16th European Signal Processing Conference (EUSIPCO 2008), Lausanne, Switzerland, August 25-29, 2008, copyright by EURASIP

EFFICIENT VISUAL FIRE DETECTION APPLIED FOR VIDEO RETRIEVAL

Paulo Vinicius Koerich Borges[*∇], *Joceli Mayer*[*], *Ebroul Izquierdo*[∇]

LPDS - Dept. of Electrical Engineering, Federal University of Santa Catarina, Brazil.[*]
MMV - Dept. of Electronic Engineering, Queen Mary, University of London, UK.[∇]

ABSTRACT

In this paper we propose a new image event detection method for identifying fire in videos. Traditional image based fire detection is often applied in surveillance camera scenarios with well behaved background. In contrast, the proposed method is applied for retrieval of fire catastrophes in newscast content, such that there is great variation in fire and background characteristics, depending on the video instance. The method analyses the frame-to-frame change in given features of potential fire regions. These features are colour, area size, texture, boundary roughness and skewness of the estimated fire regions. Because of flickering and random characteristics of fire, these are powerful discriminants. The change of each of these features is evaluated, and the results are combined according to the Bayes classifier to to achieve a decision (i.e. fire happens, fire does not happen). Experiments illustrated the applicability of the method and the improved performance in comparison to other techniques.

1. INTRODUCTION

Recently, automated retrieval of events in newscast videos has received great attention by the research community [1, 2], basically motivated by the interest of broadcasters in building large digital archives of their assets for reuse of archived materials. A huge amount of time and money is spend by news networks to find in their archives events related to a fresh event that has just happened. In this context, catastrophe related news are one of the most common topics that require automated retrieval, and this task is subject to a number of large research projects [3, 4]. In the catastrophe news, fire events are one of the most common topics, along with bombings and floods. Efficient detection of fire in video contents has proved to be an important research topic in the last few years [3, 5, 6], with application not only in video retrieval but also in surveillance in security systems.

In this paper we propose an efficient visual based event detection method for identifying fire in videos. Most of the visual based fire detection proposed in the literature are applied in surveillance camera scenarios with well behaved background. Otherwise, they propose the use of filter banks [7], frequency transforms [8] and motion tracking, requiring more computational processing time, making them unsuitable for video retrieval.

In contrast, the proposed method is applied for retrieval of fire catastrophes in news content, such that there is great variation in fire and background characteristics, depending on the video instance. The method analyses the frame-to-frame change in given features of potential fire regions. These features are colour, area size, texture, boundary roughness and skewness of the estimated fire regions. Because of flickering and random characteristics of fire due the combustion and air flow, these are efficient classifying features. The change of each of these features is evaluated, and the results are combined according to the Bayes classifier to determine whether or not fire occurs in that frame. In contrast to some works [9], the goal of this paper is not to identify fire pixels in a given image or video frame, but to determine if fire occurs in the frame.

The authors would like to thank CNPq (grants 202288/2006-4, 550745/2007-6, 472536/2007-9) and the European Commission (IST MESH, 6th Framework Programme) for the financial support.

The majority of the vision based fire detection systems employ some type of hybrid model combining colour, geometry and motion information. In general, fire detection systems use colour clues as a precondition to generate seed areas for possible fire regions (called a fire mask - FM - in this paper) since it is the most discerning feature. An effective colour model for potential fire pixel determination is thus essential for almost any vision based fire detection systems. Assuming we have an efficient fire mask, the contributions in this paper can be listed as follows:

(*i*) Unlike many works in the literature which use shape descriptors to analyse the amount of flame motion, we use the boundary roughness of the potential fire regions. This brings the same amount of discernibility with a more efficient processing speed.

(*ii*) It has been frequently observed in the literature [10] that saturation occurs in the red channel of fire regions. In order to exploit this characteristic, we propose the use of the third order statistical moment (skewness) of the potential fire region as a feature. This is a simple and powerful discriminant, which significantly enhances the detection performance.

(*iii*) We also propose the use of the variance as a feature, due to the randomness observed in fire surfaces.

(*iv*) For real fire regions, the amount of fire varies from frame to frame due to flame flickering. This is also used as a feature with classification power.

(*v*) The features are combined with the Bayes classifier to achieve an practical low detection error rate.

This paper is organised as follows. In Section 2 we review existing techniques, contrasting them with the proposed method. In Section 3 we discuss colour and dynamic fire characteristics, proposing efficient discrimination features for fire. In Section 4 we propose a framework for classification, based on the Bayes classifier. In Section 5 we present a experimental results, followed by relevant conclusions in Section 6.

2. RELATED METHODS

There is not a large number of papers about fire detection in the computer vision literature. The first works used purely a colour based model [11], which is the initial step for many other algorithms, including the one proposed in this paper.

In [12] the authors use pixel colours and their temporal variations. They use an approach that is based upon creating a Gaussian-smoothed colour histogram to determine the fire-coloured pixels, and then using the temporal variation of pixels to determine which of these pixels are actually fire. However, this algorithm is also essentially colour based, and does not exploit other statistical characteristics of potential fire regions. In addition, temporal variation in image pixel colour does not capture the temporal property of fire which is more complex and benefits from a region level representation. As observed in [8], for example, pixels in the core of the fire exhibit less temporal variation than the other pixels.

In [9] the authors propose the use of a fuzzy logic approach which uses luminance and chrominance information to replace the existing heuristic rules used to generate the FM. The implicit uncertainties in the rules obtained from repeated experiments can be encoded in a fuzzy representation that is expressed in linguistic terms. The authors argue that the single output decision quantity will then give a better likelihood that a pixel is a fire pixel. Although a very good detection rate is achieved, the method is focused on images and random changes of fire from frame to frame are not exploited.

The work presented in [7] also yields good results, where boundary of flames are represented in wavelet domain and high frequency nature of the boundaries of fire regions is also used as a clue to model the flame flicker spatially. However, this work presents two drawbacks: first, the algorithm assumes that the camera is stationary; and second, it presents a high computational complexity. Considering that newscast videos have at least 25 frames per second (fps), the analysis would be very time consuming for video retrieval applications. A similar situation occurs in the work presented in [8], which makes use of the Fourier transform for boundary description of every potential fire region.

3. STATISTICAL CHARACTERISTICS OF FIRE

It is well known that fire has unique visual signatures. Colour, geometry, and motion of fire region are all essential features for efficient classification. In general, in addition to colour, a region that corresponds to fire can be captured in terms of the statistical characteristics of the pixels in the region, and the spatial structure defined by their boundaries variation within the region. The shape of a fire region often keeps changing and exhibits a stochastic motion, which depends on surrounding environmental factors such as the type of burning elements and wind.

Based on these factors, in the following we propose several useful features for detecting fire, explaining the physical characteristics to validate their applicability.

3.1 Colour

According to most fire detection papers presented in the literature and based on our own experiments, we notice that fire has very distinct colour characteristics, and although empirical, it is the more powerful single feature for finding fire in video sequences. In tests with 120 images in different resolutions and scenarios, it is reasonable to assume that generally the colour of flames belongs to the red-yellow range.

It is noticed that for a given fire pixel, the value of red channel is greater than the green channel, and the value of the green channel is greater than the value of blue channel. Several additional characteristics also hold, such that for a given fire pixel $f(m,n)$ in an image f we have

$$f_R(m,n) > \bar{f}_R \tag{1}$$

where f_R is the red channel representation of f and $\bar{f}_R$ is the average pixel value of f_R. In addition,

$$f_R(m,n) > f_G(m,n) > f_B(m,n) \tag{2}$$

In general, using an extremely permissive threshold, the following holds true:

$$f_R(m,n) > 190, \tag{3}$$
$$f_G(m,n) > 100, \tag{4}$$
$$f_B(m,n) < 140. \tag{5}$$

There are a few cases, however, that the above thresholds may fail, specially in very bright or low quality recording conditions. For this reason, we adopt a normalised relationship, as suggested in [13]:

$$0.25 \leq f_G(m,n)/(f_R(m,n)+1) \leq 0.65 \tag{6}$$
$$0.05 \leq f_B(m,n)/(f_R(m,n)+1) \leq 0.45 \tag{7}$$
$$0.20 \leq f_B(m,n)/(f_G(m,n)+1) \leq 0.60 \tag{8}$$
$$\tag{9}$$

bringing more robustness to illumination variations.

Based on the rules above a binary image called fire mask (FM) for each frame is generated, where the values 1 or 0 indicate the presence of absence of fire at the corresponding location in the image f. We refer to the concatenation of pixels '1' as fire blobs in this paper. The FM is then processed with a connected components algorithm so that the potential fire blobs are concatenated in a contiguous region.

(a) Fire Area = 1692 pixels. (b) Fire area = 2473 pixels.

Figure 1: Illustration of the change in fire pixel area from frame to frame.

Notice that the rules above are very permissive and many non-fire regions may be included in the FM. For this reason, additional analysis is necessary to further refine the results. To define a real burning fire, in addition to using chromatics, statistical and dynamic features are usually adopted to distinguish other fire aliases [10, 8]. Examples of these fire dynamics include the change in shape, flame movement and flickering. The statistical and dynamic fire features proposed in this paper are discussed next.

3.2 Randomness of Area Size

For the estimated fire pixel area, because of the fire flickering, a change in the area of the FM necessarily occurs from frame to frame, as illustrated in Fig. 1. Non-fire areas have a less random change in the area size. The normalised area change ΔA_i for the i-th frame is given by

$$\Delta A_i = \frac{|A_i - A_{i-1}|}{A_i} > \lambda_A \tag{10}$$

where A_i corresponds to the area of the fire blobs representing the potential fire regions in the FM and λ_A is a decision threshold.

3.3 Boundary Roughness

As discussed in Section 2, in [8], for example, the authors represent the shape of fire regions using Fourier Descriptors (FD) [14], based on the coefficients of the Fourier Transform. However, in a retrieval application, this solution presents two drawbacks: (i) retrieval demands a very high processing speed, and the evaluation of the FD for every frame is a very time consuming operation; (ii) although the FD are excellent shape descriptors, for fire detection purposes what we are really interested is the randomness or roughness of the shape, and not the shape itself, as fire does not have a specific boundary characteristics. Therefore, we use the boundary roughness of the potential fire region as a feature, given by the ratio between perimeter and convex hull perimeter [15]. The convex hull of a set of pixels S is the smallest convex set containing S, as illustrated in Fig. 2. Let C_i be the boundary roughness for the blobs in the FM corresponding to the i-th frame in the video. The presence of fire is assumed if

$$\Delta C_i = \frac{C_i - C_{i-1}}{C_i} > \lambda_C \tag{11}$$

where λ_C represents a threshold in the rate of change. The inclusion of the term C_i in the denominator normalises the metric to be independent of the size of the blob.

Experiments illustrate that this is an excellent and computationally efficient discriminant for the shape of fire regions, yielding similar results to the use of FD.

Figure 2: Illustration of the convex hull (red line) used to evaluate the boundary roughness of the blob.

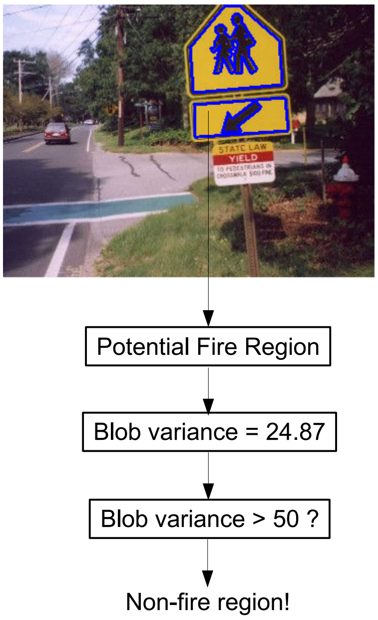

Figure 3: Eliminating potential fire regions through variance analysis.

3.4 Texture

Unlike other false-alarm regions, like an yellow traffic sign, for example, fire regions have a significant amount of texture characteristics because of its random nature. Therefore, we use the variance of the blobs as a feature to help eliminating non-fire blobs in the FM. Experiments indicate that fire blobs usually have a variance $\sigma > \lambda_\sigma$, where $\lambda_\sigma = 50$ is determined from a set of experimental analysis. Fig. 3 illustrates how the use of the variance can reduce the false-alarm rate of the FM.

3.5 Skewness

The skewness measures the degree of asymmetry of a distribution around its mean [16]. It is zero when the distribution is symmetric, positive if the distribution shape is more spread to the right and negative if it is more spread to the left, as illustrated in Fig. 4.

As discussed in Section 3.1, fire regions have high pixel values for the green and specially for the red channel. Very often, we observe a saturation in these channels, leading the histograms to the upper side of the range, as illustrated in Fig. 5. This causes the skewness of these distributions to have a high negative value. For this reason, we employ the skewness as an useful feature to identify

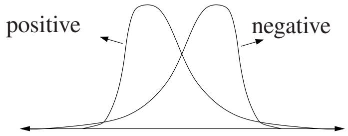

Figure 4: Illustration of the effect of positive and negative skewness.

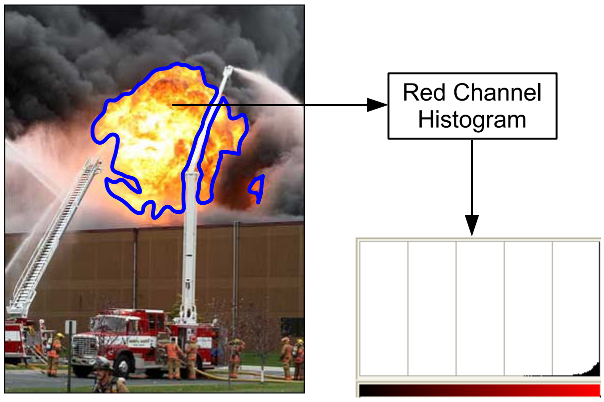

Figure 5: Illustration of the saturation effect in the histogram of the red channel of the fire regions inside the blue boundaries.

fire regions. Let the skewness γ be defined as

$$\gamma = \frac{\frac{1}{J^2}\sum_{m=1}^{J}\sum_{n=1}^{J}[f(m,n)-\bar{f}]^3}{\sigma_f^3} \qquad (12)$$

A potential fire region present at frame i is assumed as real fire if

$$\gamma_i < \lambda_\gamma \qquad (13)$$

where λ_γ is a decision threshold. Experiments illustrate that fire regions usually have $\gamma_i < -1$, and this is the value we use in the experiments in Section 5.

4. CLASSIFICATION

Considering a stochastic interpretation of the features in this work the Bayes classifier [14] is employed to combine the features, although it is clear that different classifiers could also be tested.

For each frame i, a naive FM is initially created based on the set of rules for colour, from Section 3.1. From this FM, a vector $\mathbf{d}_i$ of features is obtained and the features are combined according to the Bayes classifier. The features used are the ones discussed in the previous section (area change, roughness, variance and red channel skewness). A description on how to combine the features using this technique can be found in [15]. A block diagram of the process is given in Fig. 6. Therefore, the naive thresholds λ_A, λ_C, λ_γ and λ_σ are not applied, but only give the reader a practical reference for the typical values for these parameters.

Although some features have better classification power than others, because all the discussed features are useful to discriminate fire from non-fire, combining them increases the distance between these two classes, and consequently reduces the detection error rate [14], at the expense of increasing computational complexity.

5. EXPERIMENTS

In the experiments, we used as a test set a selection of videos from the MESH [3] database of news content. This database is formed of several instances of catastrophe related videos from the

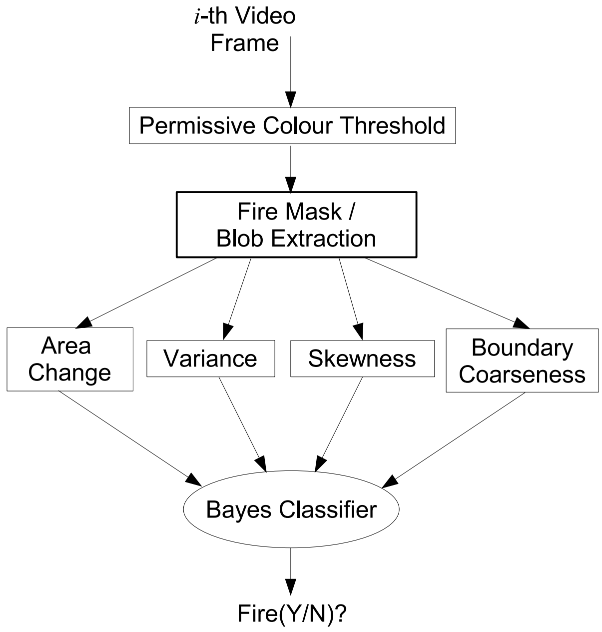

Figure 6: Block diagram illustrating the fire detection process for each frame *i*, including the FM generation, the extraction of features and the classification according to the Bayes classifier.

Figure 7: Illustration of an artificial fire-like region, with potential false-negative.

Deutsche Welle network, containing several news reports related to fire events. They include different kinds of fires such as building, wildland and residential fire, containing shots captured at day time, dusk or night time. This diversity is convenient to evaluate the performance of the system under different lighting and quality conditions. The video selection also contains many shots of objects with fire-like appearances such as sunsets and artificial background from the news program, as illustrated in Fig. 7. Notice that, for this figure, a naive classification based only on the set of rules described in Section 3.1 would cause a wrong detection.

The video resolution is 768×576 and the frame rate is 25 frames per second (fps). There are approximately 532 minutes of video (or 798,000 frames).

The frames are classified as "contains fire" or "does not contain fire," as discussed in Section 4. For training the classifier, 75 instances of fire in video sequences were used, generating a 4-D decision surface. Fig. 8 illustrates a 3-D separating surface for discriminating fire from non-fire regions, based on the features roughness, variance and normalised area change.

The results using the proposed features are presented in Table 1. This tables shows the false positive (wrongly assume the presence of fire) rate and the false negative (wrongly assume the absence of fire) rate. The rows corresponding to each feature represents the

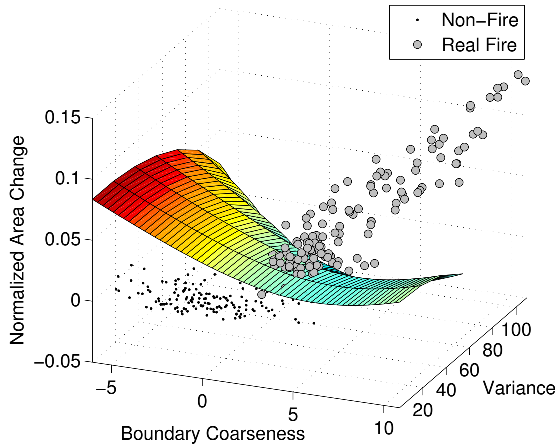

Figure 8: Illustration of a surface separating frames from fire to non-fire classes, based on the features roughness, variance and normalised area change.

Table 1: Experimental error rates.

Features Used	False-Positive	False-Negative
Colour	11.2%	0.04%
+ Skewness	2.2%	-
+ Roughness	4.5%	-
+ Area Change	3.2%	-
+ Variance	7.9%	-
Combination	0.9%	-

error rates when that feature only is used to classify fire from non-fire, after the basic colour based fire segmentation is done. The row 'Combination' corresponds to the results when all the features are combined, as illustrated in Fig. 6. The results described in [8] and [7] yield error rates similar to the system proposed here. However, [7] assumes the camera is stationary and [8] makes use of frequency transforms and motion tracking, requiring more computational processing time, making them unsuitable for video retrieval. In [9] an excellent detection rate is achieved (99%), however a false-positive rate of 9.9% is obtained, which is considerably higher the rate in the proposed method, when all the features are combined.

6. CONCLUSIONS

In this paper we have exploited important visual features of fire not previously discussed in the literature, like boundary roughness and skewness of the fire pixel distribution. The skewness, in particular, is a very useful descriptor because of the frequent occurrence of saturation in the red channel of fire regions. Also, we have proposed modifications to motion based features, yielding an efficient performance. In contrast to other methods which extract complicated features, the features discussed here allow very fast processing, making the system applicable not only for real time fire detection, but also for video retrieval in news contents, which require faster than real-time analysis. The experiments illustrate the applicability of the method, with an average false-negative rate of 0.9% and a false-positive rate of 0.4%.

REFERENCES

[1] Y. Wang, Z. Liu, and J. C. Huang, "Threading and autodocumenting news videos: a promising solution to rapidly browse news topics," *IEEE Signal Processing Magazine*, vol. 3, pp. 59–68, March 2006.

[2] Y. Wang, Z. Liu, and J. C. Huang, "Multimedia content analysis-using both audio and visual clues," *IEEE Signal Processing Magazine*, vol. 17, pp. 12–36, November 2000.

[3] MESH - *Multimedia Semantic Syndication for Enhanced News Service* - European Commission Project - IST 6th Framework Programme.

[4] K-SPACE - *Knowledge Space of Semantic Inference for Automatic Annotation and Retrieval of Multimedia Content* - European Commission Project - IST 6th Framework Programme.

[5] P.-H. Huang, J.-Y. Su, Z.-M. Lu, and J.-S. Pan, "A fire-alarming method based on video processing," in *IEEE International Conference on Intelligent Information Hiding and Multimedia Signal Processing*, December 2006.

[6] T. Celik, H. Demirel, H. Ozkaramanli, and M. Uyguroglu, "Fire detection in video sequences using statistical color model," in *IEEE International Conference on Acoustics, Speech and Signal Processing*, vol. 2, (Toulouse, France), May 2006.

[7] B. U. Toreyin and A. E. Cetin, "Online detection of fire in video," in *IEEE Conference on Computer Vision and Pattern Recognition*, June 2007.

[8] C.-B. Liu and N. Ahuja, "Vision based fire detection," in *International Conference on Pattern Recognition*, vol. 4, pp. 134–137, August 2004.

[9] T. Celik, H. Ozkaramanli, and H. Demirel, "Fire pixel classification using fuzzy logic and statistical color model," in *International Conference on Acoustics, Speech and Signal Processing*, vol. 1, pp. 1205–1208, April 2007.

[10] T.-H. Chen, P.-H. Wu, and Y.-C. Chio, "An early fire-detection method based on image processing," in *IEEE International Conference on Image Processing*, vol. 3, pp. 1707–1710, October 2004.

[11] G. Healey, D. Slater, T. Lin, B. Drda, and A. D. Goedeke, "A system for real-time fire detection," in *IEEE Conference on Computer Vision and Pattern Recognition*, June 1993.

[12] I. W. Phillips, M. Shah, and N. da Vitoria Lobo, "Flame recognition in video," in *IEEE Workshop on Applications of Computer Vision*, pp. 224–229, December 2000.

[13] T. Celik, H. Demirela, H. Ozkaramanlia, and M. Uyguroglua, "Fire detection using statistical color model in video sequences," *Journal of Visual Communication and Image Representation*, vol. 18, pp. 176–185, April 2007.

[14] S. Theodoridis and K. Koutroumbas, *Pattern Recognition*. Academic Press, 2006.

[15] R.C.Gonzalez and R.E.Woods, *Digital Image Processing*. Addison-Wesley, 1992.

[16] D. G. Manolakis, V. K. Ingle, and S. M. Kogon, *Statiscal and Adaptive Signal Processing*. McGraw-Hill, 2000.

[17] R. O. Duda, P. Hart, and D. G. Stork, *Pattern Classification*. John Wiley, 2nd ed., 2001.