Spatio-Temporal Analysis in Smoke Detection

Chen-Yu Lee, Chin-Teng Lin, *Fellow, IEEE*, and Chao-Ting Hong

Department of Electrical and Control Engineering
National Chiao Tung University, Hsinchu, Taiwan, ROC
tbrsun.ece97g@nctu.edu.tw, ctlin@mail.nctu.edu.tw, and chaotien@ieee.org

***Abstract*— Smoke detection in video surveillance images has been studied for years. However, given an image in open or large spaces with typical smoke and the disturbance of commonly moving objects such as pedestrians or vehicles, robust and efficient smoke detection is still a challenging problem. In this paper, we present a novel and reliable framework for automatic smoke detection. It exploits three features: edge blurring, the gradual change of energy and the gradual change of chromatic configuration. In order to gain proper generalization ability with respect to sparse training samples, the three features are combined using a support vector machine based classifier. This system has been run more than 6 hours in various conditions to verify the reliability of fire safety in the real world.**

I. Introduction

Owing to the limitation of the traditional concept, point-based detectors can't detect fires or smokes in early stage. In recent years many researches are devoted to video smoke detection that doesn't rely on proximity of smoke to the detector. This enables it to incorporate standard video surveillance cameras with sophisticated image recognition and processing software to identify the distinctive characteristics of smoke patterns. In most cases, smoke usually appears before ignition. Therefore, the beginning of fire can be observed rapidly before it causes any real damage.

There are four categories of video smoke detection in the literature. The first category is *Motion-Based* approaches. Kopilovic et al. [1] observed that the irregularities in motion due to non-rigidity of smoke. They apply a multiscale optical flow computation and the entropy of the motion distribution in Bayesian classifier to detect the special motion of smoke. In order to save computational time, Yuan [2] proposed a fast orientation model that produces more effective way to extract the motion characteristics. Although significant advances have been made in the development of this work, their adoption in general surveillance systems is not widely reported.

The second one is *Appearance-Based* approaches. Toreyin et al. [3] indicated that smoke of an uncontrolled fire expands in time which results in regions with convex boundaries. Chen [4] found that airflows will make the shape of smoke to be variously changed at any time. Therefore, a disorder measure, the ratio of circumference to area for the extracted smoke region, is introduced to analyze shape complexity. Growth rate is obtained by increment of smoke pixels due to the diffusion process existed in generation of smoke. Two thresholds are determined by the statistical data of experiments to verify the real smoke; furthermore, the changing unevenness of density distribution is proposed in [5]. The difference image provides a natural way to represent the attribute which has more internal information in smoke frames than non-smoke frames. While much research has been devoted to these techniques, few studies have investigated the situation that smoke and non-smoke objects exist in the same time and the presence of moving objects from the outside of video scenes.

The third one is *Color-Based* approaches. Smoke usually displays grayish colors during the burning process [4]. Two thresholds of I (intensity) component of HSI color space depend on statistical data and this implies that three components R, G and B of the smoke pixel are equal or so. Since smoke color can't be represented accurately by a single unimodal, the 3D joint probability density function can be decomposed in three marginal unidimensional distributions over each color axis to accommodate different ranges of color [6]. Independent of the fuel type, smoke naturally decreases the chrominance channels U and V values of the candidate region [3]. In spite of the early alarm capability, few experimental results have been conducted in the range of grayish or dull non-smoke objects.

The fourth one is *Energy-Based* approaches. It is well-known that wavelet coefficients contain the high frequency information of the original image [7]. Since smoke obstructs the texture and edges in the background of an image [3], a decrease in wavelet energy is an important clue for smoke detection. Piccinini et al. [8] further improved the concept by on-line modeling the ratio between the current input frame energy and the background energy. This method performs well in many real cases but needs long reaction time and more exact validation of the input data extracted from surveillance systems that operate 24 hours a day.

The objective of this paper is to analyze the characteristic of smoke in spatial and temporal domains. The results obtained from this novel approach would provide better insight to operators in the field of smoke detection to handle the problems of high false alarm rate and long reaction time.

978-1-4244-5561-4/09/$26.00 ©2009

II. BLOCK PROCESSING

A. Background Modeling

Adaptive Gaussian Mixture Model (GMM) is a reliable method to approximate the background modeling process [9]. It models each pixel as a mixture of Gaussians and uses an on-line EM algorithm to update the model. This technique deals properly with lighting changes, repetitive motion from clutter, and long-term scene changes.

B. Candidates Selection

Smoke regions come into existence and disappear continuously because of the special particle property during ignition and combustion. It is inefficient to track or analyze the target using object-based method. Block-based technique provides a better way to solve this problem. The image will be divided into non-overlapped blocks, and each block has the same size in a same image. First, we will find out the blocks with a gray-level change. The current image will be subtracted from the background to get the difference image, and we compute its summation for each block.

Foreground regions can be found by background subtraction, but they could include static objects. Next, temporal difference of two successive frames will be calculated. We also compute its summation for each block to determine the moving property. In order to reduce the computational cost, only when the value of background subtraction and temporal difference lager than the predefined thresholds will be regarded as candidates containing moving objects. We consider the information of a particular block over time as a “block process” in the following sections.

III. FEATURE EXTRACTION

A. 2-D Spatial Wavelet Analysis

Unlike the Fourier transform, which describes the original signal in frequency domain, wavelet transforms can reveal in both frequency and time (location). In spatial wavelet analysis, images often contain approximations and details, which stand for the low-frequency and the high-frequency components respectively.

Because smoke blurs the texture and edges in the background of an image, high-frequency information becomes much more invisible when smoke covers part of the scene. Therefore, details will be an important indicator of smoke due to the decrease in value of high-frequency information. Energy of details is calculated for each candidate block:

$$E(B_k, I_t) = \sum_{i,j \in B_k} \left[LH(i,j)^2 + HL(i,j)^2 + HH(i,j)^2 \right] \tag{1}$$

where B_k is the k^{th} block of the scene, I_t is the input image at time t. Wavelet sub-band images LH, HL and HH are horizontal, vertical and diagonal details.

Instead of using energy of the input directly, we prefer computing the energy ratio of the current frame to the background model due to the cancelation of negative effect on different conditions and the capability of impartial measurement in the decrease:

$$\alpha(B_k) = \frac{E(B_k, I_t)}{E(B_k, BG_t)} \tag{2}$$

where BG_t is the mean value of the distribution with a highest weight in the GMM background model. The value of the energy ratio α is our first feature in spatial domain, which supports the fact that the texture or edges of the scene observed by the camera are no longer visible as they used to be in the current input frame.

B. 1-D Temporal Energy Analysis

Ordinary moving objects such as pedestrians or vehicles have solid characteristic so we can’t see details behind through the bodies. If there is an ordinary moving object going through the candidate block then there will be a sudden energy drop because of the transition from the background to the foreground object. On the contrary, initial smoke has semi-transparent nature and becomes less visible with the time going.

A gradual change of energy is guaranteed to this process and any abrupt variation will be regarded as a noise caused by common disturbance. One-dimension temporal wavelet analysis of energy ratio α provides a proper evaluation of this phenomenon. We obtain high-band (details) and low-band (approximations) information by the 1-D DWT shown in Fig.1. Therefore, the disturbance can be measured by computing the summation of details for a predefined time interval. The likelihood of the candidate block to be a smoke region is in inverse proportion to the parameter β

$$\beta(B_k) = \frac{\sum_n |D[n]|}{n} \tag{3}$$

where D[n] is the high-frequency information of energy ratio α and n is the number of time with a non-zero value of details.

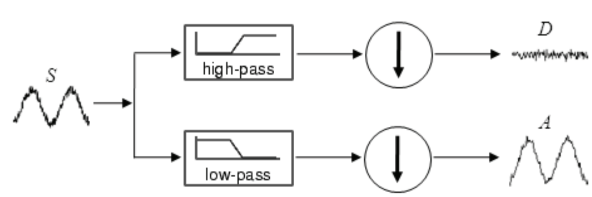


Fig. 1. One-dimensional DWT

C. 1-D Temporal Chromatic Configuration Analysis

Smoke can't be defined by a specific color appearance. However, it is possible to characterize smoke by considering its effect on the color appearance of the region on which it covers. Besides the gradual change of energy, smoke has the same property of color configuration.

Photometric invariant features are functions describing the color configuration of each image coordinate discounting local illumination variations. Hue and saturation in the HSV color space and the normalized-RGB color space are two photometric invariant features in common use. We decided to use the normalized-RGB color space for its fast computation. The transfer function is given by

$$r = \frac{R}{R+G+B},\quad g = \frac{G}{R+G+B},\quad b = \frac{B}{R+G+B} \tag{4}$$

therefore the *rgb* color system can be obtained by projecting a color vector in the RGB cube into a point on the unit plane described by $r + g + b = 1$.

From the empirical analysis, smoke lightens or darkens each component in RGB color space of the covered point but smoke doesn't severely change the values of the *rgb* color system. However, the values are likely to change in case of a material change. This constrain can be represented by

$$\begin{aligned} r(x,y,t) &\cong r(x,y,t+\Delta t) \\ g(x,y,t) &\cong g(x,y,t+\Delta t) \\ b(x,y,t) &\cong b(x,y,t+\Delta t) \end{aligned} \tag{5}$$

when the candidate blocks covered by smoke region instead of ordinary moving objects.

The details (high-frequency information) of the three channels in the *rgb* color system are obtained by the 1-D DWT again in Fig.1. We can obviously see that ordinary solid moving objects produce a great quantity of details in Fig.2a. Smoke has smooth variation in *rgb* color space and produces few details shown in Fig.2b. Therefore, the third feature ρ will be calculated by

$$\rho(B_k) = \max_{n \in \text{interval}} \left\| \left(D_r[n]\ ,\ D_g[n]\ ,\ D_b[n] \right) \right\|_2 \tag{6}$$

where $D_r[n]$, $D_g[n]$, and $D_b[n]$ stand for details of *r*, *g* and *b* channels respectively and the value of *r*, *g* and *b* are averages of candidate blocks.

IV. Classification and Verification

The difficulty in acquiring smoke or fire accident video should be concerned with practicality. Consequently, the complementary characteristic of the three features extracted from candidate blocks must be learned by a powerful classification model with robust generalization ability. Support Vector Machines (SVMs) have considerable potential as classifiers of sparse training data. This approach seeks to find the optimal separating hyperplane between classes by focusing on the training cases that lie at the edge of the class distributions, the support vectors, with the other training cases effectively discarded [10]. Here the training data represented by $\{x_i\ ,\ y_i\}$, $i = 1,\ldots,r$, are mapped into a higher dimensional space by the radial basis function (RBF) and the decision functions can be calculated by

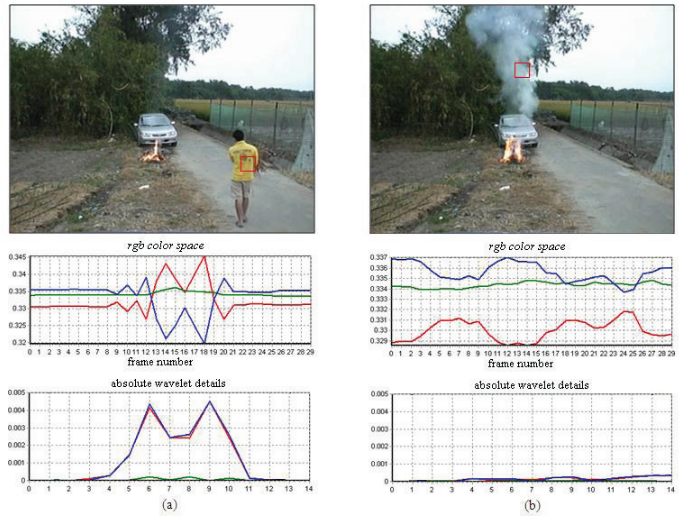


Fig. 2. Comparison of Changes in *rgb* Color Space at the Passage of Objects
a Ordinary moving objects
b Smoke

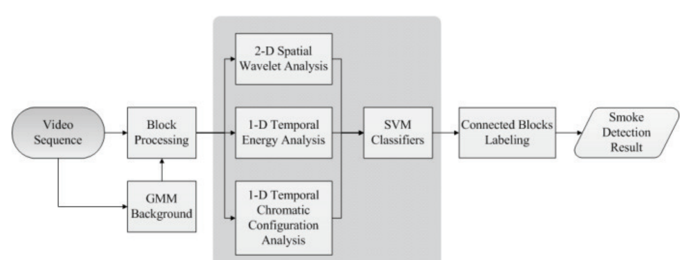
$$f(x) = \operatorname{sgn}\left(\sum_{i=1}^{r} \lambda_i y_i k(x, x_i) + b \right) \tag{7}$$

where λ_i, $i = 1,\ldots, r$ are Lagrange multipliers, b is a bias term and $k(x,x_i)$ denotes RBF kernel function.

The further verification for the presence of smoke is connected blocks labeling, which scans an image and groups its blocks into components based on block connectivity. Finally, we mark each component with red color when over 25% of its blocks are classified as smoke and mark with green color otherwise. The system overview is shown in Fig.3.

V. Results

We implemented our algorithm with a Intel Core2 Due



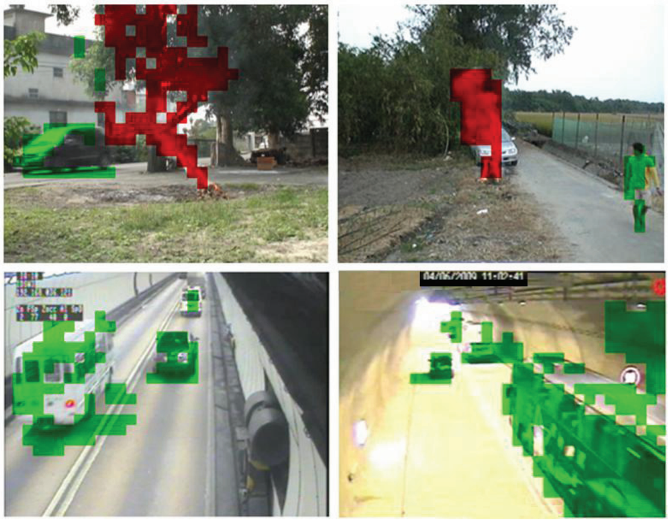


Fig. 3. Block Diagram of Smoke Detection System

Fig. 4. Experimental Results. The Red Blocks are Detected as Smoke and the Green Blocks are Detected as Ordinary Moving Objects.

TABLE I
EXPERIMENTAL RESULTS OF THE PROPOSED SYSTEM BASED ON SINGLE FRAME

	Detection Rate	False Alarm Rate	Reaction Time (sec)
2-D Spatial Wavelet Analysis	93.5%	38.0%	-
1-D Temporal Energy Analysis	91.7%	13.1%	-
1-D Temporal Chromatic Configuration Analysis	85.5%	11.2%	-
Global Analysis	**85.2%**	**1.7%**	**0.86**

2.2GHz processor and Borland C++ Builder. The processing time per frame is about 30 milliseconds for frames with sizes of 320 by 240 pixels. A large variety of conditions are tested including indoor, outdoor and sunlight variation each containing smoke events, pedestrians, bicycles, motorcycle, tourist coaches, trailers, waving leaves, etc. In our experiments, there are 18352 positive samples (smoke events) and 81503 negative samples (ordinary moving objects).

Data in Table I show the evaluations of each feature and the global testing result. The reaction time is obtained by the ratio between frames to detect and frames per second. 2-D spatial wavelet analysis can successfully extract candidate blocks with energy drop. However, pedestrians wearing flat clothing or long vehicles with flat roofs also produce energy drop. To overcome this drawback, we use 1-D temporal energy analysis which can express the gradual change of the energy ratio in smoke regions. This approach can adequately simulate the temporal characteristic of smoke. In some real cases, background model and foreground objects are so flat that there is no apparent high frequency information. It is difficult to separate smoke from non-smoke regions in this situation so we use the 1-D temporal chromatic configuration analysis to further describe the smoke's behavior and this feature operates properly. Although the detection rates are desirable by using three features individually, we are not satisfied with the false alarm rates. The SVM classifier can learn the complementary relationship among three features and gains the extremely low false alarm rate without losing the detection rate.

Most of other's algorithms are only seeking higher detection rate. It does not provide enough information on how accurate the system might be. When considering accuracy of smoke detection, people care more about how to decrease the false alarm rate and detect smoke events quickly, rather than just increase the detection rate. Here the false alarm rate of the proposed system is significantly lower than other's and the reaction time is extremely short. This system can also detect correctly even when both smoke and non-smoke objects exist in the same frame due to block processing while other systems only detect whether there is smoke existing in the whole video or the single frame.

This paper demonstrates a robust and efficient system for smoke detection, and it involves the spatial and temporal analysis for each candidate block and the SVM classifications. Experimental results show the opportunity of the real-time operation of surveillance systems and advanced applications.

REFERENCES

[1] I Kopilovic, B. Vagvolgyi and T. Sziranyi, "Application of panoramic annular lens for motion analysis tasks: surveillance and smoke detection," *In Proc. ICPR*, Barcelona, 2000.

[2] Feiniu Yuan, "A fast accumulative motion orientation model based on integral image for video smoke detection," *Pattern Recognition Letters*, vol. 29, pp.925-932, May 2008.

[3] B. U. Toreyin, Y. Dedeoglu, and A. E. Cetin, "Wavelet based real-time smoke detection in video," in *EUSIPCO '05*, 2005.

[4] Thou-Ho(Chao-Ho) Chen, Yen-Hui Yin, Shi-Feng Huang and Yan-Ting Ye, "The smoke detection for early fire-alarming system base on video processing," *Proceeding of the 2006 international Conference in intelligent information Hiding and Multimedia Signal Processing (IIH-MSP'06)*.

[5] Jing Yang, Feng Chen and Weidong Zhang, "Visual-based smoke detection using support vector machine," *Fourth International Conference on Natural Computation*, vol. 4, pp.301-305, ICNC '08, 2008.

[6] Capitan. J, Mantecon. D, Soriano. P and Ollero. A, "Autonomous perception techniques for urban and industrial fire scenarios," *IEEE International Workshop on Safety, Security and Rescue Robotics*, pp.1-6, SSRR 2007.

[7] A. E. Cetin and R. Ansari, "Signal recovery from wavelet transform maxima," *IEEE Trans. on Signal Proc.*, vol.42 (1994), pp.194-196.

[8] Piccinini. P, Calderara. S and Cucchiara. R, "Reliable smoke detection in the domains of image energy and color," *15th IEEE International Conference on Image Processing*, pp.1376-1379, ICIP 2008.

[9] Stauffer. C, Grimson. W.E.L, "Adaptive background mixture models for real-time tracking," *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, vol. 2, June 1999.

[10] G. Mercier and M. Lennon, "Support vector machines for hyperspectral image classification with spectral-based kernels," in *Proc. IGARSS*, Toulouse, France, July 21–25, 2003.

[11] Http://www.openvisor.org