

WAVELET BASED REAL-TIME SMOKE DETECTION IN VIDEO

B. Uğur Töreyn¹, Yiğithan Dedeoğlu², and A. Enis Çetin¹

¹Department of Electrical and Electronics Engineering

²Department of Computer Engineering, Bilkent University
06800, Ankara, Turkey

phone: + (90) 312 2901286, fax: + (90) 312 2664192, email: {bugur,yigithan,cetin}@bilkent.edu.tr

ABSTRACT

A method for smoke detection in video is proposed. It is assumed the camera monitoring the scene is stationary. Since the smoke is semi-transparent, edges of image frames start losing their sharpness and this leads to a decrease in the high frequency content of the image. To determine the smoke in the field of view of the camera, the background of the scene is estimated and decrease of high frequency energy of the scene is monitored using the spatial wavelet transforms of the current and the background images. Edges of the scene are especially important because they produce local extrema in the wavelet domain. A decrease in values of local extrema is also an indicator of smoke. In addition, scene becomes grayish when there is smoke and this leads to a decrease in chrominance values of pixels. Periodic behavior in smoke boundaries and convexity of smoke regions are also analyzed. All of these clues are combined to reach a final decision.

1. INTRODUCTION

Conventional point smoke and fire detectors typically detect the presence of certain particles generated by smoke and fire by ionization or photometry. An important weakness of point detectors is that in large rooms, it may take a long time for smoke particles to reach a detector and they cannot be operated in open spaces.

The strength of using ordinary video in fire detection is the ability to serve large and open spaces. Current fire detection algorithms are based on the use of color and motion information in video to detect the flames [1], [2]. However, smoke detection is vital for fire alarm systems when large and open areas are monitored, because the source of the fire and flames cannot always fall into the field of view. On the contrary, smoke of an uncontrolled fire can be easily observed by a camera even if the flames are not visible. This results in early detection of fire before it spreads around.

Smoke gradually smoothen the edges in an image when it is not that thick to cover the scene. This feature of smoke is a good indicator of its presence in the field of view of the camera. It is thus exploited in this method. Edges in an image correspond to local extrema in wavelet domain. Gradual decrease in their sharpness result in a decrease in the values of these extrema. However, these extrema values corresponding to edges, do not boil down to zero when there is smoke. In fact, they simply loose some of their energy but they still stay in their original locations, occluded partially by the semi-transparent smoke.

Independent of the fuel type, smoke naturally decrease the chrominance channels U and V values of pixels. Apart from this, it is well-known that the flicker frequency of

flames are around 10 Hz. and that this flicker frequency is not greatly affected by either the fuel type or the burner size [2], [6]. As a result, smoke boundaries also oscillate with a lower frequency at the early stages of fire.

Another important feature of the smoke that is exploited in this method is that smoke regions have convex shapes. A group of pixels are not marked as smoke even if they satisfy all of the above criteria, when the region bounded by those pixels has some extensions in arbitrary directions violating the convexity of the region.

2. SMOKE DETECTION USING THE WAVELET ANALYSIS OF VISIBLE-RANGE VIDEO

Methods of identifying fire in video include [1], [2], [3], [10] and [11]. The method in [3] only makes use of the color information. On the other hand, the scheme in [1] is based on detecting the fire colored regions in the current video first. If these fire colored regions move then they are marked as possible regions of fire in the scene monitored by a camera. In [11], a similar method is used which is based on a color model for flame and smoke. The dynamics of flame and smoke regions are described by frame differencing.

By incorporating periodicity analysis around object boundaries one can reduce the false alarms which may be due to flame colored ordinary moving objects. It is well-known that turbulent flames flicker which significantly increase the Fourier frequency content between 0.5 Hz and 20 Hz [2]. In other words, a pixel especially at the edge of a flame could appear and disappear several times in one second of a video. The appearance of an object where the contours, chrominance or luminosity oscillate at a frequency greater than 0.5 Hz is a sign of the possible presence of flames. In [2], Fast Fourier Transforms (FFT) of temporal object boundary pixels are computed to detect peaks in Fourier domain. In [10], the shape of fire regions are represented in Fourier domain, as well. Since, Fourier Transform does not carry any time information, FFTs have to be computed in windows of data and temporal window size is very important for detection. If it is too long then one may not get enough peaks in the FFT data. If it is too short than one may completely miss cycles and therefore no peaks can be observed in the Fourier domain. Another problem is that, one may not detect periodicity in fast growing fires because the boundary of fire region simply grows in video. In [2], FFT analysis inside flame regions was not carried out.

The flames of a fire may not always fall into the visible range of the camera monitoring a scene covering large areas like plane hangars or open spaces. Fire detection systems should tackle with such situations by successful detection

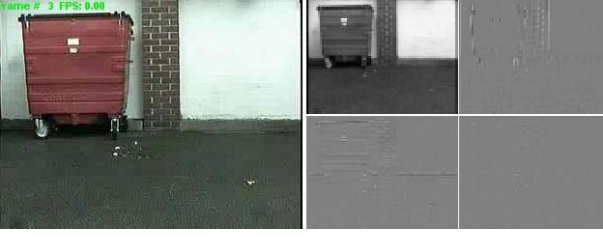


Figure 1: Original frame and its single level wavelet subimages.



Figure 2: Frame with smoke and its single level wavelet subimages. Blurring in the edges is visible.

of smoke without flame. In this paper, temporal and spatial wavelet analysis are carried out for the detection of smoke.

3. DETECTION ALGORITHM

Smoke detection algorithm consists of five steps: (i)moving pixels or regions in the current frame of a video are determined, (ii)the decrease in high frequency content corresponding to edges in these regions are checked using spatial wavelet transform. If the edges loose their sharpness without vanishing completely (iii)the decrease in U and V channels of them are checked, (iv)flicker analysis is carried out using temporal wavelet transform. Finally (v)shape of the moving region is checked for convexity.

Moving pixels and regions in the video are determined by using a background estimation method developed by Collins et.al. [8]. In this method, a background image B_{n+1} at time instant $n + 1$ is recursively estimated from the image frame I_n and the background image B_n of the video as follows:

$$B_{n+1}(k,l) = \begin{cases} aB_n(k,l) + (1-a)I_n(k,l) & (k,l) \text{ stationary} \\ B_n(k,l) & (k,l) \text{ moving} \end{cases} \quad (1)$$

where $I_n(k,l)$ represent a pixel in the n^{th} video frame I_n , and a is a parameter between 0 and 1. Moving pixels are determined by subtracting the current image from the background image and thresholding. A recursive threshold estimation is also described in [8]. Moving regions are determined by using connected component analysis. Other methods like [9] and [7] can also be used for moving pixel estimation.

It is necessary to analyze these moving regions further to determine if the motion is due to smoke or an ordinary moving object. Smoke obstructs the texture and edges in the background of an image. Since the edges and texture contribute to the high frequency information of the image, energies of wavelet subimages drop due to smoke in an image sequence. Based on this fact we monitor wavelet coefficients as in Fig.1 and we detect decreases in local wavelet energy, and detect individual wavelet coefficients corresponding to edges of objects in background whose values decrease over time in video. It is also possible to determine the location of smoke using the wavelet subimages as shown in Fig.2.

Let $w_n(x,y) = |LH_n(x,y)|^2 + |HL_n(x,y)|^2 + |HH_n(x,y)|^2$ represent a composite image containing high-frequency information at a given scale. This subband image is divided into small blocks of size (K_1, K_2) and the energy $e(l_1, l_2)$ of each block is computed as follows

$$e(l_1, l_2) = \sum_{(x,y) \in R_i} w_n(x + l_1K_1, y + l_2K_2) \quad (2)$$

where R_i represents a block of size (K_1, K_2) in the wavelet subimage. If the wavelet subimages are computed from the luminance (Y) image then there is no need to include the chrominance wavelet images. If wavelet transforms of R, G, and B color images are computed then the energy $e(l_1, l_2)$ is computed using all of wavelet subimages of the R, G, and B color images. In our implementation, subimages are computed from the luminance image and the block size is taken as 8 by 8 pixels.

The above local energy values computed for the wavelet transform of the current image are compared to corresponding local high-frequency energies computed from the wavelet transform of the background which contains information about the past state of the scene under observation. If there is a decrease in value of a certain $e(l_1, l_2)$ then this means that the texture or edges of the scene monitored by the camera no longer appear as sharp as they used to be in the current image of the video. Therefore, there may be smoke in the image region corresponding to $(l_1, l_2)^{th}$ block. One can set up thresholds for comparison. If a certain $e(l_1, l_2)$ value drops below the pre-set threshold a warning is issued.

It is also well-known that wavelet subimages contain the edge information of the original image. Edges produce local extrema in wavelet subimages [4], [5]. Wavelet subimages LH, HL and HH contains horizontal, vertical and diagonal edges of the original image, respectively. If smoke covers one of the edges of the original image then the edge initially becomes less visible and after some time it may disappear from the scene as the smoke gets thick. Let the wavelet coefficient $HL_n(x,y)$ be one of the wavelet coefficients corresponding to the edge covered by the smoke. Initially, its value decreases due to the reduced visibility, and in subsequent image frames it becomes either zero or close to zero whenever there is very little visibility due to thick smoke. Therefore locations of the edges of the original image is determined from the significant extrema of the wavelet transform of the background image in our system. Slow fading of a wavelet extrema is an important clue for smoke detection. If the values of a group of wavelet coefficients along a curve corresponding to an edge decrease in value in consecutive frames then this means that there is less visibility in the scene. In turn, this may be due to the existence of smoke. An instantaneous disappearance or appearance of a wavelet extremum in the current frame cannot be due to smoke. Such a change corresponds to an ordinary moving object covering an edge in the background or the boundary of a moving object and such changes are ignored.

In order to determine the decrease in visibility of the edges, we set two thresholds $1 > T_1 > T_2 > 0$. For a decrease in visibility to occur, at a given scale, the composite image

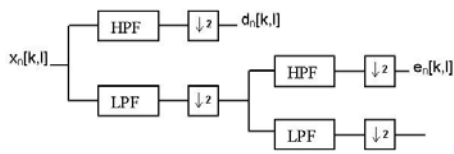


Figure 3: A two stage filter bank.

value $w_n(x,y)$ corresponding to an edge in the current image at location (x,y) and the composite image value $wb_n(x,y)$ similarly calculated for the background image at the same location, must satisfy $T_1wb_n(x,y) > w_n(x,y) > T_2wb_n(x,y)$. Since $T_2 > 0$, we guarantee to have edges that are not totally invisible due to semi-transparent nature of initial smoke.

Color information is also used for identifying smoke in video as the third step. Initially, when the smoke starts to expand, it is semi-transparent thus it preserves the direction of the RGB vector of the background image. This is another clue for differentiating between smoke and an ordinary moving object. By itself, this information is not sufficient because shadows of moving objects also have the same property. As the smoke gets thicker, however, the resemblance of the current frame and the background decreases and the chrominance values U and V of the candidate region in the current frame become smaller than corresponding values in the background image. Only those pixels with lower chrominance values are considered to be smoke.

The flicker in smoke is also used as an additional information. The candidate regions are checked whether they continuously appear and disappear over time. Flames flicker with a characteristic frequency of about 10 Hz independent from the source of the fuel and the burner dimensions [6]. This, in turn induces a less frequent flicker in the smoke boundaries with a frequency range of 1-3 Hz. FFT is used to estimate the unusual activity in [2]. In this paper we describe a wavelet domain approach which is used to determine the temporal high-frequency activity in a pixel. A two-stage filterbank is used for a pixel which satisfy the conditions in steps (i), (ii) and (iii) as shown in Fig.3. Input $x_n[k,l]$ to the filterbank is a one-dimensional signal representing the temporal variations at location $[k,l]$. The signal $x_n[k,l]$ is the luminance (Y component) of the image. We examine the wavelet subsignals $d_n[k,l]$ and $e_n[k,l]$ at 5 Hz image capture rate. In a stationary pixel, values of these two subsignals should be equal to or very close to zero because of high-pass filters used in subband analysis. If there is an ordinary 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. If the pixel is part of a smoke boundary then there will be several spikes in one second due to transitions from background to smoke and smoke to background. Therefore, if $|e_n[k,l]|$ and/or $|d_n[k,l]|$ exceed a threshold value several times in a few seconds then an alarm is issued for this pixel.

The number of wavelet stages that should be used in smoke flicker analysis is determined by the video capture rate. In the first stage of dyadic wavelet decomposition we obtain the low-band subsignal and the high-band wavelet subsignal $d_n[k,l]$ of the signal $x_n[k,l]$. The subsignal $d_n[k,l]$ contains [1.25 Hz, 2.5 Hz] frequency band information of the original signal $x_n[k,l]$ in 5 Hz video frame rate. In the second stage the lowband subsignal is processed once again



Figure 4: Sample images from test videos. Smoke regions are successfully detected. Edge points satisfying all the conditions defined by the method are marked.

using a dyadic filterbank and the subsignal $e_n[k,l]$ is obtained containing [0.625 Hz, 1.25 Hz] frequency band information of the original signal. This means that by monitoring the wavelet subsignals $e_n[k,l]$ and $d_n[k,l]$ one can detect 0.625 to 2.5 Hz fluctuations in the pixel $[k,l]$.

At the last step, the convexity in the shape of the smoke regions is checked. Smoke of an uncontrolled fire expands in time which results in regions with convex boundaries. Boundaries of the moving regions that contain candidate smoke pixels are checked for their convexity along equally spaced vertical and horizontal lines. In our implementation we take five horizontal and five vertical lines and carry out the analysis on them. Analysis simply consists of checking whether the pixels on each of the lines belong to the moving region or not. At least three consecutive pixels on the lines intersecting moving regions must belong to the background, in order to have the moving region violate the convexity condition. If along any one of the lines, convexity is not met, the smoke pixels in that region are discarded.

These clues are then combined to give a final decision. If all of the above mentioned criteria are satisfied for a pixel, the moving region comprising that pixel is determined as smoke.

4. EXPERIMENTAL RESULTS

The proposed method (Method1) is implemented in a laptop with an AMD AthlonXP 2000+ 1.66GHz processor and tested for a large variety of conditions including real-time and off-line videos containing only smoke, both flame and smoke, and videos with no smoke or flame.

The computational cost of the wavelet transform is low. The filterbank in Fig.3 have integer coefficient low and high pass Lagrange filters. The same filters are used for a single level wavelet decomposition of image frames in the spatial wavelet analysis step. Smoke detection is achieved in real-time. The processing time per frame is about 10 msec for frames with sizes of 320 by 240 pixels.

Sample images showing the detected smoke regions are presented in Fig.4. Edge points satisfying all of the conditions are marked inside the detected regions. Detection results for some of the test sequences are presented in Table 1. Smoke is successfully detected in all of the shots con-

Table 1: Detection results of Method1 for some live and off-line videos.

Video Sequences	# of Shots with Smoke	# of Shots detected as Smoke	Description
Movie 1	0	0	A smoke-colored parking car
Movie 2	5	5	Fire in a garden
Movie 3	5	5	Fire in a garden in snow
Movie 4	7	7	A burning box
Movie 5	6	6	A burning pile of woods
Movie 6	3	3	Fire in a waste basket in the garden monitored from an indoor camera
Movie 7	0	0	Three men walking in a room
Movie 8	8	8	Fire in a fireplace
Movie 9	0	1	A parking car in the night

Table 2: Smoke and flame detection time comparison of Method1 and Method2, respectively. Smoke is an early indicator of fire. In Movie 11 and 12, flames are not in the viewing range of the camera.

Video Sequences	Method	Detection at frame#	Description
Movie 10	Method1	25	Fire in a garden
	Method2	87	
Movie 11	Method1	14	Fire in a trash bin
	Method2	-	
Movie 12	Method1	97	Fire at a distance of 25m.
	Method2	-	

taining smoke. No false alarms are issued in live tests and off-line videos recorded in the day time. A false alarm is issued in Movie 9 which is recorded at night. A parking car is captured from its front in this video. The driver intentionally varies the intensity of the front lights of the car. The light beams directed towards the camera at night defines artificial edges around them. These edges appear and disappear continuously as the intensity of the lights change. The U, V channel values of the pixels decrease as the light intensities are lowered, since everywhere in the scene is dark other than the car lights. In this way, car lights at night mimic the smoke characteristics in the day time and a false alarm is issued. Method1 is developed based on the day time characteristics of smoke. The proposed method assumes a well lighted scene, it is not intended for night use.

In videos containing both smoke and flame, Method1 is compared with the flame detection method proposed in [12] (Method2). The comparison results in some of the test sequences are presented in Table 2. At the early stages of fire, smoke is released before flames become visible. Method1 successfully detects smoke in such situations earlier than Method2. Hence, early detection of fire is possible with the proposed smoke detection method. In Movies 11 and 12, flames are not in the viewing range of the camera. A fire detection system without smoke detection capability fails in detecting the fire before it spread around.

5. CONCLUSION

A method for detecting smoke in video is developed. The algorithm is mainly based on determining the edge regions whose wavelet subband energies decrease with time. These regions are then analyzed along with their corresponding

background regions with respect to their RGB and chrominance values. The flicker of the smoke and convexity of smoke regions are also set as clues for the final decision.

The method can be used for detection of smoke in movies and video databases as well as real-time detection of smoke. It can be incorporated with a surveillance system monitoring an indoor or an outdoor area of interest for early detection of fire. It can also be integrated with the flame detection method in [12] in order to have a more robust video based fire detection system.

REFERENCES

- [1] W. Phillips III, M. Shah, and N. V. Lobo, "Flame Recognition in Video," *Pattern Recognition Letters*, v.23(1-3), pp.319-327, Jan. 2002.
- [2] Fastcom Tech.SA, Blvd. de Grancy 19A, CH-1006 Lausanne, Switzerland, "Method and Device for Detecting Fires Based on Image Analysis," *Patent Coop. Treaty(PCT) Appl.No: PCT/CH02/00118, PCT Pubn.No: WO02/069292*.
- [3] G. Healey, D. Slater, T. Lin, B. Drda, A. D. Goedeke, "A system for real-time fire detection," *CVPR'93*, pp.15-17, 1993.
- [4] S. Mallat, S. Zhong, "Characterization of Signals from Multiscale Edges," *IEEE Trans. on PAMI*, v.14 n.7, pp.710-732, July 1992.
- [5] A. E. Cetin and R. Ansari, "Signal Recovery from Wavelet Transform Maxima," *IEEE Trans. on Signal Proc.*, v.42 (1994), pp.194-196.
- [6] D. S. Chamberlin and A. Rose *The First Symposium (Int.) on Combustion, The Combustion Institute, Pittsburgh*, pp.27-32, 1965.
- [7] C. Stauffer and W. E. L. Grimson "Adaptive Background Mixture Models for Real-Time Tracking," in *Proc. of IEEE Computer Soc. Conf. on CVPR*, v.2, 1999.
- [8] R. T. Collins, A. J. Lipton, T. Kanade, "A System for Video Surveillance and Monitoring" *Proc. of American Nuclear Society 8th Int. Topical Meeting on Robotics and Remote Systems*, Pittsburgh, PA, Apr.25-29, 1999.
- [9] M. Bagci, Y. Yardimci, and A. E. Cetin, "Moving Object Detection Using Adaptive Subband Decomposition and Fractional Lower Order Statistics in Video Sequences," *Elsevier, Signal Proc.*, pp.1941-1947, Dec. 2002.
- [10] C. B. Liu and N. Ahuja, "Vision Based Fire Detection," in *Proc. of Int. Conf. on Pattern Recognition, ICPR '04*, Vol. 4, 2004.
- [11] T. Chen, P. Wu and Y. Chiou, "An Early Fire-Detection Method Based on Image Processing," in *Proc. of IEEE ICIP '04*, pp. 1707-1710, 2004.
- [12] Y. Dedeoglu, B. U. Toreyin, U. Gudukbay and A. E. Cetin, "Real-time Fire and Flame Detection in Video," in *Proc. of IEEE ICASSP '05*, pp. 669-672, 2005.