Vision Based Fire Detection

Che-Bin Liu and Narendra Ahuja
Beckman Institute
University of Illinois at Urbana-Champaign
Urbana, IL 61801
{cbliu, ahuja}@vision.ai.uiuc.edu

Abstract

Vision based fire detection is potentially a useful technique. With the increase in the number of surveillance cameras being installed, a vision based fire detection capability can be incorporated in existing surveillance systems at relatively low additional cost. Vision based fire detection offers advantages over the traditional methods. It will thus complement the existing devices. In this paper, we present spectral, spatial and temporal models of fire regions in visual image sequences. The spectral model is represented in terms of the color probability density of fire pixels. The spatial model captures the spatial structure within a fire region. The shape of a fire region is represented in terms of the spatial frequency content of the region contour using its Fourier coefficients. The temporal changes in these coefficients are used as the temporal signatures of the fire region. Specifically, an autoregressive model of the Fourier coefficient series is used. Experiments with a large number of scenes show that our method is capable of detecting fire reliably.

1. Introduction

The objective of this work is in the general context of modeling and recognizing shape evolution in stochastic visual phenomena. In particular, this paper focuses on detection of fire in image sequences. Fire has diverse, multispectral signatures, several of which have been utilized to devise different methods for its detection. Most of the methods can be categorized into smoke, heat, or radiation detection. A detailed survey can be found in [2]. Each fire detection method is better suited to a distinct environment. Vision based fire detection has the following advantages over the other methods. First, it has fast response to fires. Like the radiation based method, it detects fires as soon as they appear in sight. Second, it directly senses the location of fire (in 2-D), not just radiation which comes from its general vicinity. Last, but not least, it is capable of analyzing existing images or image sequences so that it can be used for multimedia database retrieval. Line of sight visual methods like this complement other methods that use associated cues of smoke and heat.

1.1. Related Work

There are only a few papers about fire detection in computer vision literature. Healey et al. [3] use a purely color based model. Phillips et al. [10] use pixel colors and their temporal variations. These methods have the following two drawbacks. First, a region composed of fire-colored pixels is too simple a model of fire since fire also has spatial structure, namely the core is brighter than the periphery. Second, temporal variation in image pixel color does not capture the temporal property of fire which is more complex and benefits from a region level representation. For example, pixels of the core of the fire exhibit less temporal variation than the other pixels.

The fire detection method described in this paper includes recognition of evolving region shapes. There has been an enormous amount of literature related to static shape analysis. A survey can be found in [4]. Our method is more relevant to works on modeling and recognition of deformable shapes/objects [1, 7]. These methods implicitly assume all shapes have to be observed before learning the subspace or the manifold. Thus, they are very likely to fail to recognize objects with stochastic appearances, such as fire. The shapes of fires with different burning materials could be of a large degree of variability. These methods do not have good representation in shapes and their evolution.

2. Fire Models

Fire has unique visual signatures. Color, geometry, and motion of fire region are all essential for recognition. A region that corresponds to fire can be captured in terms of (1) spectral characteristics of the pixels in the region, and (2) the spatial structure defined by their spectral variation within the region. The shape of a fire region usually keeps changing and exhibits a stochastic motion, which depends

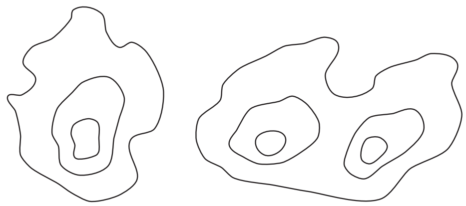

Figure 1: Examples of the nested ring structure of fire regions. (a) A fire region with a single core. (b) A fire region with two cores.

on surrounding environmental factors such as the type of burning materials and air flow.

The pixels in a fire region have characteristic color spectra and the pixels with different spectra have characteristic relative locations. In color images, we might see bright white color in the core, and yellow, orange and red away from the core. In grayscale images, we see that core is brighter than the periphery. Note that a fire region may include multiple bright cores which correspond to multiple hot spots. This can be viewed as a large fire composed of multiple sources of fires as illustrated in Figure 1. Thus, the fire region in a single image can be modeled as follows: (i) It stands in high contrast to its surroundings; (ii) It exhibits a structure of nested rings of colors, changing from white at the core to yellow, orange and red in the periphery.

A fire in motion has a relatively static general shape (determined by the shape of burning materials) and rapidly changing local shape in the unobstructed part of the border. The lower frequency components of fire region boundary are relatively steady over time, and the higher frequency components change in a stochastic fashion. Accordingly, we use a stochastic model to capture the characteristic random motion of fire boundaries over time.

3. Fire Detection Algorithms

Our algorithms for video based fire detection make use of spectral, spatial, and temporal properties of fire regions. First, we extract potential fire regions from an image using fire spectral and spatial models. Second, we represent boundaries of these regions using Fourier coefficients. Third, we estimate parameters of an AR model of each region with its correspondence in previous images in the image sequence. Last, Fourier coefficients and AR model parameters are used as features of each region for a classifier that recognizes fire regions.

3.1. Potential Fire Region Detection

We first detect potential fire regions based only on the fire spectral and spatial models described in Section 2. We first extract high intensity regions (in grayscale) possibly corresponding to fire cores, which we called seed regions. We grow each seed region by following spectral gradients of the image and adding neighbor pixels if they have colors given by the fire color model with sufficiently high likelihood. The interior color probability density functions of fire are modeled as a mixture of Gaussian distributions in HSV space [13].

For each extracted region, we traverse its boundary to check if half of boundary points is of interior fire color. This check eliminate regions with most pixels having extremely high intensity such as a purely bright white regions.

3.2. Shape Representation

We represent the shape of fire regions in Fourier domain. Fourier Descriptors (FD), the Fourier Transform coefficients of the shape boundary, represents a 2-D shape using an 1-D function. There are several variations of Fourier based 1-D boundary representation in literature [6]. In this paper, we use Persoon and Fu's method [9].

Given an extracted region, we first retrieve its boundary using eight-connected chain code. Assume that we have N points from the chain code representation of the boundary. We express these points in complex form: $\{z_i | z_i = x_i + jy_i\}_{i=1}^{N}$ where (x_i, y_i) are the image coordinates of boundary points as the boundary is traversed clockwise. The coefficients of the Discrete Fourier Transform (DFT) of $\{z_i\}_{i=1}^{N}$ are

$$a_k = \frac{1}{N} \sum_{i=1}^{N} z_i \exp(-j \frac{2\pi}{N} ik) \tag{1}$$

where $k = -\lfloor \frac{N-1}{2} \rfloor, \ldots, \lfloor \frac{N}{2} \rfloor$. If M harmonics are used ($M \leq \lfloor \frac{N-1}{2} \rfloor$), the coefficients $\{a_m\}_{m=-M}^{M}$ are the Fourier Descriptors used to characterize the shape. Note that $a_0 = \frac{1}{N} \sum_{i=1}^{N} z_i$ represents the center of gravity of the 1-D boundary, which does not carry shape information. We neglect it to achieve translation invariance.

Related works in Fourier based shape description usually discuss about similarity measures that make FD invariant to relevant transformations, e.g., rotation, translation and scale. The requirement for each invariance depends on the applications. In this paper, we do not consider rotation invariance because fire shapes are not rotation invariant. Since rotation invariance is not relevant, we can always choose the starting point as the topmost boundary pixel along the vertical axis through the center of gravity of the entire shape. Our representation approximates scale invariance (by dropping a_0 term) since we retrieve boundary points in the chain code fashion. Chain code expression discretizes the arc and Equation (1) normalizes the arc length [1].

For detection purpose, we represent only the fire boundary as Fourier coefficients. Generally, to model a fire region, each ring in the fire region (Fig. 1) can be represented as a set of Fourier coefficients.

3.3. Stochastic Temporal Variation of Shape

The stochastic characteristics of fire boundary motion are estimated by an autoregressive model of changes in Fourier coefficients of the region boundary. The autoregressive (AR) model is used based on the assumption that each term in the time series depends linearly on several previous terms along with a noise term [5]. Since the lower frequency coefficients are likely to remain static and higher frequency coefficients have higher temporal variation, this model will capture different levels of temporal variation of FDs.

Suppose v_k are the m-dimensional random vectors observed at equal time intervals. The m-variate AR model of order p (denoted as AR(p) model) is defined as

$$v_k = \sum_{i=1}^{p} A_i v_{k-i} + n_k. \quad (2)$$

The matrices $A_i \in R^{m\times m}$ are the coefficient matrices of the AR(p) model, and the m-dimensional vectors n_k are uncorrelated random vectors with zero mean.

Since we use Fourier Descriptors to represent fire region shape, we represent the stochastic characteristics of the temporal changes in the magnitude of each FD using the AR model. If M harmonics of FDs are used, then a $2M$-dimensional random vectors v_k represents the region shape at time k. We further assume that different FDs at any given time k are independent of each other, so we have diagonal coefficient matrices A_i, where $A_i(m,n) = 0$ if $m \neq n$. Thus the problem can be seen as modeling of $2M$ independent time series.

To select the optimum order of the AR model, we adopt Schwarz's Bayesian Criterion [11] which chooses the order of the model so as to minimize the forecast mean-squared error. We have found that the AR(1) model yields the minimal error. Thus, we have $v_k = Av_{k-1} + n_k$. We estimate the parameters of our AR(1) model using Neumaier and Schneider's algorithms [8].

[1]Note: scale invariance is achieved if the distances between a pixel and its eight neighbors are considered as equal.

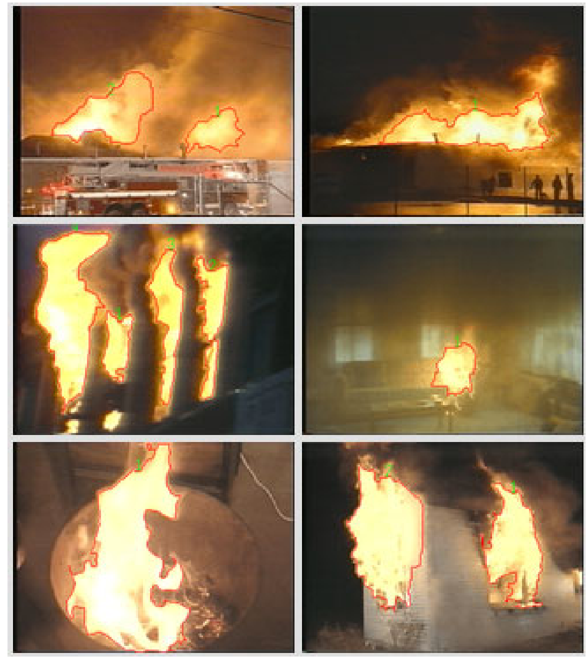

Figure 2: Selected fire images used in experiments.

4. Experimental Results

The video clips used in our experiments are real-world image sequences taken from a random selection of commercial/training video tapes. They include different types of fires such as residential fire, warehouse fire, and wildland fire. We use images captured at day time, dusk or night time to evaluate system performance under different lighting conditions. We also use other image sequences containing objects with fire-like appearances such as sun and light bulbs as negative examples. Most image sequences involve camera motion. The video clips that we tested our algorithm on contain a total of 3956 image frames in 36 sequences. Figure 2 shows some selected fire images used in our experiments. The (red) contours depicted in the images are the detected fire region contours. As seen in some images, fire sometimes complements with smoke nearby. Our spectral and spatial models of fire regions define boundaries between fire and smoke.

Our potential region extraction algorithm extracts almost all the true fire regions. It also extracts other fire-like objects. What it does not extract are mainly spark-like, small fire regions emanating from the main fire regions which are detected. In our test data, the algorithm extracted a total of 1319 fire-like region contours, 1089 of which were true fire region contours. These contours are used for test of our fire region classification.

For shape representation in terms of Fourier Descrip-

Table 1: Average recognition rates of fire and non-fire contour recognition.

Experiments	Fire	Non-Fire
A: Use FD	0.996	0.904
B: Use FD and AR	0.999	1.0

tors, we find that using 40 coefficients (i.e. $M = 20$) is sufficient to approximate the relevant properties of the fire region contours. Accordingly, we use 40 AR coefficients (diagonal of matrix A) to represent the stochastic characteristics of the temporal changes in FDs. We use a two-class Support Vector Machine (SVM) classifier with radial basis function (RBF) kernel [12] for fire region recognition.

We tested our algorithms in two ways: The first set of experiments was performed with only spatial properties of region contours (only FD as feature vectors), and the second set of experiments was performed with both spatial and temporal properties of region contours and their changes (FD and AR parameters as feature vectors). In the second set of experiments, we required that the fire contour be seen in at least previous four frames. Note that three frames are the minimum requirement to estimate parameters of our AR(1) model. For each set of experiments, we repeated the test ten times using one-tenth of fire and one-tenth of non-fire region contours to train the SVM classifier, and using the other region contours for test only. In this way, we used many more fire examples than counter examples on training. This was intended to tilt the detector in favor of false positives vs. false negatives as can be seen from the average recognition rates shown in Table 1.

The results of experiments in Table 1 row A show that our method is capable of recognizing fire contours using single images, for which we credit our spatial model that searches for fire-like regions instead of fire-like pixels [3, 10]. False positive examples are images of sun and mask lamps. However, it is clear from row B that temporal information of shape dynamics indeed improve the detection performance significantly. In particular, we also tested our algorithms with image data sets provided on the web site of 2003 Workshop on Visual Surveillance and Performance Evaluation of Tracking and Surveillance (VS-PETS 2003). These images had no positive fire examples and our algorithms gave no false alarms in any of the image sequences.

4.1. Limitation

In Section 3.2, we approximate scale invariance for FD by dense sampling of region boundary. However, spatial quantization errors for small regions are likely to introduce considerable noise in the FD. To avoid this problem, we place a threshold to eliminate regions of small size (number of pixels). We also exclude large but thin regions. Consequently, our algorithm does not detect very small or far away fire.

5. Conclusion

In this paper, we have proposed a vision based fire detection algorithm based on spectral, spatial and temporal properties of fires. Experiments show that our algorithm detects fire with high accuracy, both in single images as well as in image sequences. Our approach extends beyond fire detection. The stochastic model that we use to represent the dynamics of fire region can be applied to many other stochastic visual phenomena, which is the underlying general motivation for this paper.

References

[1] A. Baumberg and D. Hogg. An adaptive eigenshape model. In *Proceedings of the British Machine Vision Conference*, volume 1, pages 87–96, 1995.

[2] W. D. Davis and K. A. Notarianni. NASA fire detection study. Technical report, National Institute of Standards and Technology, March 1996.

[3] G. Healey, D. Slater, T. Lin, B. Drda, and D. Goedeke. A system for real-time fire detection. In *Computer Vision and Pattern Recognition*, pages 605–606, 1993.

[4] S. Loncaric. A survey of shape analysis techniques. *Pattern Recognition*, 31(8):983–1001, 1998.

[5] H. Lütkepohl. *Introduction to Multiple Time Series Analysis*. Springer-Verlag, 1991.

[6] S. Mori, H. Nishida, and H. Yamada. *Optical Character Recognition*. John Wiley & Sons, 1999.

[7] H. Murase and S. K. Nayar. Visual learning and recognition of 3D objects from appearance. *International Journal of Computer Vision*, 14(1):5–24, 1995.

[8] A. Neumaier and T. Schneider. Estimation of parameters and eigenmodes of multivariate autoregressive models. *ACM Transactions on Mathematical Software*, 27(1):27–57, 2001.

[9] E. Persoon and K. Fu. Shape discrimination using fourier descriptors. *IEEE Transactions on Systems, Man and Cybernetics*, 7(3):170–179, March 1977.

[10] W. Phillips, III, M. Shah, and N. da Vitoria Lobo. Flame recognition in video. In *Fifth IEEE Workshop on Applications of Computer Vision*, pages 224–229, December 2000.

[11] G. Schwarz. Estimating the dimension of a model. *Annals of Statistics*, 6:461–464, 1978.

[12] V. N. Vapnik. *The Nature of Statistical Learning Theory*. Springer, second edition, 1999.

[13] M.-H. Yang and N. Ahuja. Gaussian mixture model for human skin color and its application in image and video databases. In *Conference on Storage and Retrieval for Image and Video Databases (SPIE 99)*, volume 3656, pages 458–466, 1999.