Available online at www.sciencedirect.com
SCIENCE DIRECT®
ELSEVIER
International Journal of Thermal Sciences 41 (2002) 1113–1120
International Journal of Thermal Sciences
www.elsevier.com/locate/ijts

An image processing technique for automatically detecting forest fire

Jerome Vicente *, Philippe Guillemant

Institut universitaire des systèmes thermiques industriels, UMR 6595, Technopôle de Château-Gombert, 5 rue Enrico Fermi, 13453 Marseille, France

Received 24 May 2001; accepted 3 January 2002

Abstract

In this paper we present an automatic system for early smoke source detection through the real time processing of landscape images. The first part describes the segmentation technique we use to extract persistent dynamical envelopes of pixels into the images. We describe the temporal algorithm at the pixel level (filtering) and the spatial analysis to bring together connected pixels into the same envelopes (object labeling). The second part deals with the method we use to discriminate the various natural phenomena that may cause such envelopes. We describe the image sequence analysis we developed to discriminate distant smokes from other phenomena, by extracting the transitory and complex motions into little pre-processed envelopes. We present then our main criterion for smoke recognition based on the analysis of the smoke plumes velocity.

Keywords: Computer vision; Smoke recognition; Temporal filtering; Real time image processing; Spatio-temporal analysis; Non-linear analysis; Data clustering; Fractal space-filling curves

1. Introduction

In most cases, it is difficult to contain a forest fire beyond 15 minutes following ignition, and rapid detection is therefore critical. To assist human surveillance, infrared technology has been proposed to detect forest fire with thermal infrared cameras [1]. Until now, these methods do not yield good results for the main reason that the fire itself is often hidden by the trees at the start of its ignition, and the smoke plumes are too quickly cooled to be detected by infrared. More recently, a semi-automatic fire detection system uses infrared satellite images from the Very High Resolution Radiometer (AVHRR) [2–4]. Nevertheless, the satellite permits a detection service at a continental scale, and only at the moments when it passes over the same region. We present in this paper a ground-based video system able to detect fires much more rapidly and which is immune from cloud obscuration. Our system consists of a set of remote CCD cameras covering the supervised zone (Fig. 1) and able to recognize smoke plumes across the visible spectrum at about 3 frames per seconds. Each detector ensures autonomously the basic functions of smoke detection and data transmission. When fire detection occurs on a particular remote analyzer, alarm position on a map and fire images are sent to the control station to obtain quick visualization and the location of the growing blaze.

The fire detection algorithm that we present requires a landscape image analysis in two stages; first the tracking of local dynamical envelopes of pixels, which consist of local time-varying grey levels of connected pixels, and second the discrimination between the various natural phenomena that may cause such envelops. The next section describes the technique we used to extract the local and persistent dynamical envelopes from landscape images. In the last part, we present the identification criteria that we use to discriminate smoke from various natural phenomena envelopes.

2. The dynamical envelope pre-processing

The difficulties of segmenting landscape images are due to their varying nature, to the various illumination conditions and to the large number of dynamical events that may appear. In the case of the distant smokes, the sets of pixels constituting the smoke envelopes are quasi-static and composed of variegated and fleeting motion, dim contrasts,

* Corresponding author.
E-mail address: jevic@iusti.univ-mrs.fr (J. Vicente).

1290-0729/02/$ – see front matter © 2002 Éditions scientifiques et médicales Elsevier SAS. All rights reserved.
PII: S1290-0729(02)01397-2

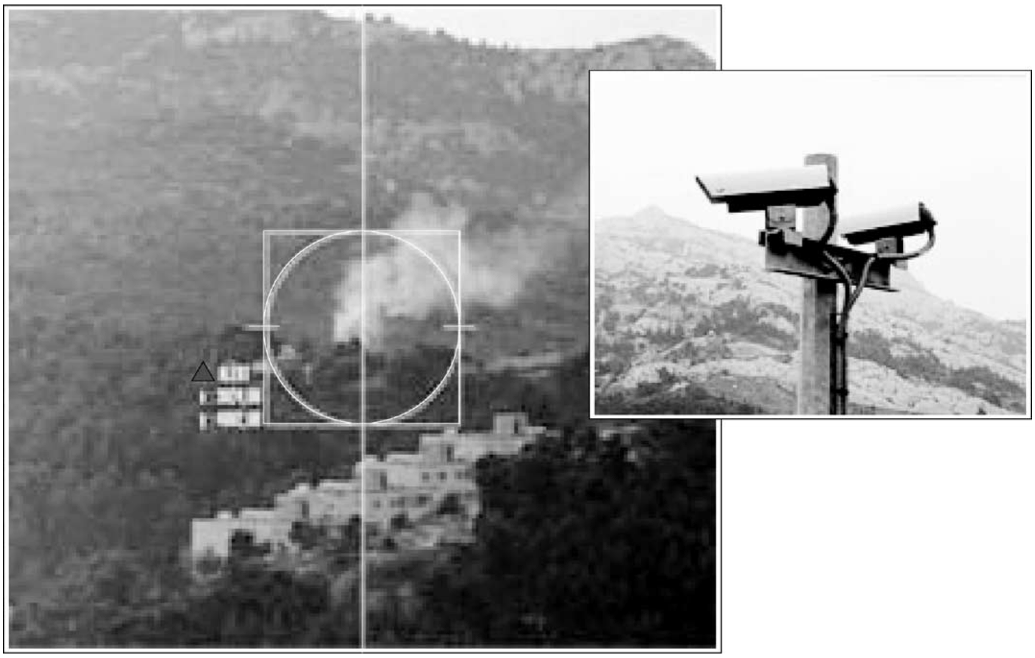


Fig. 1. Automatic forest fire detector system.

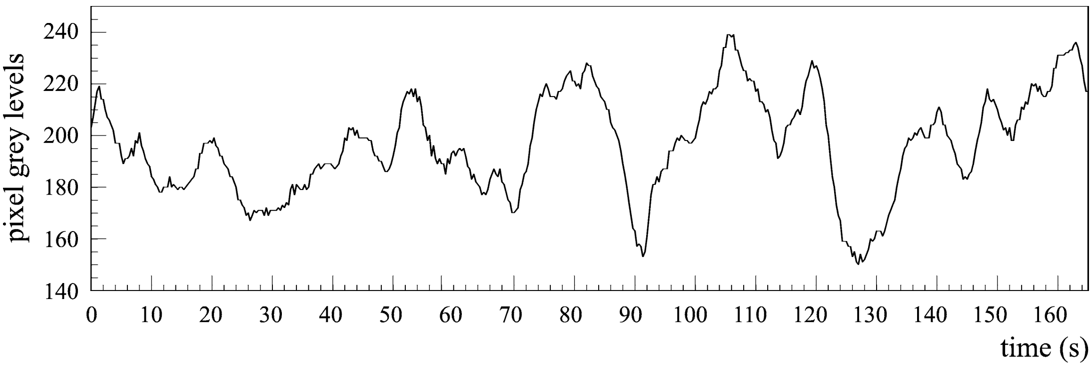


Fig. 2. Smoke pixel grey levels for a 4 km distant forest fire.

and changing contours. This prevent any spatial or texture analysis. To extract the persistent dynamical envelops into the images, we use a temporal analysis at the pixel level (filtering) and a spatial analysis to bring together connected pixels into the same envelopes (object labeling).

2.1. Temporal analysis

2.1.1. Instantaneous dynamical data

The purpose of this treatment is to build a representative image of the dynamical activity of each pixel. According to many distant smoke observations, we propose to retain pixels presenting important dynamical activity but low frequency variations only (Fig. 2). Actually, because of spatial resolution effects, close moving objects present higher pixel grey level variation frequencies than distant moving objects. The only cases where distant pixels present important dynamical and high frequencies correspond to the rapid motions of rigid objects. In the other cases, the grey level variations of distant pixels result from large-scale events, like cloud shadows or local luminosity variations, and present low frequencies variations. This is the case of distant smoke pixels for which the smoke plumes are propagating into a same envelope.

At each sampling time t and for each pixel x of the image, we calculate the instantaneous dynamical information $I(x,t)$, which quantifies the ratio of the slow variations compared with the fastest ones. To avoid calculating a Fourier transform, which is a high cost calculus, the system uses a buffer store composed of the last N images of the scene that allows operating for each pixel a linear combination of its last grey levels:

$$I(x,t) = \sum_{i,j=0,N} \alpha_{ij} \left| l(x,t-i\Delta t) - l(x,t-j\Delta t) \right| \tag{1}$$

with $l(x,t)$ being the grey level of pixel x at time t, and Δt the sampling time of the images. The choice of α_{ij} and the number N of successive images depend on the frequencies we want to sample. This choice is limited by the memory size of the system and by the computing time remaining to carry out this operation in real time.

The sampling period is fixed by the detector system at the value $\Delta t = 1/3$ s. Δt is fast enough to quantify the fast variations of close smokes (500 m for 60-deg CCD lens), whose puffs cross quickly over pixels. At the opposite $N\Delta t$ should be longer enough to quantify the slow variations of distant smokes, because the puff apparent sizes are larger and consequently, the crossing time over pixels is longer. A balance guides to the choice of the optimal segment duration $N\Delta t$, to quantify low frequency in the pixels activity; a higher value of $N\Delta t$ would have required an excessive quantity of memory for the system and a significant increase of the calculation time. A smaller value would have reduced the detection ability of quantifying the slow grey level variations of distant smoke. Due to these conditions and according to the multiple smoke records, we have chosen $N = 16$ and $\Delta t = 1/3$ s.

To quantify the information above the electronic noise, we suppose that for the given images sampling rate (3 frames per second), two successive pixel grey levels differences represent a sample of this electronic noise. We suppose also that the variations of two grey levels temporally separated by N sampling time represent, on the contrary, a significant estimation of the low frequencies variations. As a result, we define $I(x,t)$ as the difference between the low frequencies and the high frequencies estimation as following:

$$I(x,t) = \overbrace{\left|I(x,t)-I(x,t-N\Delta t)\right|}^{\text{slow variation term}} - \Big(\underbrace{\left|I(x,t)-I(x,t-\mathrm{d}t)\right|+\left|I(x,t-(N-1)\Delta t)-I(x,t-N\Delta t)\right|}_{\text{noise variation terms}}\Big) \quad (2)$$

2.1.2. Cumulated dynamical data

We use then the accumulation in time of $I(x,t)$, that we call the cumulated dynamical data $I_c(x,t)$ to localize the areas of persistent dynamical activity, and to obtain stable pixels envelops. The purpose of $I_c(x,t)$ is to obtain a more consistent dynamical information, and to connect the pixel inducing dynamical information in a sufficient time in order to assure the stability of the spatial envelops. We have to join pixels covering the same local dynamical event into stable envelopes.

Actually, a smoke source can generate, during its evolution, some transitory dynamical data that corresponds to the consecutives smoke plumes. We use the cumulated data to extract the smoke pixels and then to obtain stable smoke envelopes. As we will see later in Section 2.2.2, the matching during time of a given envelope depends on the correspondence of its pixels between two successive images.

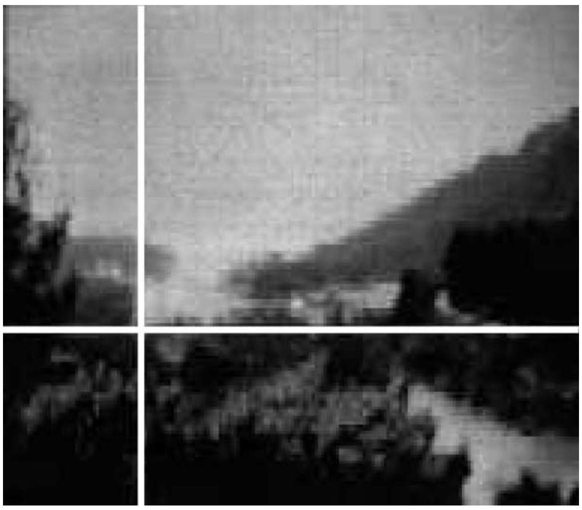
$I_c(x,t)$ is calculated, for each pixel x and at each sampling time t, with a mean algorithm and a temporal weighting as following:

$$I_c(x,t) = (1-\rho)I_c(x,t-\Delta t) + \rho I(x,t) \quad (3)$$

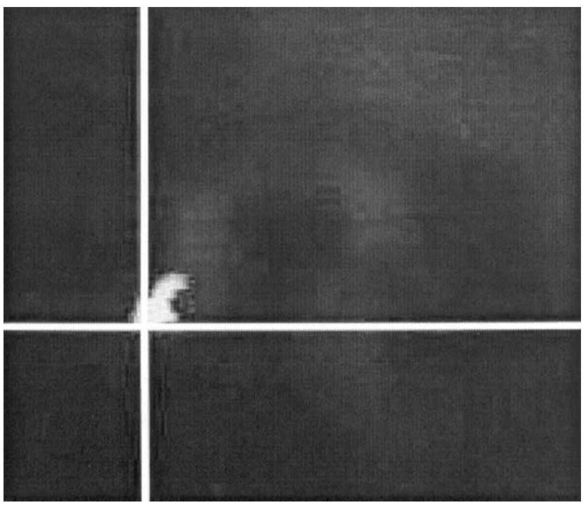
where $\rho = 1/N$ is the time constant depending on the buffer store size.

We could use a new buffer store and a sliding algorithm to calculate $I_c(x,t)$, but to save memory and computing time we prefer the use of the aforementioned method. By cumulating the instantaneous dynamical data $I(x,t)$, we definitely eliminate the pixels affected by the electronic noise or pixels presenting important variations induced by rapid or solid motions.

Eqs. (2) and (3) applied to smoke images sequences allow to obtain the image $I_c(x,t)$ where the intensity is quantifying the dynamical activity according to the previous hypothesis (Fig. 3). As a result, $I_c(x,t)$ is used to identify persistent dynamical local envelopes and directly remove



(a)



(b)

Fig. 3. Smoke plumes into a natural landscape image (a) and cumulated dynamical information $I_c(x,t)$ image (b).

from the analysis too brief events when no dynamical data is locally re-enforced during time.

2.2. Spatial analysis

2.2.1. Envelops decomposition

Once the dynamical data estimation is made, we have to segment the images in order to retain only pixels that constitute the envelopes of dynamical phenomena. We define the envelopes as the areas of the images where the cumulated data $I_c(x,t)$ is important. We obtain such envelops by applying at each sampling time a specific threshold T_c to $I_c(x,t)$ as following:

Let M be the maximum number of pixels with significant value of $I_c(x,t)$, M is also the maximum number of pixels selected into the images. We calculate the $I_c(x,t)$ values histogram to determine the threshold T_c for which at least M pixels present a cumulated data $I_c(x,t)$ above T_c. If this threshold is less than a minimal value admitted, I_o, we choose $T_c = I_o$ to remove too noisy data. After this operation, only pixels with significant dynamical contents and time stability are selected.

The first remark is that we have to increase M for the highest resolution images or images presenting important dynamical events. Secondly, we apply the same threshold to the whole image of $I_c(x,t)$. In the presence of important dynamical activity, the sensibility of the system decreases. In this case, we can reduce the value of T_c and report later on the analysis of the supplementary pixels through an off-line process.

2.2.2. Objects labeling

The object labels must persist to the probably changing shape of each envelope but also to the weak displacement of the whole envelope. Fig. 4 illustrates our label affectation principle applied to two consecutive envelopes which present the maximum number of common pixels. Nevertheless, the segmentation presented here is not enough to perform the smoke recognition at only one go. Actually, we observe that many natural events can generate similar dynamical envelopes (clouds shadow, illumination variations on the ground relief, ...). In the next part we present the identification criteria we use to definitely discriminate smoke envelop from the other dynamical events by the analysis of the spatio-temporal content of each envelope.

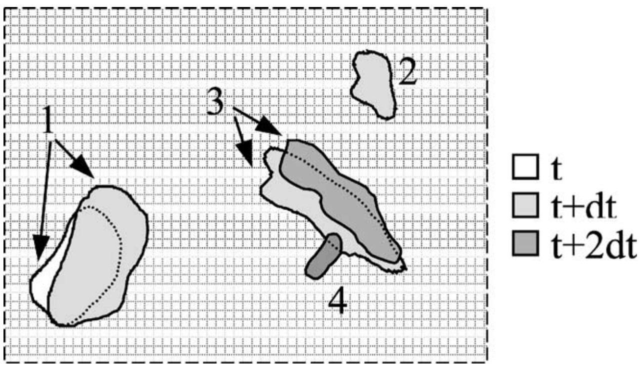


Fig. 4. Label affectation for the temporal following of dynamical envelopes.

3. Smoke pattern analysis

3.1. Decorrelation criteria

We have demonstrated in [5] the non-linear behavior of the grey levels variations of a single smoke pixel using the estimation of the fractal dimension [6] combined with the surrogate data method [7]. More precisely, we observed that the smoke envelope forms a complex dynamical structure, which is comparable to a low frequency spatio-temporal noise [8]. Fig. 5 shows typical grey levels variations of smoke pixels displayed perpendicularly (Fig. 5(a)) and along (Fig. 5(b)) to the main orientation of the smoke plumes propagation. We observe in Fig. 5 typical spatial decorrelation of the grey levels that corresponds to the spatial vanishing of the smoke plumes and that depends on the spatial resolution of observation.

The smoke identification described in the patent [8] consists in the calculus of several spatial decorrelation criteria. To quantify such criteria, we calculate, at each sampling time, the distribution value histogram of the grey levels differences of all the pairs of pixels contained into a pre-processed envelop. By cumulating this histogram in time, we obtain, for all the smoke envelopes analyzed, a gaussian distribution of the grey levels differences that is translated into a typical spatio-temporal behavior of smoke patterns [5].

However, events as fog/haze, clouds but mainly clouds shadows on changing relief, show a complex dynamical structure, comparable to the one produces by the smoke. This is principally due to the similar spatial texture of the envelops produced by such phenomena. We propose to discriminate those phenomena by analyzing the temporal correlations into each envelops. Fig. 5(b) shows the temporal signatures that propagate trough pixels with different temporal delays induced by the motions of the smoke plumes at various propagation speeds. These observations proceed from a large temporal analysis window. We present in the following a real time method to identify these features by analyzing shorter temporal pixels evolution.

3.2. Motion detection with a temporal embedding method

To improve definitively the forest fire detector, we have proposed to calculate the complex smoke plumes motions from the previous extracted envelopes. The difficulty in processing such envelopes is that the motions to detect are composed of fleeting and not sharply contrasted fronts that propagate slowly on an already smoke-filled background. Given the small size of the envelopes to process, the motion detection methods based on spatial analysis, which are

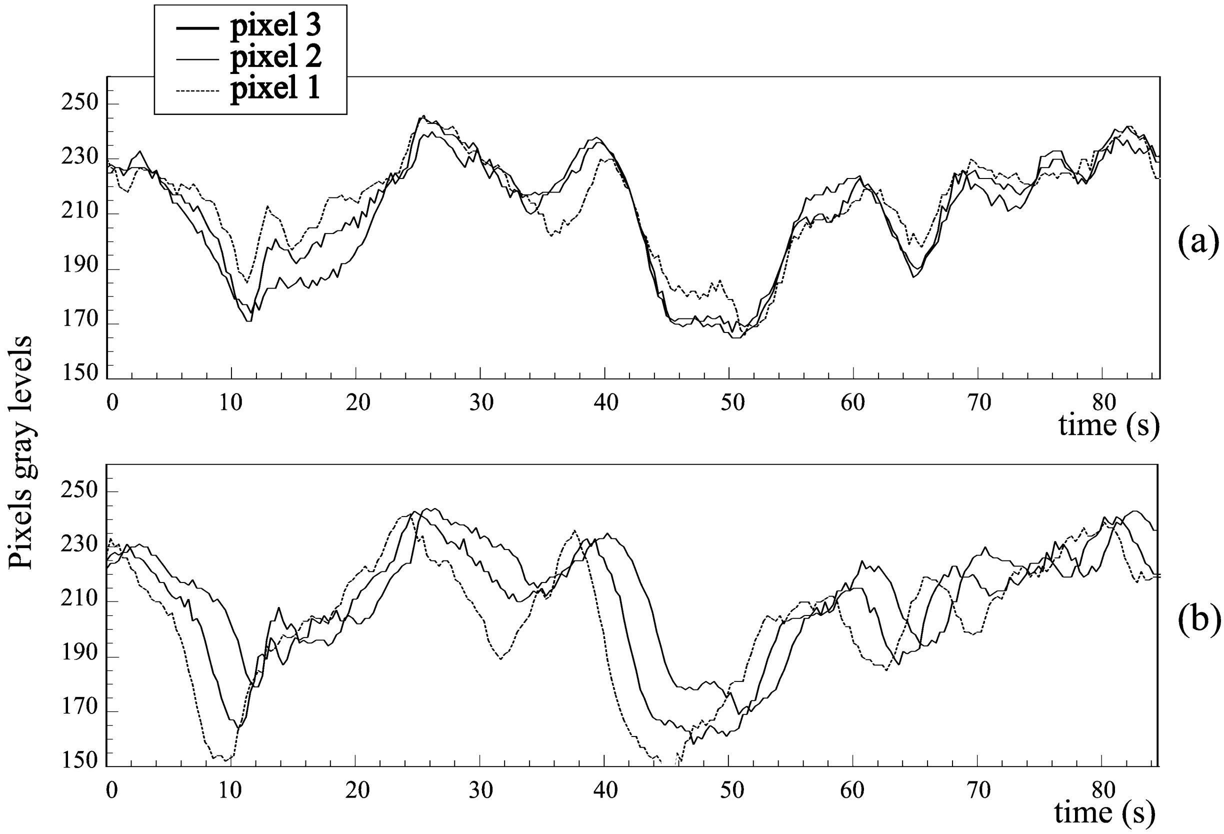


Fig. 5. Temporal variations of smoke pixels displayed perpendicularly (a) and along (b) the smoke plumes propagation direction.

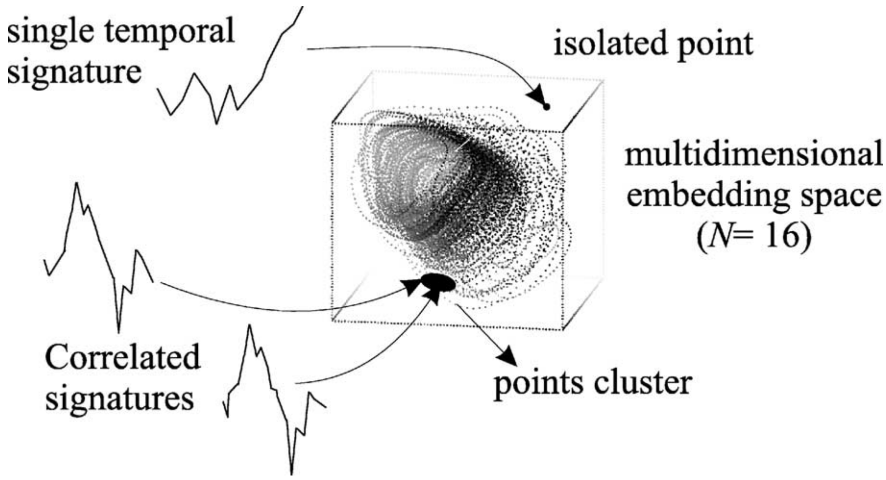
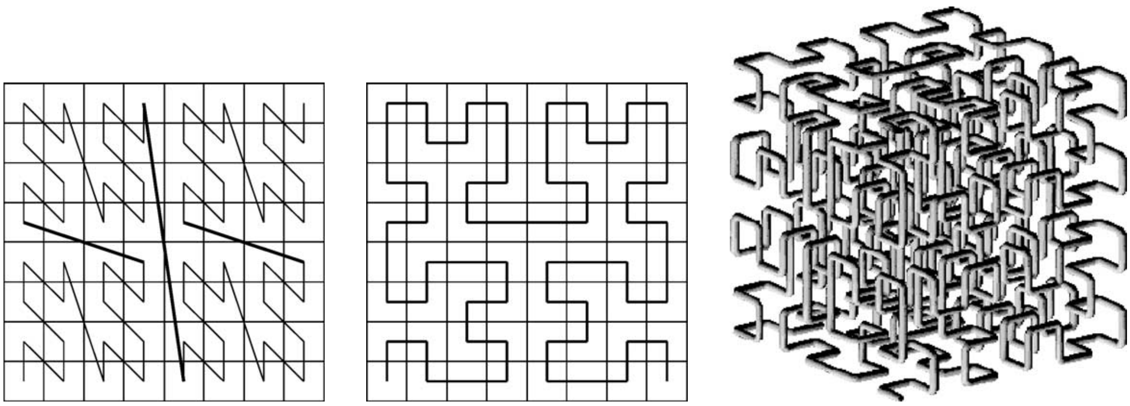


Fig. 6. Cluster formation of the correlated signatures into the multidimensional embedding space.

rather adapted to solid motion tracking, are unsuited [9,10]. However, we observe that the motion of a smoke plume creates the repetition of identical-looking temporal variation along its propagation trajectory (Fig. 5(b)). We can clearly observe an average time delay that corresponds to the average puff propagation speed. To retrieve the trajectory of a single puff only, we have to track short temporal variations. These temporal signatures correspond to the average crossing time ($N\Delta t$) of the puff over the area covered by on pixel. The method we propose is to identify all the correlated temporal signatures in order to reconstruct the smoke plume trajectories. The number of signatures to process in a real time analysis for a single envelope is approximately equal to the product of the envelope area (some hundred of pixels) with the number of temporal samples (some tens time steps). It is then current to compare 10 000 points in only one time step. To extract all the correlated signatures in real time, we propose to perform a cluster extraction into a multidimensional space, where the temporal signatures are embedded [11] (Fig. 6).



(a) (b) (c)

Fig. 7. Space-filling curves examples: (a) z-curve in dim 2, (b) Hilbert curve in dim 2, (c) Hilbert curve in dim 3.

Clustering techniques are most often used for data mining, which is the process of extracting useful information from very large data sets. Many clustering techniques exist, such as k-mean algorithm or r-tree techniques [12]. Such indexations are obviously not general and a recalculation is necessary for any new configuration of the data. The majority of these techniques are not adapted to real time process, and mainly not adapted to a dynamical process when the points are sequentially embedded.

Our real time clustering technique uses the indexing of the multidimensional embedded points along a space-filling curve [13], with an original chaining technique of these points that we detail in [11]. The fractal property of the space-filling curves (Fig. 7) allows us to consider that neighbor points in the space are neighbor points, or consecutives, on the fractal curve. If we disregard the recovering limitation of the one-dimensional space-filling indexing [14], finding cluster points in space boils down to find close consecutive points on the curve. With this method, we can recuperate the closest signature of any new incoming signature and then estimate if this signature belongs to an identified trajectory or not. In the following section, we detail the three successive stages of our cluster extraction method: indexing, chaining and cluster identification.

3.2.1. Fractal indexing and chaining

At each sampling time and for each pixel of the envelope, we embed a multidimensional point formed with the last N grey level of the pixel. To index each embedded point we use its curvilinear coordinates on the space-filling curve. We call the index the fractal rank. To achieve this indexing, we use the "Z-curve" initially proposed by Orenstein [15] (Fig. 7(a)) which yields to the fastest fractal rank calculation through a transposition of the bit coordinates matrix:

$$z(X) = \sum_{j=0}^{e-1} \sum_{i=0}^{N-1} 2^{i+jd} x_i^j \tag{4}$$

where x_i^j is the jth bit of the e-bytes grey level $x_i = l(x, t - i\Delta t)$ of the signature X.

The next step of the method consists in storing each new incoming point into a memory table in the upward order of their fractal rank. By truncating the fractal rank, we classify all possible ranks within a limited number of classes, thus not exceeding the standard memory of computers. Each new point is then chained inside a hypercube, or class, containing a very small number of already-chained points. This limits the number of rank comparisons to be carried out, and therefore the computing time.

This chaining technique allows retrieving, for each new point, its best neighbors points into the embedded space, just by scanning up or down the linked list. Thanks to a few operations, considering the clustering properties of the space-filling curves, we are then able to estimate the number of occurrences of one temporal signature into the past of the pixels envelope history and consequently to extract the temporal correlated data.

3.2.2. Real time motion detection algorithm

The objective of our real time processing algorithm (Fig. 8) is to provide a diagnosis based on the cumulative motion estimation. At each sampling time, temporal signatures of the last $N = 16$ grey levels of each pixel are embedded, indexed and chained into the linked list of the envelope. Then we compare, during the clustering procedure, each new embedded point with the Δ points in its immediate neighborhood into the list. The number Δ of the comparisons depends on the calculation time available for the process. As this calculation time depends on the number of pixels that are embedded at each sampling time, Δ varies in inverse proportion to the number of pixels of the envelope. Minimum $\Delta_{\min}$ and maximum $\Delta_{\max}$ are used to keep Δ within reasonable values. We calculate then the histogram of the instantaneous velocities (IVH) of the Δ couples of points in the neighborhood points of the list with the new embedded one ($\Delta/2$ before this point, $\Delta/2$ after this point). Knowing the space and time coordinates of all points, the quotient of their differences with the new embedded point coordinates is calculated and added to the IVH. ISD is the instantaneous standard deviation of the histogram IVH.

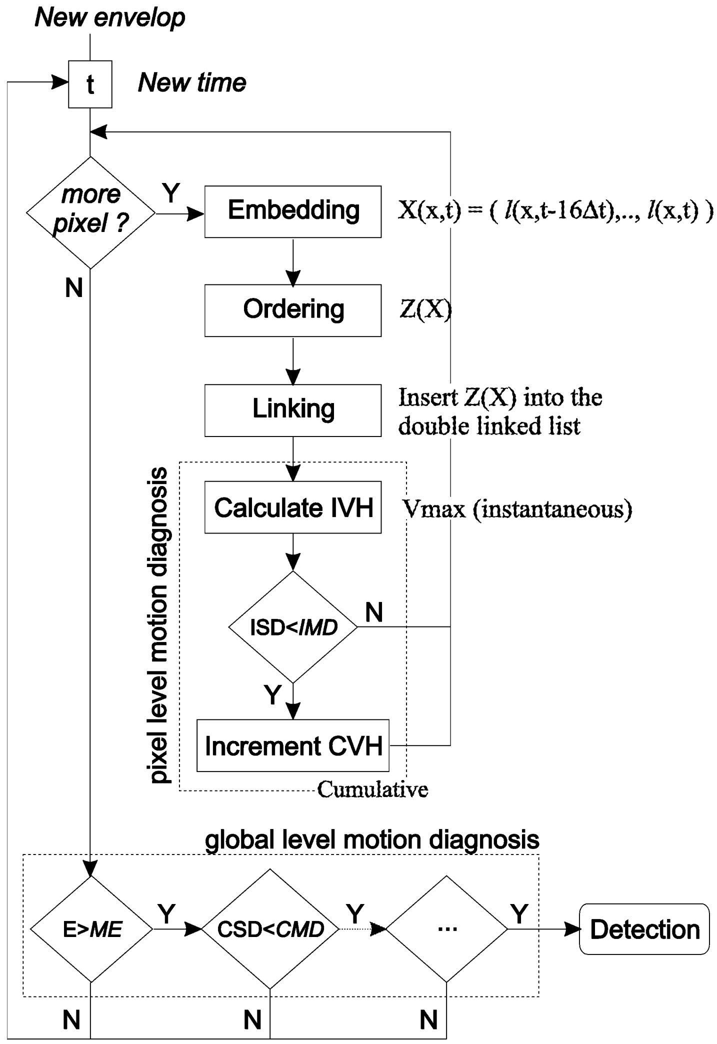


Fig. 8. Real time detection algorithm using the 4 steps: Temporal embedding, fractal ordering, linking points and motion clustering.

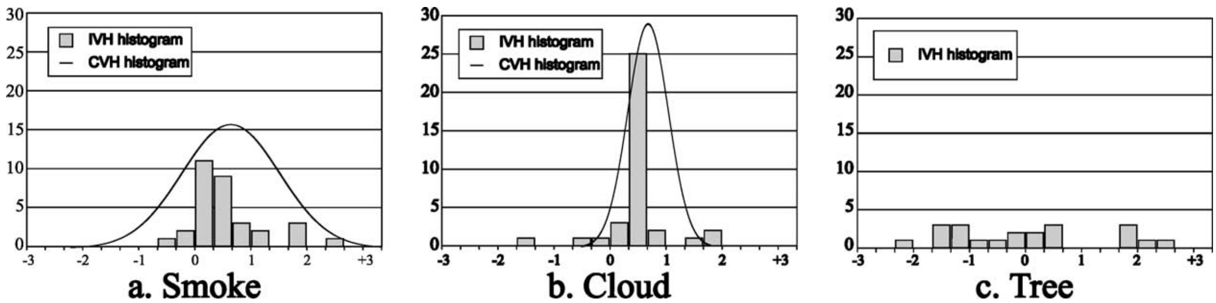


Fig. 9. Instantaneous velocity histograms (IVH) for on point neighborhood and cumulative velocity histograms (CVH) for all the point neighborhoods.

Fig. 9 shows representative results calculated for the envelopes of a cloud, a smoke and a wind tossed tree, the smoke being at a distance of 4 km. We notice no significant results in the case of the tree and in contrast a very well defined peak for the cloud example.

When examining various types of smoke results, we always observe a spreading histogram with a certain variability of the standard deviation, but that always keeps above that of the cloud. Conversely, we consider that there is no motion detected if the value of the instantaneous standard deviation ISD is above a certain threshold IMD. This parameter is included in the initialization file of the system. If any motion is detected, the velocity of the maximum of IVH, $V_{\max}$ is then incremented into a cumulative velocity histogram CVH. CSD is the standard deviation of CVH. CVH is at last used to calculate criteria in order to distinguish smoke from other phenomena.

In the case of smoke and clouds, we observe that CVH (Fig. 9(a), (b)) has the same shape as IVH, but is smoother and the standard deviation CSD is stable. When CSD is less than a minimum standard deviation CMD, we invalidate smoke detection for this sampling time (Fig. 8). The use of the CMD test poses a problem only in the rare case when certain smoke envelopes are stretched by a very stable wind, so that their velocity distribution resembles that of a cloud. However, that cannot last, and the only consequence is that the smoke detection time is increased. When wind speed is zero, the motions in the case of smoke are induced by the thermal convection. These motions are still detected by our method, but are slower than those obtained during important wind conditions. The result of this is also to increase the detection time.

In the case of wind tossed trees (Fig. 9(c)) and many other phenomena, the cumulative histogram is most often rather poor in data, because motions are detected in the case of point like situations and for precise pixels only. A second criterion is then the minimum average energy ME of the cumulative histogram per embedded point, which is in other words the average number of instantaneous motion diagnosis per embedded point. This criterion is the most selective one, as it separates smoke from a lot of various pseudo-dynamic envelopes that are not eliminated by the envelope extraction pre-processing.

In addition to CMD and ME, other criteria are used, for example, those based on the shape and on the smoothness of the cumulative histogram; but they are more questionable and have less generality. Among the other criteria is the sense of propagation (top or bottom) that we calculate from the angular distribution analysis.

One of the main interests of this temporal embedding algorithm is that it is sensitive enough to detect local motions by means of only a few pixels. So it is particularly adapted to the analysis of small dynamic deformable objects with multiple transitory motion paths, such as little envelopes of smoke. The different processing steps described in this part lead first to a local motion diagnosis at the level of one pixel, and second to a global motion diagnosis at the level of the whole envelope. These diagnoses are used for a complete smoke identification and are actually implemented into the fire forest detector ARTIS FIRE, commercialized by T2M Automation.[1] Validation campaigns permitted to estimate an average smoke detection time below 3 minutes following the first visible smoke emissions.

4. Conclusion

We have experienced that the more efficient data for smoke identification is the velocity distribution of smoke plumes, whose energy is higher than the energy of many other landscape phenomena, except clouds. However, in the case of clouds, the standard deviation of velocity distribution is generally lower than that of the smoke. Then our main criterion for smoke detection is based on the analysis of velocity distribution, using a minimum energy threshold and a minimum standard deviation threshold. We have shown that our method is able to provide a direct access to the useful correlated data when a fast process is needed. The "fractal embedding" method does not use any model or parameter introduced a priori in the calculation for the extraction of a result. For example, no threshold is used to quantify any correlation between two short temporal segments. One should be especially aware that the definite advantage of fractal chaining and clustering is to considerably reduce computing time. This is of paramount importance when processing large amount of data, such as image sequences, in particular in the case of real time computing.

[1] Aérospatiale-Matra, Tests et Services T2M, Zone d'Activités Louis Bréguet, BP 6–Bâtiment T2, 78141 Vélizy-Villacoublay cedex, France.

References

[1] S.H. Chan, M. Osumi, M.J.D. Low, An analytical feasibility study of forest fire detection by infrared dual sensors, in: Symposium on Environmental and Geophysical Heat Transfer, November 28–December 2, 1971.

[2] L. Giglio, J. Kendall, C.O. Justice, Evaluation of global fire detection algorithms using simulated AVHRR data, Internat. J. Remote Sensing 20 (1999) 1947–1986.

[3] T.M. Lillesand, R.W. Kiefer, Remote Sensing and Image Interpretation, 4th Edition, Wiley, New York, 2000.

[4] T. Häme, Y. Rauste, Multitemporal satellite data in forest mapping and fire monitoring, EARSeL, Adv. Remote Sensing 4 (3) (1995) 93–101.

[5] J. Vicente, P. Guillemant, Identification de sources de fumées par plongement fractal, in: Proceedings of French Thermal Society, SFT 2000, Lyon, 2000, pp. 631–636.

[6] P. Grasberger, I. Procaccia, Generalized dimension of strange attractors, Phys. Lett. A 97 (1983) 227–230.

[7] J. Theiler, S. Eubank, A. Longtin, B. Galdrikian, D. Farmer, Testing for nonlinearity in time series—the method of surrogate data, Phys. D 58 (1992) 77–94.

[8] Brevet d'invention : N° de publication : 2750870, N° d'enregistrement national : 9609030, Procédé de détection automatique de feux, notamment de feux de forêts.

[9] P. Bouthemy, Motion segmentation and qualitative dynamic scene analysis from an image sequence, Internat. J. Comput. Vision 10 (2) (1993) 157–182.

[10] D. Terzopoulos, A. Witkin, M. Kass, Constraints on deformable models: Recovering 3D shape and nonrigid motion, Artif. Intell. 36 (1988) 91–123.

[11] P. Guillemant, J. Vicente, Real time identification of smoke images by clustering motions on a fractal curve with a temporal embedding method, Optical Engrg. 40 (4) (2001) 554–563.

[12] T. Sellis, N. Roussopoulos, C. Faloutsos, The R+Tree: A dynamic index for multi-dimensional objects, in: Proceedings of the 13th International Conference on Very Large Data-bases, Brighton, England, 1987, pp. 507–518.

[13] H. Sagan, Space-Filling Curves, Springer-Verlag, Berlin, 1994.

[14] H.V. Jagadish, Linear clustering of objects with multiple attributes, in: Proceedings of ACM SIGMOD, 1990, pp. 332–342.

[15] J.A. Orenstein, Spatial query processing in an object-oriented database system, in: Proceedings of ACM SIGMOD, 1986, pp. 326–336.