Vision Based Fire Detection Using Mixture Gaussian Model

FEINIU YUAN[1], GUANGXUAN LIAO[1], WEICHENG FAN[1], and HEQIN ZHOU[2]
[1] State Key Lab of Fire Science
University of Science and Technology of China
Hefei 230027, Anhui, China
[2] School of Information Science and Technology
University of Science and Technology of China
Hefei 230027, Anhui, China

ABSTRACT

Vision based fire detection has many advantages over traditional methods. In vision based fire detection approaches, it is required that systems must have enough robustness and be insensitive to environment. We mainly take advantage of mixture Gaussian model and frame difference techniques to adaptively extract a background image from image sequences captured by ordinary color cameras. These techniques are able to mostly eliminate influences of artificial lights, wind and moving objects disturbance. By subtracting the background image from the incoming frame, foreground objects which are possible fire pixels are thus obtained. After analyzing behavior and spectroscopy of fire, color, shape fluctuation and growth rate are used to determine if a possible pixel is an actual fire pixel. Experiments show that our algorithm is robust for a stationary camera.

KEYWORDS: vision based fire detection, mixture Gaussian model, background subtraction, experimental testing

INTRODUCTION

Many existing fire detectors are based on particle sampling, temperature sampling, relative humidity sampling and smoke analysis [1]. Those sensors are applied very widely due to low cost and simplicity, but those sensors must be installed in the close proximity of fire, otherwise they can not detect fire at all. That's one of major drawbacks of traditional fire detection techniques. So those detectors can not be applied in large spaces. And detectors often get fouled and disabled by dust or other non-fire particles due to direct contiguity with them. Some detectors adopt infrared sensors and that leads to high cost for surveillance. To avoid fire disasters and reduce the cost of surveillance system, a lot of novel techniques have been proposed to improve the performance of fire detection. One of them is vision based fire detection. Vision based fire detection approaches not only detect fire far from the detectors, but also provide additional information about fire, such as location, size, growth rate, and so on.

Up to now, however, some issues must be addressed, such as instability, low accuracy when there is slightly strong wind or there exists complicated artificial lighting condition. To solve these problems, Yamagishi *et al* [2,3] presented a fire flame detection method based on a space-time fluctuation data of the contour of flame region. But the method only used color information to segment fire pixels. We must notice that many non-fire objects share the same color with fire. Though accurate color parameters can be acquired by learning, false segmentation exists inevitably. False segmentation severely influences

subsequent decision. Noda *et al* [4] developed a specified fire detection based on gray scale images applied in tunnels. They analyzed the relationship between temperature and G/R ratio, and used gray level histogram features to recognize fire. Only using one dimensional histogram is not capable of adapting to general fires. Shah *et al* [5] proposed a flame recognition technique based on video. They used Gaussian-smoothed color histogram to generate a color lookup table of fire pixel and then took advantage of temporal variation of pixels to determine if a pixel is actually a fire one. The approach is based on each frame, so it is insensitive to camera motion. But it can not segment fire pixels very well when there are objects sharing the same color with fire, and it is not capable of locating the position of fire.

Due to wide application of inexpensive color cameras in military and commercial securities [6], we use a stationary color camera to monitor experimental fires. In this paper, a background image is estimated by mixture Gaussian model. The background image is adaptive to environmental changes. Possible fire pixels are obtained by subtracting the background image from the current input frame, instead of directly extracting possible fire pixels from each frame. The technique can eliminate disturbances of many still objects sharing the same color with fire. And then color, shape fluctuation and growth rate features are used to decide if a fire alarm should be given. This paper is organized as follows. Firstly, segmentation of foreground pixels using an adaptive background model is presented in detail, and then these foreground pixels are further examined according to color and temporal features. Secondly, several experiments are described and our future work is also outlined. At last, this paper is concluded.

SEGMENTATION OF POSSIBLE FIRE PIXELS

Adaptive Background Model

We consider the values of a particular pixel over time as a stochastic process [7], i.e., a time series of scalars for gray values or vectors for color pixel values. At any time, color value for each pixel belongs to one of several Gaussian distributions when there is no fire or moving objects in the scene monitored by surveillance cameras. Let us to observe color values over time at the center of the video in an experimental fire. As shown in Fig. 1, color values at the same position marked by a red cross are automatically recorded per frame by our developed software. The top row shows four different frames captured by a color CCD camera where a red cross represents the location to be tracked over time, the middle row illustrates the tracking results of color value in the HSV color space and the bottom row gives the tracking results in the RGB color space. The RGB color model is widely used in computer monitors but it is nonlinear to our visual perception. The RGB tricolor model is based on light emission theory. R, G and B stand for red, green and blue, respectively. The HSV color model is subconsciously used by artists. H, S and V denote hue, saturation and intensity value, respectively. So it is easy to be understood with regard to each component. When there is no fire, pixel values (R, G, B) or (H, S, V) match a Gaussian distribution with a low mean value. Once fire is just burning up, pixel value changes violently at once and leaps into another Gaussian with a high mean value. That's one of the major statistical differences between fire and non-fire. We just use these statistical properties to extract foreground objects.
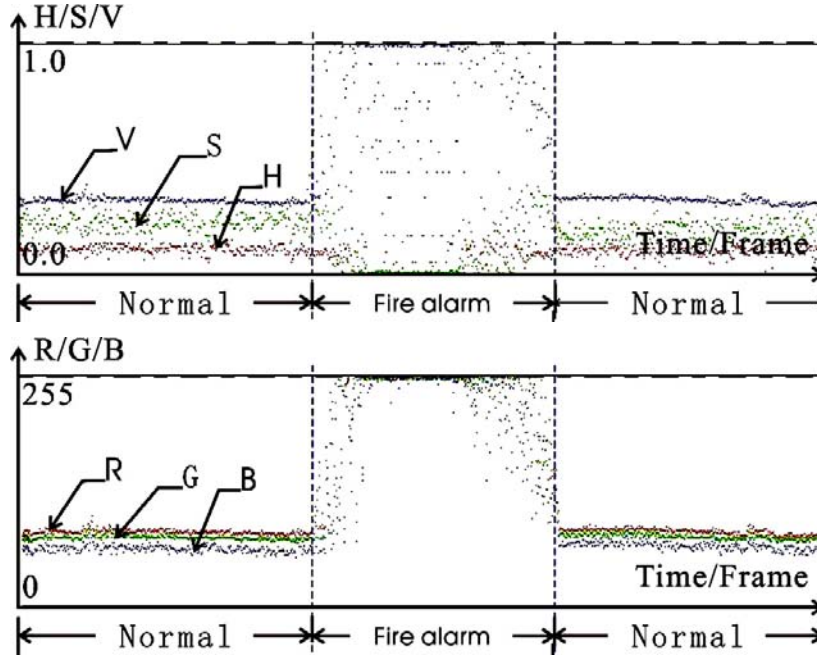
1576

Fig. 1. Tracking color values at the center of the video
over time in HSV and RGB color spaces.

If only lighting condition changes over time, a single adaptive Gaussian per pixel would be sufficient. When the background becomes complicated, such as sway of tree leaves, multiple Gaussians are required. We model the recent history of each pixel, $\{\mathbf{X}_1, \mathbf{X}_2, \ldots, \mathbf{X}_t\}$, as a mixture of K Gaussian distributions [8]. The probability of observing the incoming pixel value is

$$P(\mathbf{X}_t) = \sum_{i=1}^{K} w_{i,t} \cdot \eta(\mathbf{X}_t, \boldsymbol{\mu}_{i,t}, \Sigma_{i,t}) \tag{1}$$

Where K is the number of distributions, $w_{i,t}$ is the normalized weight of the ith Gaussian at time t, $\boldsymbol{\mu}_{i,t}$ is the mean color value, $\Sigma_{i,t}$ is covariance matrix of the ith Gaussian at time t, and η is a Gaussian density function as follows.

$$\eta(\mathbf{X}, \boldsymbol{\mu}, \Sigma) = \frac{1}{(2\pi)^{D/2} |\Sigma|^{1/2}} e^{-\frac{1}{2}(\mathbf{X}-\boldsymbol{\mu})^T \Sigma^{-1}(\mathbf{X}-\boldsymbol{\mu})} \tag{2}$$

1577

Where D is the dimension of color values, D often equals 3 for chromatic pixels and 1 for gray scale pixels. K is determined by the available memory and computational power, currently from 3 to 5. In order to simplify the computation, and the covariance matrix is assumed to be diagonal matrix (Eq. 3).

$$\Sigma_{i,t} = \begin{bmatrix} \sigma_{i,1}^2 & 0 & 0 \\ 0 & \sigma_{i,2}^2 & 0 \\ 0 & 0 & \sigma_{i,3}^2 \end{bmatrix} \tag{3}$$

For true color videos, we can process video in RGB or HSV color space. So stochastic variable vector $\mathbf{X}$ is defined as

$$\mathbf{X} = \begin{bmatrix} R \\ G \\ B \end{bmatrix} \quad \text{or} \quad \mathbf{X} = \begin{bmatrix} H \\ S \\ V \end{bmatrix} \tag{4}$$

After analyzing the model, another problem is how to specify parameters of the model at the beginning. We use the expectation maximization algorithm (EM) to automatically learn the parameters, $w_{i,t}$, $\mu_{i,t}$ and $\Sigma_{i,t}$, of the K Gaussian model. Because the EM algorithm [9] is computationally extensive, we can do it offline since there is no fire most of time and environment often changes slowly.

Model Matching and Updating

Before matching, the K distributions are ordered in descending order based on each priority value V_p (Eq. 5). And then the algorithm selects the first B distributions that account for a predefined fraction of the evidence T_w. The first B distributions are considered as background distribution while the remaining K-B distributions are considered as foreground distribution.

$$B = \arg\min_b \{ \sum_{i=1}^{b} w_{i,t} > T_w \} \tag{5}$$

$$V_p = w_{i,t} \cdot |\Sigma_{i,t}|^{-1/2} = \frac{w_{i,t}}{\sigma_{i,1} \cdot \sigma_{i,2} \cdot \sigma_{i,3}} \tag{6}$$

Every input pixel value of current frame, $\mathbf{X}_t$, is checked against the existing K Gaussian distributions until a match is found. To simplify computation, a match is defined as a pixel value within 2.5 standard deviation of a distribution instead of calculating the time-consuming Eq. 1. A deviation of a Gaussian is defined as the Mahalanobis distance r_t.

$$r_t^2 = (X_t - \mu_{i,t})^T \cdot \Sigma_{i,t}^{-1} \cdot (X_t - \mu_{i,t}) \tag{7}$$

If a match is found and the matched distribution m is within the first B distributions, the incoming pixel can be considered as a background pixel, otherwise foreground. And we must update the matched distribution parameters using the following equations.

$$w_{i,t} = (1-\alpha)w_{i,t-1} \quad i=1, \dots, K \quad \text{and} \quad i \neq m$$

$$w_{i,t} = (1-\alpha)w_{i,t-1} + \alpha$$
$$\boldsymbol{\mu}_{i,t} = (1-\alpha)\boldsymbol{\mu}_{i,t-1} + \alpha\mathbf{X}_t \qquad \text{otherwise} \qquad (8)$$
$$\Sigma_{i,t} = (1-\alpha)\Sigma_{i,t-1} + \alpha \cdot Diag\{(\mathbf{X}_t - \boldsymbol{\mu}_{i,t}) \cdot (\mathbf{X}_t - \boldsymbol{\mu}_{i,t})^T\}$$

Where α is the learning coefficient, relatively low value for fire case, $Diag()$ is an operator which makes a matrix diagonal by setting non-diagonal elements to 0 while keeping diagonal elements unchanged.

If no match is found, the least probable distribution is replaced by a Gaussian with the incoming value as its mean value, an initially high variance and low weight.

RECOGNITION OF FIRE EVENT

Possible fire pixels extracted in the previous section must be further checked to determine if there is a fire. A fire often has smoke and flame. Smoke pixels have low color saturation. Color of flame pixels often changes from red to yellow. When flame temperature is extremely high and materials have fewer impurities, blue color may be produced. When pixel value is within the range from red to yellow and blue color with low saturation and high intensity, the pixel may be considered as a fire pixel. These color features can discard a lot of no fire pixels. The shape of fire region varies from time to time. Temporal variance of shapes is adopted to exclude rigid objects that share the same color with fire. Therefore, possible fire pixels are checked in the following two stages.

COLOR FEATURES

If pixel values are within the fire color range, the pixel passes color testing. What we really care about is how to exactly determine the color parameters in the chroma circle (Fig. 2). We can also obtain accurate parameters by automatically learning from a large amount of video data captured from all kinds of natural fires, practical thermal disaster and experimental fires. In our algorithm, as shown in Fig. 2, we manually specified the color ranges of flame for simplicity. With our research going on, we will try to generate more general, complicated and accurate parameters for color features of fire, for example, establishing color database or look up table by statistical training and even building texture feature databases.
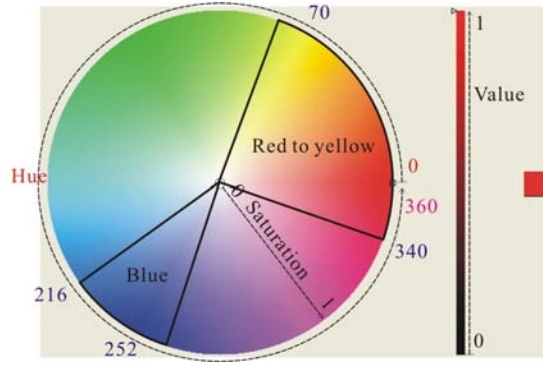
Fig. 2. Color range of fire pixels in the chroma circle.

Temporal Features

Pixels, which just passed the color testing, cannot be directly regarded as a fire pixel due to many things sharing the same color with fire. To further reduce false alarm, temporal features must be adopted for further identification. An important temporal characteristic is the fluctuation of shape. To measure this feature, we pay more attention to area and shape of fire region. Only early warning of fire is significant for fire suppression. When the area of fire region is larger than some threshold and then tends to expand, there are possible fire events occurring. Fire shape changes frequently over time, while shapes of sun, flashlight and other artificial light often change slowly.

The area can be calculated by counting the total number of pixels. So the area and growth rate of fire pixels can be expressed as

$$Area_t = \sum_{x=0}^{w-1}\sum_{y=0}^{h-1} Mask(x,y,t) \qquad (9)$$

$$GrowthRate_t = \frac{1}{N}\sum_{n=0}^{N-1} \frac{Area_{t-n} - Area_{t-n-1}}{Area_{t-n} + Area_{t-n-1}} \qquad (10)$$

where $Mask(x,y,t)$ is a boolean function where it returns 1 if the corresponding location is a fire pixel, 0 if not; N is the number of frames to average growth rate in order to filter high frequency noises.

The shape feature is measured by the distance between the edge points and gravity center of the shape, as shown in Fig. 3. In fact, it is another description of edge points (r, θ) in the polar coordinates system. For our test, we divide the edge into 8 points at the even angle interval (45 degrees). To improve the accuracy, a small angle interval may be used. The temporal fluctuation are defined as

$$ShapeFluctuation_t = \frac{1}{N}\sum_{n=0}^{N}\frac{1}{8}\sum_{i=0}^{7} \frac{(r_{t-n,i} - r_{t-n-1,i})}{(r_{t-n,i} + r_{t-n-1,i})} \qquad (11)$$
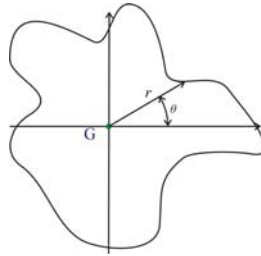
1580

Fig. 3. Polar coordinates features of shape.

Now, let's summarize the complete fire detection approach:

(1) Adaptively extract foreground pixels using mixture Gaussian model

(2) Recognition of fire pixels

      (a) Foreground pixels are testing by color values

      (b) Testing by some temporal features.

EXPERIMENTAL RESULTS

We have implemented our method using Visual C++ and performed several experiments. First, we test our approach on a cotton rope fire video. The fire consumed sixty pieces of ropes. As shown in Fig. 4, the top row shows source frames captured by a color CCD camera and the bottom row illustrates corresponding possible fire or smoke pixels extracted by our approach where pure blue color value stands for non-foreground pixels. From Fig. 4, we can find that the algorithm can extract fire pixels and even smoke pixels efficiently, and a fire alarm was given during early stage of the fire.
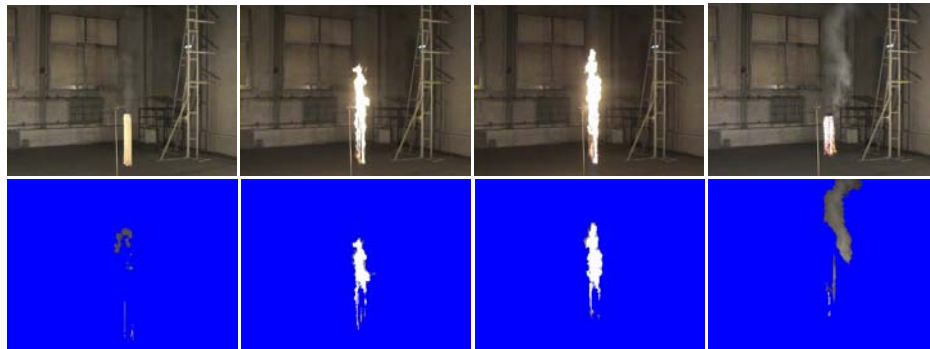

Fig. 4. Cotton cope fire.

We also have done another experiment of coal gas fire in an ordinary daily kitchen, as shown in the left of Fig. 5. The top left shows frames of coal gas fire captured by a color CCD camera and the bottom left illustrates corresponding possible fire pixels of the coal gas fire extracted by our approach. Pure blue color value denotes background pixels. Mixture Gaussian model can extract blue flame very well from the complex background. The blue flame passed the color and temporal testing, and a fire alarm was given when the number of flame pixels reached to the pre-specified threshold.
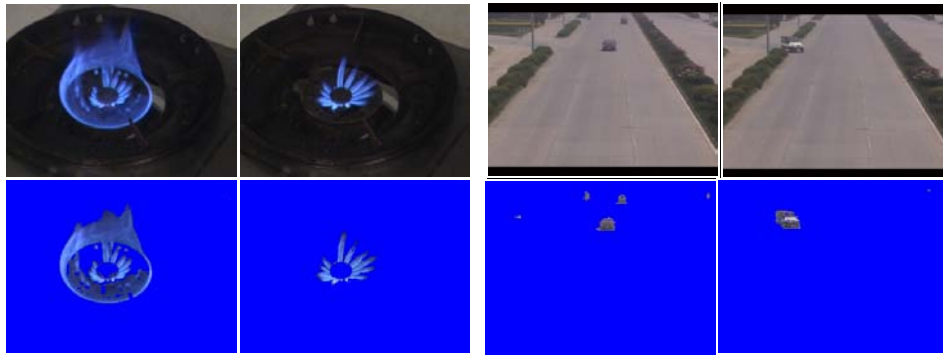
Fig. 5. Fire of coal gas and traffic video.

At last, we tested our method on a traffic video captured outdoors, as shown in the right of Fig. 5. The top right illustrates frames of traffic video captured by a color CCD camera and the bottom right gives corresponding non-fire pixels of moving vehicles segmented by our approach. The mixture Gaussian model exactly segments the foreground moving objects. But these moving objects failed to match the color testing.

FUTURE WORK

The work presented in this paper is only a primary phase of vision based fire detection. We will go on with the research by fully making use of computer vision, pattern recognition and fire science together. It is also possible to take advantage of texture features of smoke and flame for recognition. And two cameras may be used as binocular stereo for obtaining the position, size and expanding path of fires, which are very important for joint fire suppression. These directions are some choices of our future work.

CONCLUSIONS

Vision based fire detection has many advantages over conventional methods. In vision based fire detection approaches, it is required that the system must have enough robustness and be insensitive to environment. We mainly use mixture Gaussian model and frame difference techniques to extract a background image from video captured by color surveillance camera. Mixture Gauss model is able to partially eliminate disturbances of artificial lights, wind and moving objects. By subtracting the background image from the incoming frame, we can obtain the foreground objects that are possible fire pixels. Then, we use color, fluctuation, growth rate, and so on, to reduce false fire alarm. Experiments show that our algorithm is robust and insensitive to environment.

ACKNOWLEDGES

REFERENCES

[1]   Thou-Ho Chen, Cheng-Liang Kao, and Sju-Ma Chang, "An Intelligent Real-time Fire-detection Method Based on Video Processing," *IEEE 37th Annual 2003 International Carnahan Conference on Security Technology*, Oct. 14-16, 2003, pp. 104-111.

[2]   Yamagishi, H., and Yamaguchi, J., "Fire Flame Detection Algorithm Using a Color Camera," *Proceedings of 1999 International Symposium on Micromechatronics and Human Science*, Nov. 23-26, 1999, pp. 255-260.

[3]   Yamagishi, H., and Yamaguchi, J., "A Contour Fluctuation Data Processing Method for Fire Flame Detection Using a Color Camera," *IEEE 26th Annual Conference on IECON of the Industrial Electronics Society*, Vol. 2, Oct. 22-28, 2000, pp. 824-829.

[4]   Noda, S., and Ueda, K., "Fire Detection in Tunnels Using an Image Processing Method," *Proceedings of Vehicle Navigation and Information Systems Conferenc*, Aug. 31 - Sept. 2, 1994, pp. 57-62.

[5]   Phillips III, W., Shah, M., and Da Vitoria Lobo, N., "Flame Recognition in Video," *Fifth IEEE Workshop on Applications of Computer Vision*, Dec. 4-6, 2000, pp. 224-229.

[6]   Healey, G., Slater, D., Lin, T., Drda, B., and Goedeke, A.D., "A System for Real-time Fire Detection," *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, June 15-17, 1993, pp. 605-606.

[7]   Stauffer, C., and Grimson, W.E.L., "Learning Patterns of Activity Using Real-time Tracking," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 22, Iss. 8, Aug. 2000, pp. 747-757.

[8]   Pavlidis, I., Morellas, V., Tsiamyrtzis, P., Harp, S., "Urban Surveillance Systems: From the Laboratory to the Commercial World," *Proceedings of the IEEE*, Vol. 89, Iss. 10, Oct. 2001, pp. 1478-1497.

[9]   Dempster, A.P., Laird, N. M., Rubin, D.B., *Maximum Likelihood from Incomplete Data via the Em algorithm*, J. Roy Stat. Soc. B, Vol. 39, 1977, pp. 1-38.