VISION-BASED FIRE DETECTION USING VIDEO SEQUENCES

[1] *Yuan-Hsiang Chang* (*張元翔*), [2] *Tung-Che Hsieh* (*謝東哲*), [3] *Li-Der Jeng* (*鄭立德*)

[1] Dept. of Information & Computer Engineering,
Chung Yuan Christian University, Taiwan
[2, 3] Master Program in Communication Engineering,
Chung Yuan Christian University, Taiwan
E-mail: changyh@ice.cycu.edu.tw

ABSTRACT

Fire, if improperly used, could pose great threats to peoples' security, life, and property. Motivated by the requirement to detect fire at its early stage, we aimed to develop an automatic system for *vision-based fire detection using video sequences*. Our system included four major steps, namely *image preprocessing*, *foreground region analysis*, *fire dynamic behavior analysis*, and *fire flow energy analysis*. Overall, our system could achieve the detection rates of over 91% in either indoor or outdoor environments. In addition, our system could achieve the system response time within 1 second (average delay of ~25 frames) once the fire occurred. In summary, our system could be used in surveillance systems, leading to prevent damage to peoples' security, life, and property.

Keywords *Eddies Effect*; *Fire detection; Image processing; Video analysis;*

1. INTRODUCTION

Fire, if improperly used, may pose great threats to peoples' security, life, and property. Conventional fire detection technologies are generally based on physical/chemical properties of fire, such as smoke analysis, particle sampling, and temperature sampling. These technologies are often limited in detection distance and rely on fire alarms which are installed at the ceiling for monitoring. The fire alarms generally require the fire to burn for a while to create large amount of smoke to detect. Therefore, conventional fire alarms can't detect fire at its early stage, which may result in delay response. Besides, fire alarms can't be used at outdoor environments effectively.

Recently, with the development of image/video processing techniques, surveillance cameras are widely seen in many indoor/outdoor environments. These cameras can help monitor the environment for potential risk or crime that may cause damage to peoples' security, life, and property. With the advance of image/video processing techniques, smart video surveillance systems could be developed for fire detection using vision-based approaches.

Many vision-based fire detection techniques have been presented in past years. For example, Healey *et al*. [1] presented a real-time system for automatic fire detection using color video sequences. They used a color model for the fire detection algorithm. However, there are many objects which may exhibit similar color property of fire in real scenes. As a result, the system using color as the detection rule only may easily generate false alarms. Liu *et al*. [2] presented a vision-based fire detection system. The detection system was based on spectral, spatial, and temporal models of fire regions in image sequences. The spectral model was represented in terms of the color probability density of fire pixels. The spatial model was used to capture the spatial structure within a fire region. Further, the shape of a fire region was represented in terms of the spatial frequency of the region contour using its Fourier coefficients. Duong *et al*. [3] described the color property of a fire region as a bright-white, yellow, orange, red-like region that is extended from its interior core to the exterior contour. They used the fuzzy clustering technique to classify the fire regions into 5 different clusters using pixels' red, green, blue values. However, their method was only suitable for dark environments. Wei *et al*. [4] presented a method to detect smoke generated by fire using the Discrete Wavelet Transform (DWT) to estimate the edge energy of fire regions. Along with the increasing amount of smoke, the fire regions would become blurred. As a result, the high frequency components of the fire regions were decreased. Although the system could be used to detect the smoke successfully, the fire could have been burning terribly.

Motivated by the requirement to detect fire at its early stage, we aimed to develop an automatic system for *vision-based fire detection using video sequences*. Our system was developed using surveillance cameras to capture video sequences for fire detection. The method was based on image and/or video processing techniques that were designed to automatically analyze the fire (or non-fire) images in real scenarios.

2. METHOD

Our research assumptions include the following:

1. Fire has specific color property in digital images [5]. In this study, physical/chemical properties (e.g., temperature, smoke, etc.) of fire are not considered;
2. Fire will expand and spread gradually during the burning process. That is, fire won't suddenly shift its location in a short period of time;
3. Fire will create the *Flickering* phenomenon known as the Eddies Effect [6].

System limitations are described as follows:

1. Single surveillance camera is used and fixed at its location;
2. Fire region can't be occluded by other objects;
3. The fire region's size can't be too small.

Fig. 1 shows the flow chart of our system for the *vision-based fire detection using video sequences*. Our method includes four major steps: *image preprocessing*, *foreground region Analysis*, *fire dynamic behavior analysis*, and *fire flow energy analysis*. Detail descriptions of our system are described herein.

2.1. Image Preprocessing

The purpose of *image preprocessing* is to detect candidate fire regions in video sequences. The method is based on the *background subtraction* to extract foreground regions in motion from their surrounding background regions. Then, the *fire-like color filtering* is used to remove non-fire pixels. Further, the *morphological processing* is used to refine the contour of the foreground region. Finally, the *connected component labeling* is used to identify each foreground region for further analysis.

A. Background Subtraction

Background subtraction is an important segmentation technique for motion detection. First, we build the initial background model as follows. Let $B(x, y; 0)$ be the initial background model as defined by:

$$B(x, y; 0) = \frac{1}{N}\sum_{t=1}^{N} I(x, y; t) \tag{1}$$

where t is the frame index in the video sequence, and $I(x, y; t)$ is the intensity of pixel(x, y) in the t-th frame.

Then, the background subtraction is used to extract moving foreground pixels if they satisfy the following equation:

$$|\, I(x, y; t) - B(x, y; t-1)| > T(x, y; t) \tag{2}$$

where $T(x, y; t)$ is the threshold of pixel(x, y) in the t-th frame, and the threshold is defined as:

$$T(x, y, t) = c \times \sigma(x, y) \tag{3}$$

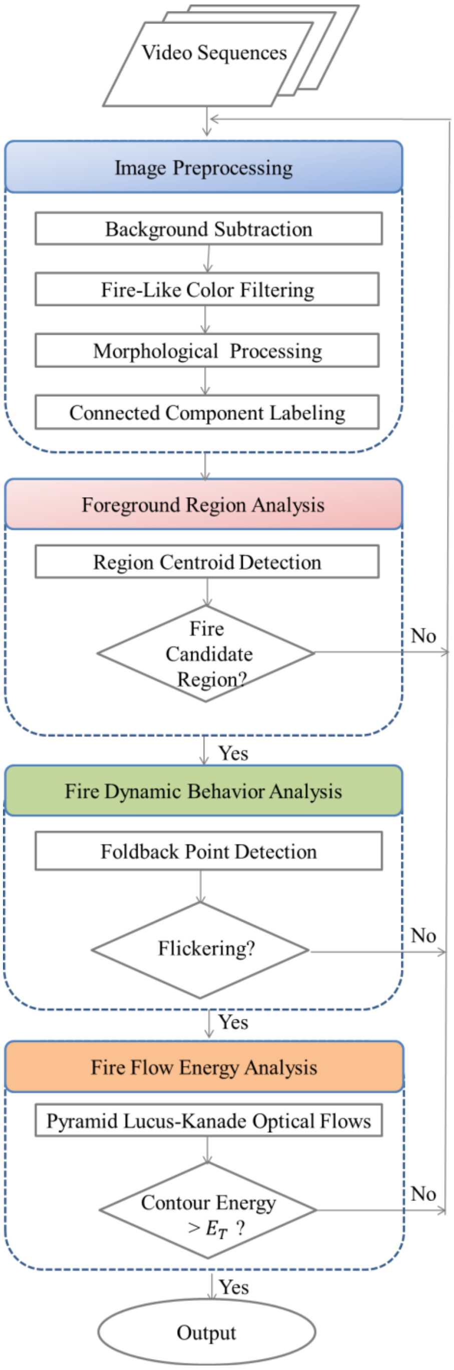


Fig. 1: System flowchart for the *vision-based fire detection using video sequences*.

where c is a constant and $\sigma(x, y)$ is the standard deviation as defined by:

$$\sigma(x, y) = \sqrt{\frac{1}{N}\sum_{t=1}^{N}[I(x, y; t) - B(x, y; 0)]^2} \tag{4}$$

The background information may change with respect to varying illumination of the surveillance environment. Therefore, we build and update the background model adaptively using the following equations:

$$B(x,y;t) = \begin{cases} \alpha I(x,y;t-1) + (1-\alpha)B(x,y;t-1), \\ \qquad\qquad\qquad if\ (x,y)\ is\ stationary \\ B(x,y;t-1), \\ \qquad\qquad\qquad otherwise \end{cases} \tag{5}$$

and

$$T(x,y;t) = \begin{cases} \alpha|I(x,y;t-1) - B(x,y;t-1)| + \\ \qquad (1-\alpha)T(x,y;t-1), \\ \qquad\qquad\qquad if\ (x,y)\ is\ stationary \\ T(x,y;t-1), \\ \qquad\qquad\qquad otherwise \end{cases} \tag{6}$$

where α is an updating parameter between 0 to 1.

An example is shown in Fig. 2. In a dark environment, a person was asked to light a fire to burn papers. Fig. 2 (a) is the original image, (b) is the adaptive background model, and (c) is the foreground regions after the background subtraction. As seen, the foreground regions in motion were detected accordingly.

(a)

(b)

(c)

Fig. 2: An example of the *background subtraction*: (a) original image; (b) background model; (c) foreground regions after the background subtraction.

B. Fire Color Filtering

The objective of *fire color filtering* is to remove non-fire-like pixels, while retaining candidate pixels of fire. Here, the color model as described in [5] is used. The fire color filtering is performed using two different color spaces, namely the RGB and the HSI color spaces.

In the RGB color space, the criterion for fire pixels is:

$$R \geq G > B \,\&\, R > R_T \tag{7}$$

where R, G, B are the red, green, blue values of the pixel, respectively. R_T is a selected threshold.

In the HSI color space, the criterion is:

$$\begin{cases} 0^\circ \leq H \leq 60^\circ \\ S \geq \left[(255 - R) * \frac{S_T}{R_T}\right] \\ I > I_T \end{cases} \tag{8}$$

where

$$S = 1 - \frac{1}{(R+G+B)}[\min(R,G,B)] \tag{9}$$

and

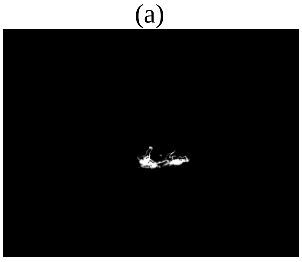
$$I = \frac{1}{3}(R+G+B) \tag{10}$$

where H, S, I are the Hue, Saturation, and Intensity, respectively. In addition, R_T, S_T, and I_T are selected thresholds.

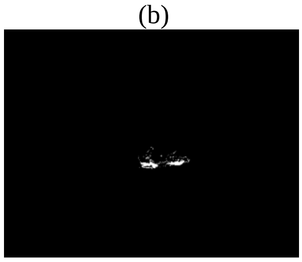
An example of the fire color filtering is shown in Fig. 3, where (a) is the original image, (b) is the foreground regions after the background subtraction, (c) contains the candidate fire regions after the RGB color filtering; (d) contains the candidate fire regions after the HSI color filtering.

(a)

(b)



(c)

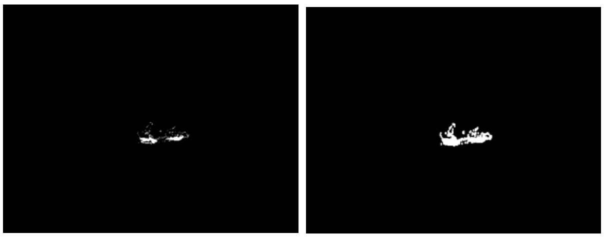


(d)

Fig. 3: An example of the *fire color filtering*: (a) original Image; (b) foreground regions after the background subtraction; (c) candidate fire regions after the RGB color filtering; (d) candidate fire regions after the HSI color filtering.

C. Morphological Processing

The candidate fire regions may have irregular and/or broken shapes (contours). Therefore, morphological processing is applied in order to refine the shape (contour) of the fire region. Because the fire region generally spreads/expands vertically, we use a vertical structuring element with the size of 3 ×5 pixels to refine the shape of the candidate fire regions. Here, the morphological-process, i.e., dilation, is selected for the purpose as shown in Fig. 4. Fig. 4 (a) is the candidate fire regions after the fire color filtering, (b) is the resulting fire region after the morphological processing, and (c) is the structuring element being used.



(a) (b)

1	1	1
1	1	1
1	1	1
1	1	1
1	1	1

(c)

Fig. 4: An example of the *morphological processing*: (a) candidate fire regions after the fire color filtering; (b) candidate fire regions after the morphological processing; (c) structuring element used for dilation.

D. Connected Component Labeling

The *Connected Component Labeling* [7] is used to label sets of pixels that are connected to each other by eight-connectivity, such that the sets of pixels share a unique label. As a result, sets of pixels belonging to their corresponding connected components can be identified. An example of the connected component labeling is shown in Fig. 5, where (a) is the candidate fire regions, and (b) is the connected component as marked with a rectangular region.

(a) (b)

Fig. 5: An example of the *connected compoenent labeling*: (a) image containing the candidate fire region; (b) the fire region is identified as a connected component (rectangular region).

2.2. Foreground Region Analysis

While fire is burning, the motion of the fire region is not like normal moving objects (e.g., rigid or articulated objects). That is, during the burning process, the fire region will not suddenly shift its location in a short period of time. In our system, this property of fire region is applied as a criterion to distinguish between fire and non-fire regions.

Assuming the candidate region is not likely a fire region, the region centroid of the moving object will generally shift its location along a path away from its initial location. An example of a candidate region not likely a fire region is shown in Fig. 6.

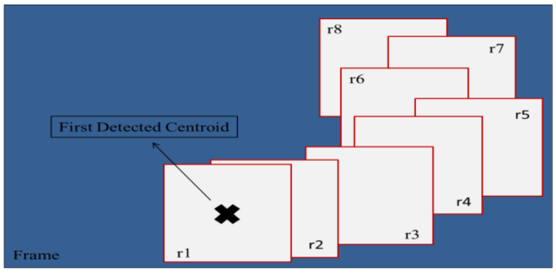


Fig. 6: An example of a candidate region not likely a fire region. The region centroid of a moving object will shift its location along a path away from its initial location, where r1, r2 stand for the candidate regions as detected at the t-th frame and the $t + 1$-th frame, etc. The location of the first detected centroid is marked, and will likely shift its location during motion.

Assuming the candidate region is actually a fire region, the region centroid will generally remains at its initial location in a short period of time during the fire burning process. An example of a candidate fire region is shown in Fig. 7.

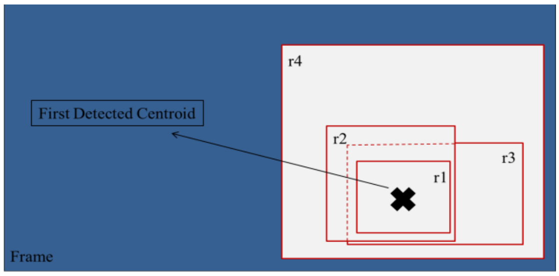
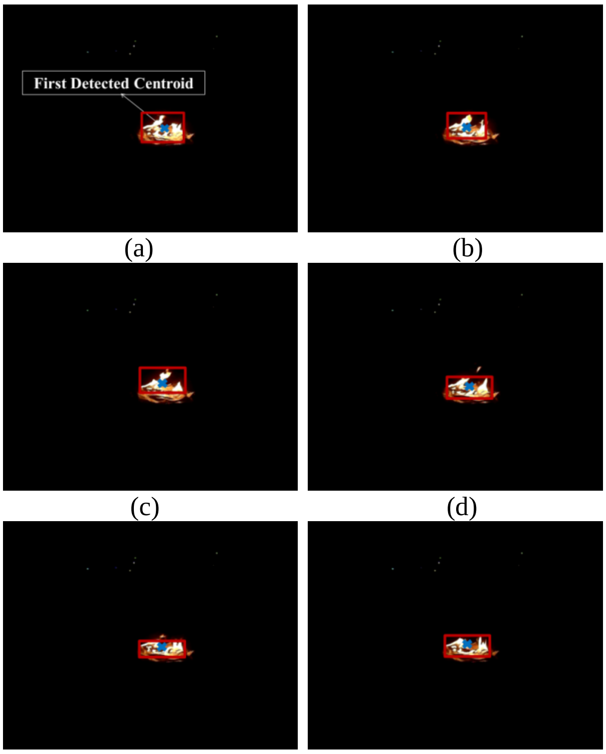


Fig. 7: An example of a candidate fire region. In a short period of time, the region centroid will remains at its initial location, where r1, r2 stand for the candidate regions as detected at the t-th frame and the $t + 1$-th frame, etc. The location of the first detected centroid is marked.

The property described above is suitable to describe real fire burning process. An example is shown in Fig. 8. While the fire is burning, the region centroid will likely remain at its initial location in adjacent frames.



(a) (b)
(c) (d)
(e) (f)

Fig. 8: An example of the fire region during the fire burning process: (a)~(f) are adjacent frames. Here, the first detected region centroid is marked (blue). The fire regions detected at each frame are marked (red).

2.3. Fire Dynamic Behavior Analysis

The Fire dynamic behavior analysis is developed according to the third assumption, i.e., fire will create the *Flickering* phenomenon known as the Eddies Effect. Eddies Effect is caused by the buoyant force, which causes eddies by air entrainment during the fire burning process. The schematic diagram of the Eddies effect is shown in Fig. 9.

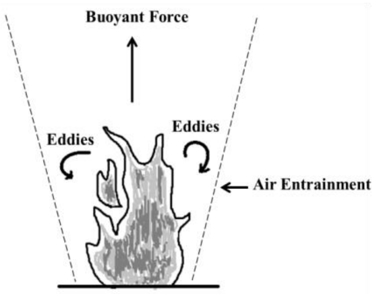


Fig. 9: Schematic Diagram of the Eddies Effect.

The top portions of the fire will fluctuate because of the Eddies effect, called the *Flickering* Phenomenon as shown in Fig. 10 and 11.

Fig. 10: Schematic Diagram of the Eddies Effect (From left to right).

Fig. 11: An example of the Eddies Effect in a real fire scene (From left to right).

After previous steps, those remaining object are non-moving. Now, we will determine the objects whether they are *Flickering* or not. Based on fire flickering, the upper edge on the rectangle may fluctuate as shown in Fig. 12. The line chart representing the height fluctuations of fire regions in Fig. 12 is shown in Fig. 13.

Fig. 12: Example of upper-edge fluctuations of the fire.

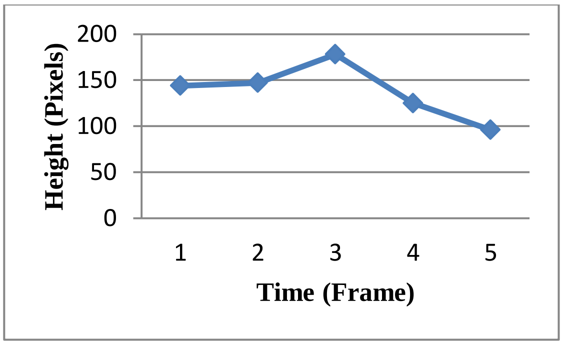


Fig. 13: Line chart of height fluctuations of the fire regions with respect to Fig. 12.

Because of the fire flickering, a *Foldback Point* (Fp) ocuurs in a short period of time during the fire burning process. For example, in Fig. 13, the 3rd frame is the corresponding Foldback Point. In our system, the number of the Foldback Points must satisfy the following criterion to be considered for further analysis:

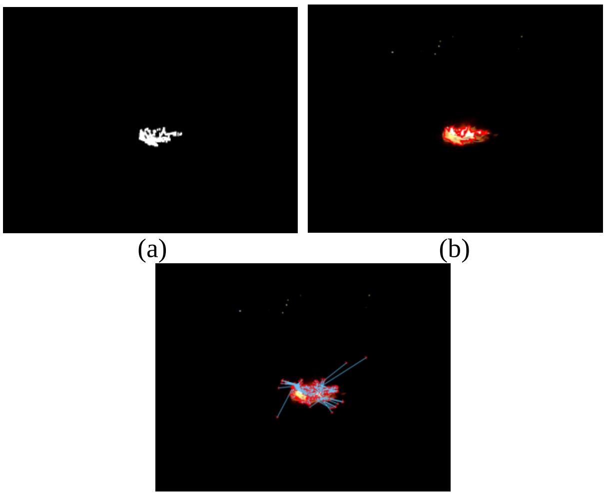
$$\mathrm{FpCount} > F_T \qquad (10)$$

where FpCount is the number of the Foldback Points, and F_T is a selected threshold.

2.4. Fire Flow Energy Analysis

In this section, our objective is to estimate the optical flows of the pixels on the fire contour in temporal video sequences. The Lucus-Kanade optical flows are based on three assumptions: (1) Brightness Constancy: The brightness of a pixel in a moving object will not change in adjacent frames; (2) Temporal Persistence: Small movements are observed from fram to frame. That is, it must not move quickly from frame to frame; (3) Spatial coherence: the pixels which belong to same object will have similar motion.

The *Pyramid Lucus-Kanade Optical Flows* method [8] is used to analyze the fire region, which is a technique to first solve for optical flows at the top layer, then at each layer below. The process is repeated through the pyramid until the lowest layer is reached. An example of the pyramid Lucas-Kanade optical flows method applied to a real fire scenario is shown in Fig. 14.



(a) (b) (c)

Fig. 14: An example of the *Pyramid Lucus-Kanade Optical Flows* method: (a) image containing the candidate fire region after image proprocessing; (b) region contour of (a); (c) estimated moition vectors.

According to the motion vectors as estimated by the pyramid Lucus-Kanade optical flows, we estimate the fire flow energy of the fire region as described herein. Let P denote the set of starting points for a feature point p, and Q denote the ending point for a feature point q, the fire flow energy of the fire region can be defined accordingly:

$$P = \{ (x_p, y_p) \}_{p \in 1,\ldots,n} \quad (12)$$

$$Q = \{ (x_q, y_q) \}_{q \in 1,\ldots,n} \quad (13)$$

Fire Flow Energy =

$$\frac{\Sigma\Sigma_{(x,y)\in \mathrm{Contour}}\left[(x_q\text{-}x_p)^2+(y_q\text{-}y_p)^2\right]}{\mathrm{Nc}} \quad (14)$$

where *Nc* denotes the number of the contour pixels in the fire region. Here, we require that the fire flow energy is greater than the selected threshold E_T.

3. RESULTS

A database containing 67 digital videos were collected for the development and evaluation of our system. Among which, 45 digital videos were acquired from Youtube, the remaining 22 videos were acquired in our laboratory. Table 1 summarizes the database used for system evaluation. During system software development, the research environment is summarized in Table 2.

Table 1: The database for system evaluation

	Youtube	*Our Videos*
Fire Scene	33	19
Non-Fire Scene	12	3

Table 2: System software environment

CPU	AMD Athlon II X640 3.0G
Main Memory	2GB
Operating System	Microsoft Windows 7
Integrated Development Environment	Microsoft Visual Studio C++ 2010 With OpenCV2.4.6

Table 3 summarizes the detection results using the database. Our results showed that the detection rate are 91.11% and 95.45% in Youtube videos and our videos, respectively.

Table 3: Fire Detection Results-Youtube

	Detection Results	
	Positive	*Negative*
Fire Scene	32	1
Non-Fire Scene	3	9

Table 4: Fire Detection Results-Our Videos

	Detection Results	
	Positive	*Negative*
Fire Scene	18	1
Non-Fire Scene	0	3

Table 5 summarizes the detection rate and average delay (Frames) of our system using the database. The experimental results showed that our system's average delays are 24.4 and 26.6 frames in Youtube videos and our videos, respectively. Therefore, an immediate fire alarm could be triggered in < 1 second (30 frames per second in a surveillance camera) once the fire occurred.

Table 5: Fire Detection Rate and Average Delay Frame

	Detection Rate	*Average Delay (Frames)*
Youtube	91.11%	24.4
Our Videos	95.45%	26.6

Fig. 15 and Fig. 16 show the fire detection results for indoor and outdoor fire scenarios, respectively. As shown, our system could be used to detect fire in various scenarios.

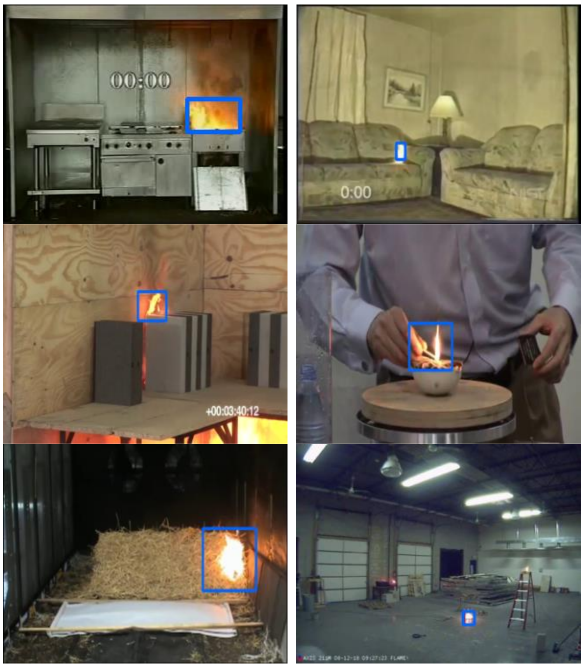


Fig. 15: Fire detection results for indoor scenearios.

Fig. 16: Fire detection results for outdoor scenearios.

Although our system was quite sensitive (> 91%), localization of the fire region may still subject to minor errors. Fig. 17 showed an example of the fire detection, where the location of the fire was not accurately localized.

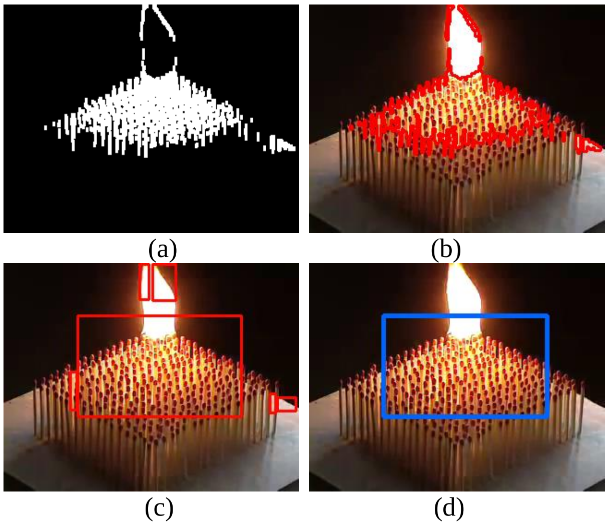


Fig. 17: An example of inaccurate fire localization: (a) candidate fire regions; (b) contour of the fire region; (c) candidate fire regions; (d) detected fire region.

Fig. 18 shows an example of the mis-detections in which our system failed to detect the fire in a real scenario. Fig. 18 (a) is the original image, (b) is the foreground regions as detected by our system, (c) is the resulting image after fire color filtering, and (d) is the candidate fire regions after the image preprocessing. In this example, the fire didn't exhibit color property as modeled in our system, resulting in mis-detection.

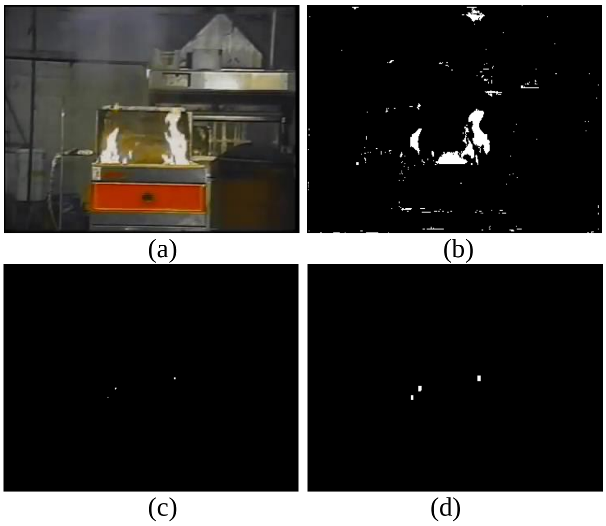
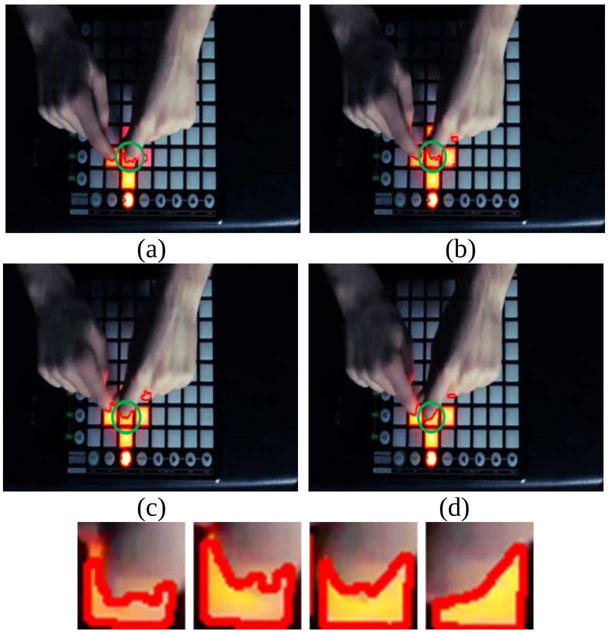


Fig. 18: An example of mis-detection in a real scenario: (a) original image; (b) foreground image; (c) resulting image after the fire color filtering; (d) resulting image after the morphological processing.

Fig. 19 shows an example of false alarms triggered by our system. The scenario contains a Novation LaunchPad, a media play platform on which a user can operate to trigger or process a music clip. Fig. 19 (a)~(d) are selected frames in the video sequence. As shown in Fig. 19 (e), the image characteristics of candidate regions in adjacent frames does mimic the property of fire, therefore detected by our system.



(a) (b) (c) (d) (e)

Fig. 19: False alarm of the fire detection: (a)~(d) is the selected images in the video suquence; (e) Partial magnification of regions in (a)~(d) that mimics the image characteristics of fire.

4. CONCLUSION

We proposed a vision-based fire detection system using video sequences which could potentially be used for early detection of fire. Based on the assumptions of fire properties and dynamics (e.g., color, region, Eddies effect, etc.), our system was designed using image/video processing techniques and included four major steps: *image preprocessing*, *foreground region analysis*, *fire dynamic behavior analysis*, and *fire flow energy analysis*.

Our experiment results showed that our system had been evaluated using a video database containing fire and/or non-fire scenarios. Overall, our system could achieve the detection rates of over 91% in either indoor or outdoor environments. In addition, our system could achieve the system response time within 1 second (average delay of ~25 frames) once the fire occurred. Therefore, the fire detection techniques being developed could be quite effective than traditional fire detectors (e.g., temperature, smoke detectors, etc.)

Despite the high accuracy yielded by our system, our system may still generate false alarms when a moving object does mimic the image properties of fire. However, we anticipate that these situations are rather rare in real scenarios.

In this paper, we have developed fire-detection methods (algorithms) and have presented a fire detection system that could be used to automatically detect fire at its early stage. In addition, we have carefully evaluated the fire dynamic behavior with respect to its image characteristics in real scenarios. In summary, our system could be used in conjunction with other fire detectors (e.g., smoke detectors and the like), leading to prevent damage to peoples' security, life, and property.

REFERENCES

[1] G. Healey, D. Slater, T. Lin, B. Drda, and D. Goedeke. "A system for real-time fire detection," *IEEE International Conference on Computer Vision and Pattern Recognition*, pp. 605 – 606, January 1993

[2] C. B. Liu and N. Ahuja "Vision Based Fire Detection, *IEEE International Conference on Pattern Recognition*,Vol. 4, pp. 134 – 137, August 2004.

[3] H.D. Duong, D.D. Nguyen, L.T. Ngo, and D.T. Tinh, "On Approach to Vision Based Fire Detection Based on Type-2 Fuzzy Clustering," *IEEE International Conference on Soft Computing and Pattern Recognition*, pp. 51-56, October 2011.

[4] Y. Wei, Y. Chunyu, Z. Yongming, "Based on wavelet transformation fire smoke detection method," ICEMI International Conference on Electronic Measurement and Instruments, pp. 2-872 – 2-875, August 2009.

[5] T. H. Chen, P.H. Wu, and Y. C. Chiou, "An early fire-detection method based on image processing," *Image Processing 2004*, Vol. 3, pp. 1707 - 1710,October 2004.

[6] S. Rinsurongkawong, M. Ekpanyapong, and M. Dailey, "Fire Detection For Early Fire Alarm Based On Optical Flow Video Processing," *IEEE International Conference on* Electronic Engineering/Electronics, Computer, Telecommunications and Information Technology(ECTICON), pp. 1 – 4, May 2012

[7] R.M. Haralick and L. G. Shapiro, *Computer and Robot Vision Volume I*, Addison-Wesley Publishing Company Inc., United States of America, 1992.

[8] J. Y. Bouguet, "Pyramidal implementation of the affine Lucas Kanade feature tracker description of the algorithm," *Intel Corporation* 5, 2001.