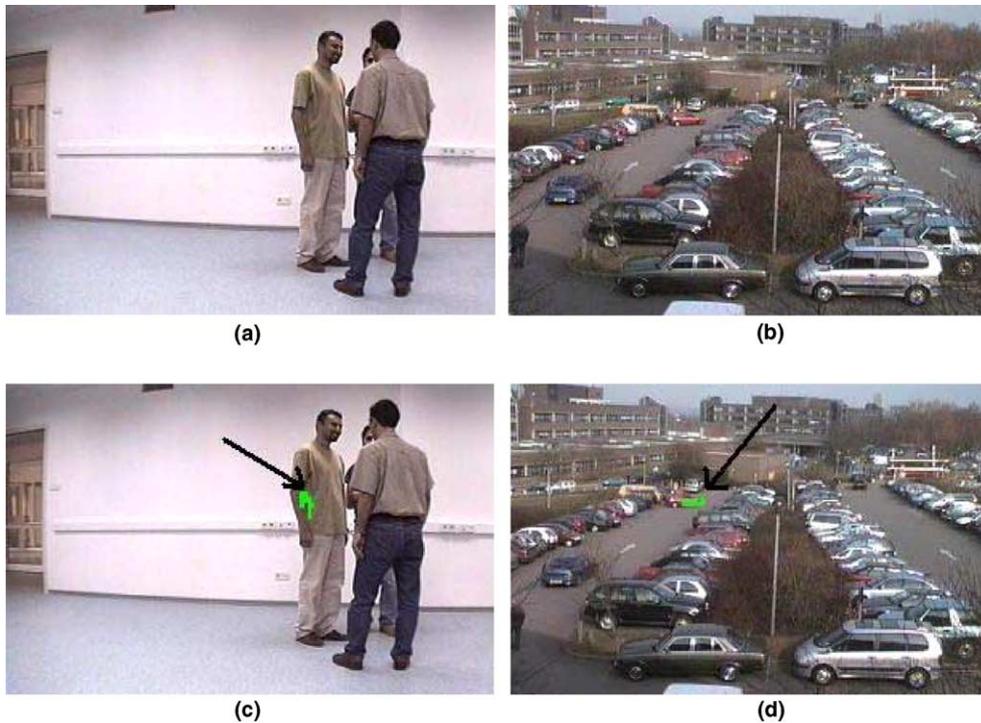


Fig. 8. Sample images (a) and (b) are from Movies 7 and 9, respectively. (c) False alarms are issued for the arm of the man with the method using color and temporal variation only (Method 2) (Phillips et al., 2002) and (d) on the fire-colored parking car. Our method does not give any false alarms in such cases (see Table 1).

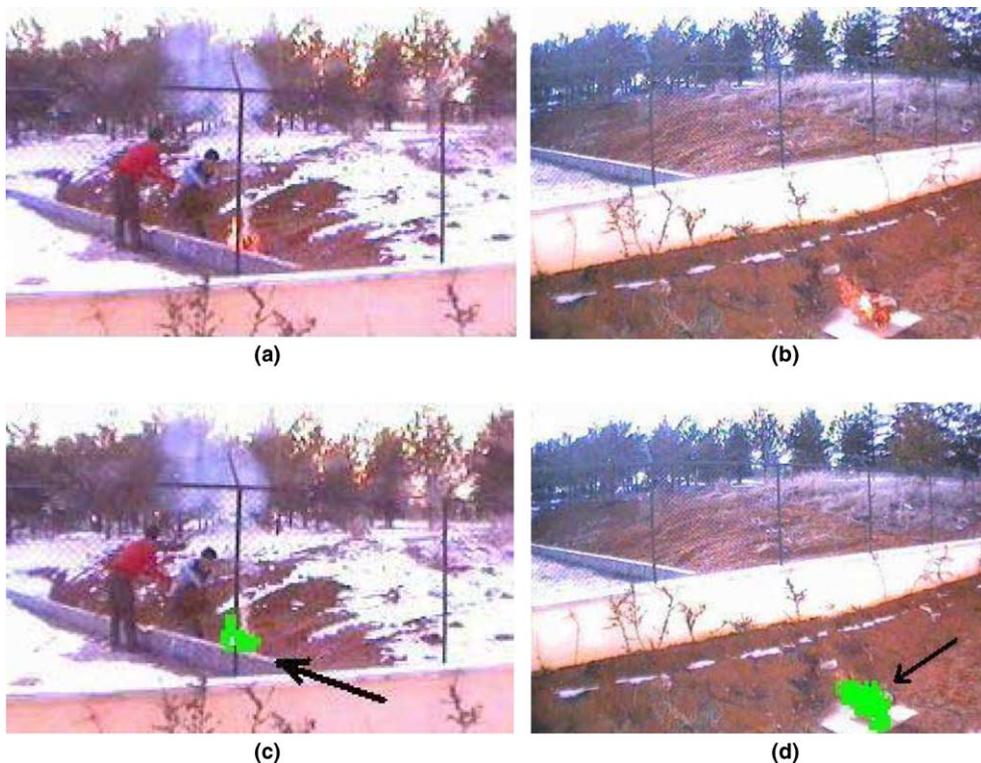


Fig. 9. Sample images (a) and (b) are from Movies 2 and 4, respectively. Flames are successfully detected with our method (Method 1) in (c) and (d). In (c), although flames are partially occluded by the fence, a fire alarm is issued successfully. Fire pixels are painted in bright green.

In Movie 11, a man wearing a fire-colored shirt intentionally waves his arms to mimic the quasi-periodic flicker

behavior in flames. Although all of the methods produce false alarms in this Movie, Method 1 significantly decre-

Table 2

Time performance comparison of Methods 1, 2 and 3 for the movies in Table 1

Videos	Method 1	Method 2	Method 3
Movie 1	16	12	14
Movie 2	16	12	14
Movie 3	16	12	14
Movie 4	16	12	14
Movie 5	17	13	15
Movie 6	17	13	15
Movie 7	17	13	15
Movie 8	17	13	15
Movie 9	16	12	14
Movie 10	16	12	14
Movie 11	16	12	14

The values are the processing times per frame in milliseconds.

ases the number of false positives relative to Methods 2 and 3.

These methods are also compared to each other in terms of computational cost (as shown in Table 2). Movies in Tables 1 and 2 are all captured at 10 fps with a frame size of 320 by 240 pixels. The average processing times per frame are 16.5 ms, 12.5 ms and 14.5 ms, for our method, Methods 2 and 3, respectively. Our method is computationally more demanding due to additional wavelet analysis based steps. Since only shift and add type operations take place when convolving signals with the wavelet filters, additional cost is not high. Our implementation works in real-time for videos with frame size 320 by 240 pixels, captured at 10 fps or higher in a PC.

The video clips that we tested our method contain a total of 83,745 frames in 61 sequences. In 19 of the sequences fire takes place. Our method is successful in detecting fire in all of these sequences. This corresponds to a fire detection rate of 1.0. A fire contour recognition rate of 0.999 is reported in (Liu and Ahuja, 2004) which corresponds to a fire detection rate of 0.999. Our overall false alarm (false positive) rate is 0.001. It is reported that non-fire contour recognition rate is 1.0 in (Liu and Ahuja, 2004) which corresponds to a false alarm rate of 0. The video sequences containing fire in (Liu and Ahuja, 2004) are not publicly available. Therefore we used our own data set. We also test our method with the data set of the EU funded CAVIAR project/IST 2001 37540, publicly available at URL: <http://homepages.inf.ed.ac.uk/rbf/CAVIAR/>. Although there are a lot of clips with moving fire-colored objects, none of the clips in this data set contains fire. Our method gives no false alarms in any of these sequences.

## 5. Conclusion

An algorithm for fire and flame detection in color video is developed. The algorithm not only uses color and temporal variation information, but also checks flicker in flames using 1-D temporal wavelet transform and color variation in fire-colored moving regions using 2-D spatial wavelet transform. Methods based on only color information and ordinary motion detection may produce false alarms in real

scenes where no fires are taking place. The experimental results show that false alarms can be drastically reduced by temporal and spatial wavelet analysis.

The method can be used for detection of fire in movies and video databases, as well as real-time detection of fire. It can be incorporated into a surveillance system monitoring an indoor or outdoor area of interest for early detection of fire.

## Acknowledgements

We are grateful to Erin Maloney and Kirsten Ward for proofreading and suggestions.

## References

- Albers, B.W., Agrawal, A.K., 1999. Schlieren analysis of an oscillating gas-jet diffusion. *Combust. flame* 119, 84–94.
- Bagci, M., Yardimci, Y., Cetin, A.E., 2002. Moving object detection using adaptive subband decomposition and fractional lower order statistics in video sequences. *Signal Process.*, 1941–1947.
- Cetin, A.E., Ansari, R., 1994. Signal recovery from wavelet transform maxima. *IEEE Trans. Signal Process.* 42, 194–196.
- Chamberlin, D.S., Rose, A., 1965. The First Symposium (International) on Combustion. The Combustion Institute, Pittsburgh, 27–32.
- Chen, T., Wu, P., Chiou, Y., 2004. An early fire-detection method based on image processing. In: *Proc. IEEE Internat. Conf. on Image Processing, ICIP '04*, pp. 1707–1710.
- Collins, R.T., Lipton, A.J., Kanade, T., 1999. A system for video surveillance and monitoring. In: *Proc. American Nuclear Society (ANS) Eighth International Topical Meeting on Robotics and Remote Systems*, Pittsburgh, PA.
- Davis, J.W., Bobick, A.F., 1997. The representation and recognition of action using temporal templates. In: *Proc. IEEE Computer Vision and Pattern Recognition Conference (CVPR'97)*, pp. 928–934.
- Fastcom Technology SA., 2002. Boulevard de Grancy 19A, CH-1006 Lausanne, Switzerland, Method and Device for Detecting Fires Based on Image Analysis. Patent Cooperation Treaty Application No. PCT/CH02/00118, PCT Publication No. WO02/069292.
- Gerek, O.N., Cetin, A.E., 2000. Adaptive polyphase subband decomposition structures for image compression. *IEEE Trans. Image Process.* 9 (10), 1649–1660.
- Haering, N., Qian, R.J., Sezan, M.I., 2000. A semantic event-detection approach and its application to detecting hunts in wildlife video. *IEEE Trans. Circ. Syst. Video Technol.* 10 (6), 857–868.
- Healey, G., Slater, D., Lin, T., Drda, B., Goedeke, A.D., 1993. A system for real-time fire detection. In: *Proc. IEEE Computer Vision and Pattern Recognition Conference (CVPR'93)*, pp. 605–606.
- Heijden, F., 1996. *Image Based Measurement Systems: Object Recognition and Parameter Estimation*. Wiley.
- Javed, O., Shah, M., 2002. Tracking and object classification for automated surveillance. In: *Proc. European Conf. on Computer Vision (ECCV'02)*, pp. 343–357.
- Kim, C.W., Ansari, R., Cetin, A.E., 1992. A class of linear-phase regular biorthogonal wavelets. In: *Proc. Internat. Conf. on Acoustics, Speech, and Signal Processing, ICASSP-92*, vol. 4, pp. 673–676.
- Liu, C.B., Ahuja, N., 2004. Vision based fire detection. In: *Proc. Internat. Conf. on Pattern Recognition (ICPR'04)*, vol. 4, pp. 134–137.
- Mallat, S., Zhong, S., 1992. Characterization of signals from multiscale edges. *IEEE Trans. Pattern Anal. Machine Intell.* 14 (7), 710–732.
- Naphade, M.R., Kristjansson, T., Frey B., Huang, T. S., 1998. Probabilistic multimedia objects (multijets): A novel approach to video indexing and retrieval in multimedia systems. In: *Proc. IEEE Internat. Conf. on Image Processing, ICIP '98*, pp. 536–540.
- Parhami, B., 1994. Voting algorithms. *IEEE Trans. Reliability* 43 (4), 617–629.

- Phillips III, W., Shah, M., Lobo, N.V., 2002. Flame recognition in video. *Pattern Recognition Lett.* 23 (1–3), 319–327.
- Reynolds, D.A., Rose, R.C., 1995. Robust text-independent speaker identification using gaussian mixture speaker models. *IEEE Trans. Speech Audio Process.* 3 (1), 72–83.
- Stauffer, C., Grimson, W.E.L., 1999. Adaptive background mixture models for real-time tracking. In: *Proc. IEEE Computer Society Conf. on Computer Vision and Pattern Recognition*, vol. 2, pp. 246–252.