

Détection de fumée dans des zones de forêt à partir de séquences vidéo codées

Noura HAMOUDA¹, Amel BENAZZA-BENYAHIA¹, Fethi TLILI²,

¹Unité de Recherche en Imagerie Satellitaire et ses Applications,
SUP'COM, Univ. de Carthage, Cité Technologique des Communications, Tunisie

²Unité de Recherche en Circuits et Techniques Avancées pour les Systèmes de Communications,
SUP'COM, Univ. de Carthage, Cité Technologique des Communications, Tunisie *

noura.hamouda@yahoo.fr, benazza.amel@supcom.rnu.tn, fethi.tlili@supcom.rnu.tn

Résumé – Des fonctions dites intelligentes commencent à être intégrées dans les systèmes de vidéosurveillance au niveau de l'acquisition et de la compression de la séquence dynamique par la caméra. A cet égard, la détection de la fumée et l'envoi d'alarmes dans le flux binaire de la vidéo codée constituent des fonctionnalités désirables pour des applications de surveillance des zones de forêts. Dans cet article, nous proposons une procédure de détection rapide et précoce qui exploite le caractère fractal de la fumée uniquement à partir des informations de codage de la séquence au moment de son acquisition et ce, selon les deux normes de compression vidéo MJPEG ou MPEG2 couramment disponibles dans la plupart des caméras.

Abstract – Smart functions begin to be integrated into videosurveillance systems at the acquisition and the compression steps of the dynamic sequence by the camera. In this regard, early detection of smoke through emission of alarms into the encoded video bitstream are attractive features for monitoring applications of forest areas. In this paper, we propose a fast and early smoke alarming that exploits only informations generated at the compression step of the video after its acquisition according to the two coding video standards MJPEG or MPEG2 standards widely available in the operating cameras.

1 Introduction

Les forêts ainsi que les zones boisées du bassin méditerranéen présentent des risques élevés d'incendie notamment durant les périodes de sécheresse estivale : en moyenne, 50.000 incendies détruisent chaque année près de 600 000 ha [1]. La détection précoce de feu s'avère donc indispensable pour préserver ces écosystèmes méditerranéens et le patrimoine culturel placé dans de tels sites et à cet égard, il est d'usage de déployer des capteurs de fumée, température et humidité [2]. Toutefois, l'efficacité de leur détection dépend fortement de leur proximité de la source de feu et s'avère souvent limitée dans des espaces ouverts et/ou de grande taille. De plus, de tels capteurs ne fournissent aucune information sur la localisation du feu, ni sur son étendue. Pour pallier de telles insuffisances, une solution facile et peu coûteuse est de recourir à des techniques de vision par ordinateur [3, 4]. A ce titre, le projet FIRESENSE¹ auquel nous participons s'attache à développer un système de détection d'incendie dans des sites naturels du patrimoine archéologique [5]. A cet égard, un intérêt soutenu a été porté à la détection de fumée qui est un indicateur annonciateur de feu de type méditerranéen et permet donc une action précoce de la part

des équipes de surveillance. Néanmoins, détecter la fumée est un réel défi principalement en raison de la difficulté de modéliser fiablement la fumée (densité variable spatialement et temporellement) et aussi des variations de l'illumination, des occlusions [6]. Dans ce contexte, plusieurs méthodes de détection de fumée ont été développées [7, 8, 9, 10]. Généralement, elles exploitent différents types d'information sur la fumée comme la couleur, le mouvement et le désordre de la forme de la fumée [9]. Concernant la couleur, l'espace de représentation majoritairement retenu est l'espace RVB dans lequel les pixels des zones de fumée ont des valeurs égales dans les 3 composantes [4, 11, 12] ce qui permet de construire des tests comme celui reposant sur la décroissance dans le temps de la valeur de la chrominance [11]. Les paramètres dynamiques reflétant les oscillations des contours de la fumée peuvent être extraits dans le domaine transformé en ondelettes [8, 13]. L'évolution désordonnée au cours du temps des zones de fumée (variations de leur taille, irrégularité de leur contour ...) s'avère également utile pour la détection de la fumée [6, 11]. Récemment, il a été proposé une approche s'appuyant sur le principe que la fumée lisse les contours des objets présents dans la scène [14]. Une autre stratégie consiste à caractériser la texture des zones de fumée par le biais de la matrice de co-occurrence [15] ou par l'exposant de Hurst mesurant la rugosité de la région d'intérêt [16]. Dans ce travail, nous conservons le même principe à savoir ex-

* Ce travail a été financé par le projet FIRESENSE (# 244088) dans le cadre du 7ème programme sur l'environnement de l'Union européenne.

¹<http://www.firesense.eu>

exploiter le caractère fractal des zones de fumée mis en évidence récemment dans les travaux de Maruta *et al.* [16]. Notre principale contribution réside dans le fait que le système de détection conçu a pour spécificité de n'utiliser que l'information de codage de la séquence au moment de son acquisition. En effet, nous visons à intégrer dans le système de capture/compression d'images un détecteur précoce de fumée capable d'alerter rapidement le centre de surveillance. En ce sens, nous voulons doter la caméra d'une fonction dite intelligente de détection. Plus précisément, notre principale contribution est de mesurer l'exposant de Hurst reflétant le caractère fractal local sur les coefficients de la Transformée en Cosinus Discrète (TCD) selon les deux normes de compression MJPEG ou MPEG2 couramment disponibles dans la plupart des caméras. Cet article est organisé comme suit. Un bref rappel sur de ces deux normes sera établi et le principe général de la stratégie proposée sera décrit en section 2. La section 3 est dédiée à la caractérisation fractale des zones de fumée. Les résultats expérimentaux sont présentés en section 5 et les conclusions sont fournies en section 6.

2 Méthode proposée

2.1 Bref rappel sur les normes de compression

La norme MJPEG effectue la compression d'une séquence vidéo en codant les images séparément selon la norme JPEG. Seules les redondances spatiales sont exploitées par le biais de la TCD appliquée sur des blocs disjoints de taille 8×8 . La TCD permet de représenter l'image dans le domaine fréquentiel avec des valeurs réelles des coefficients contrairement à la transformée de Fourier discrète. Les propriétés remarquables de la TCD sont la décorrélation des coefficients et la concentration de l'énergie de la plupart des images naturelles sur un faible nombre de coefficients des basses fréquences. Il s'agit d'un avantage majeur en compression. En effet, étant peu énergétiques, les coefficients majoritaires peuvent être quantifiés grossièrement voire annulés. Par conséquent, la qualité de reconstruction peut être acceptable pour un débit relativement faible. La norme de compression vidéo MPEG2 opère sur des séquences vidéo découpées en sous-séquences GOP (*Group Of Pictures*) et le codage d'un GOP se fait indépendamment des autres afin de garantir une robustesse par rapport aux erreurs de transmission et un accès rapide aux images de la séquence. Un GOP est composé de 3 types d'images : les images I, P et B. En général, chaque GOP commence par une image I codée en mode intra selon un schéma similaire au codeur JPEG. Le codage des images P et B est hybride combinant la prédiction sur des macroblocs 16×16 et la TCD. Pour les images P et B, la prédiction peut être en mode intra-image ou inter-images. Le premier mode vise à réduire la redondance spatiale entre blocs voisins tandis que le second mode exploite la redondance temporelle par compensation du mouvement. La quantification et la génération du flux binaire sont les dernières opérations de cette chaîne de codage. Enfin, notons que la composition com-

mune à tous les GOP d'une séquence est choisie au début de la session de codage.

2.2 Principe

Le principe adopté est de caractériser les propriétés texturales spatiales des zones de fumées. Si la norme MJPEG est utilisée, la détection se fait sur des images séparées d'un décalage k_0 . Dans le cas de la norme MPEG2, le détecteur s'applique sur la première images I de chaque GOP et k_0 désignera alors la longueur commune à tous les GOP. Le principe est de déterminer dans ces images (appelées images-cibles) les blocs en mouvement. Puis, il s'agit de discriminer parmi ces blocs en mouvement ceux susceptibles de contenir des zones de fumée. Pour ce faire, nous supposons que les blocs de fumée correspondent à une texture de rugosité particulière caractérisée par une plage de valeurs de l'exposant de Hurst [19]. Autrement dit, une classification binaire des blocs candidats est effectuée sur la base de leur attribut fractal (exposant de Hurst). Dans ce qui suit, nous détaillons chacune des étapes de notre méthode.

3 Caractérisation fractale

3.1 Détection des objets en mouvement

Déterminer les objets en mouvement se fait en ligne et les contraintes de temps réel incitent à recourir à une technique basique mais peu coûteuse comme la soustraction d'images. Dans le cas de la norme MJPEG, il faut choisir la trame $I(m, n, k - k_1)$ à laquelle est soustraite la trame candidate $I(m, n, k)$. La valeur du décalage k_1 dépend de la fréquence temporelle d'échantillonnage et de la vitesse moyenne de la fumée : typiquement dans nos tests, un décalage d'environ 3 à 5 secondes s'avère suffisant. On considère qu'un bloc 8×8 de l'image courante appartient à une zone en mouvement si :

$$\sum_{m=1}^8 \sum_{n=1}^8 |I(m, n, k) - I(m, n, k - k_1)| > S \quad (1)$$

où S est un seuil ajusté après une étape d'apprentissage vu les contraintes de temps réel. Plus précisément, des seuils adaptatifs ont été calculés automatiquement sur les images-cibles de différentes séquences d'apprentissage en s'appuyant sur l'entropie des images [18]. En phase de test, la valeur retenue pour S correspond à la moyenne de ces seuils.

3.2 Calcul de l'attribut fractal

Les blocs en mouvement ainsi déterminés dans les images-cibles sont alors considérés comme des blocs candidats potentiels pouvant contenir de la fumée ou pas. La deuxième étape vise donc à leur assigner un label binaire c tel que pour un bloc d'une région de fumée, $c = 1$. L'attribut retenu pour guider cette classification est le paramètre de Hurst qui reflète le

caractère fractal des zones de fumée comme l’ont mis en évidence Maruta *et al.* dans [16]. Par conséquent, le défi est d’estimer pour chaque bloc candidat b la valeur de H_b en n’utilisant que l’information disponible au niveau du codeur puis de mettre en œuvre un classificateur binaire agissant sur les attributs ainsi estimés. Dans le cas 1D, rappelons qu’un signal auto-similaire a sa Densité Spectrale de Puissance (DSP) qui évolue en $1/f^{2H_b+1}$ où f est la fréquence et H_b le paramètre de Hurst [19]. Par conséquent, il est possible d’approximer H_b par la pente du logarithme d’un estimateur de la DSP. A cet égard, recourir à un algorithme itératif robuste aux données aberrantes tel que l’algorithme Ransac n’est pas envisageable en temps réel. Il est préférable d’opter pour un estimateur de H_b moins complexe tel que celui des moindres carrés. Dans notre cas, en 2D, nous devons d’abord disposer d’un estimateur de la DSP en fonction des fréquences cartésiennes (u, v) . Le périodogramme est un estimateur simple et assez naturel de la DSP. Toutefois, il repose sur les coefficients de la transformée de Fourier discrète alors que seuls les coefficients de la TCD sont disponibles. Nous proposons de calculer le périodogramme à partir de ces seuls coefficients TCD. Par la suite, nous passons du système de coordonnées fréquentielles cartésiennes (u, v) à celui des coordonnées polaires (f, θ) :

$$u = f \cos(\theta), v = f \sin(\theta). \quad (2)$$

Pour une valeur fixée de θ , la DSP d’un signal auto-affine fractal est en $1/f^{2h_b(\theta)+1}$ où f est la fréquence radiale. Un estimateur de $h_b(\theta)$ peut être facilement obtenu selon le critère des moindres carrés en considérant le logarithme de la DSP estimée. Il suffit alors que θ décrive un ensemble fini de valeurs Θ pour moyenniser les valeurs de $h_b(\theta)$ et obtenir un estimateur \hat{H}_b du paramètre de Hurst du bloc b candidat [20] :

$$\hat{H}_b = \frac{1}{\text{card}[\Theta]} \sum_{\theta \in \Theta} \hat{h}_b(\theta). \quad (3)$$

4 Classification des blocs candidats

Une fois l’estimation des exposants de Hurst effectuée, il faut les classer en deux classes. Pour cela, une étape préalable d’apprentissage est envisagée. A partir d’une collection de blocs pris dans des zones de fumée (classe $c = 1$) et d’une collection de blocs pris dans des régions où la fumée est absente ($c = 0$), sont ainsi formés deux ensembles des valeurs estimées de leurs paramètres de Hurst. La Figure 1 affiche pour chaque classe de la vérité-terrain, les valeurs des quantiles de la loi empirique comparées aux quantiles de la loi normale centrée réduite. Ces tracés confirment l’hypothèse gaussienne de la répartition des valeurs de l’attribut fractal. Par conséquent, la répartition des valeurs se fait pour chaque classe c selon une distribution gaussienne dont les paramètres sont les moyenne et écart-type approchés par les estimateurs empiriques \bar{H}_c et σ_c calculés une fois pour toutes en phase d’apprentissage. Par la suite, en phase test, il suffit d’estimer la valeur \hat{H}_b du paramètre de Hurst de chaque bloc b en mouvement. Le bloc

candidat b est affecté à la classe fumée ($c = 1$) si :

$$\left| \frac{\hat{H}_b - \bar{H}_1}{\sigma_1} \right| < \left| \frac{\hat{H}_b - \bar{H}_0}{\sigma_0} \right|.$$

Dans le cas non-trivial $\sigma_0 \neq \sigma_1$, il est facile de voir que cela revient à des comparaisons selon le seuil η défini par $\eta = \frac{\sigma_1 \bar{H}_0 - \sigma_0 \bar{H}_1}{\sigma_1 - \sigma_0}$ et calculé au préalable une fois pour toutes en phase d’apprentissage.

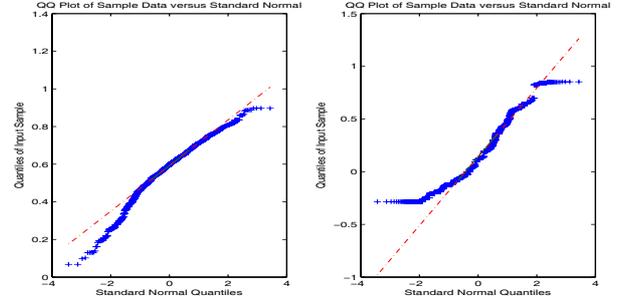


FIG. 1 – Tracé des quantiles des lois empiriques des attributs fractals de blocs 8×8 et ceux de la loi normale centrée réduite pour la classe $c = 1$ (à gauche) et l’autre classe (à droite).

5 Résultats expérimentaux

Pour l’apprentissage, chaque classe est formée de 1680 blocs 8×8 extraits manuellement de 60 images de luminance de 3 séquences vidéo donnant les paramètres statistiques suivants : $(\bar{H}_1, \sigma_1) = (0.582, 0.1359)$ et $(\bar{H}_0, \sigma_0) = (0.1543, 0.2826)$. Ensuite, la détection a été appliquée à L images de luminance des séquences tests V1, V4, V5 et V7 du projet FIRESENSE dont les caractéristiques sont résumées dans le tableau 1. Les performances du détecteur sont mesurées par les taux moyens de fausses alarmes P_f , de non détections incorrectes P_m et d’erreurs globales de classification P_e . Le décalage Δ entre l’image où apparaît réellement la fumée et l’image où notre détecteur commence par détecter de la fumée permet de mesurer la précocité de la détection. Les tableaux 2-(a) et 2-(b) fournissent ces mesures pour les 2 normes étudiées. Il apparaît que l’algorithme de détection proposé aboutit à un taux de fausse alarme inférieur à 0.43 et un taux de non détection faible inférieur à 0.18. La rapidité de la détection est assez appréciable puisqu’un décalage temporel Δ quasi nul existe entre l’apparition de la fumée et sa détection. La figure 2 visualise les blocs détectés en mouvement classés en $c = 0$ (étoiles bleues) et $c = 1$ (étoiles rouges) dans 2 images test et indique que le détecteur parvient à localiser correctement les zones de fumée.

6 Conclusion

Dans cet article, nous avons présenté une nouvelle méthode de détection de fumée exploitant le caractère fractal de la fumée. La nouveauté dans notre approche est que la détection se

fait directement à partir du flux vidéo codé en estimant l'exposant de Hurst reflétant le caractère fractal local sur les coefficients de la TCD. L'algorithme s'articule autour de 2 étapes : l'extraction des objets en mouvement par simple soustraction d'images et la classification des blocs détectés après une phase d'apprentissage. Les résultats sur des séquences de scènes naturelles indiquent des résultats encourageants. Une perspective laissée ouverte par ce travail est d'étudier l'apport d'autres informations complémentaires comme la chrominance pour une détection précoce et peu coûteuse de la fumée.

TAB. 1 – Caractéristiques des séquences tests.

Séquence	Cadence d'acquisition	Taille des images
V1	7 images/s	720 × 576
V4	25 images/s	352 × 288
V5	9 images/s	352 × 288
V7	5 images/s	352 × 288

TAB. 2 – Performances du détecteur opérant avec $k_1 = 30$ sur L images codées par MJPEG (a) et MPEG2 (b).

Séquence	k_0	L	P_f	P_m	P_e	Δ
V1	1	30	0.0129	0.09076	0.1036	0
V4	1	30	0.2012	0.0061	0.2073	0
V5	1	30	0.0159	0.0425	0.0584	0
V7	1	30	0.1234	0.1511	0.2745	0
V1	1	350	0.4242	0.0442	0.4684	1
V4	1	1000	0.1654	0.0145	0.1799	0
V5	1	950	0.0261	0.1530	0.1791	0
V7	1	2470	0.0969	0.1790	0.2759	0

(a)

Séquence	k_0	L	P_f	P_m	P_e	Δ
V1	7	30	0.0037	0.1209	0.1246	1
V4	25	30	0.1759	0.0169	0.1928	0
V5	9	30	0.0251	0.0522	0.0773	0
V7	5	30	0.0647	0.1548	0.2195	0
V1	7	50	0.018	0.1549	0.1729	1
V4	25	40	0.1745	0.0141	0.1886	0
V5	9	106	0.0279	0.1503	0.1782	0
V7	5	494	0.0946	0.1805	0.2751	0

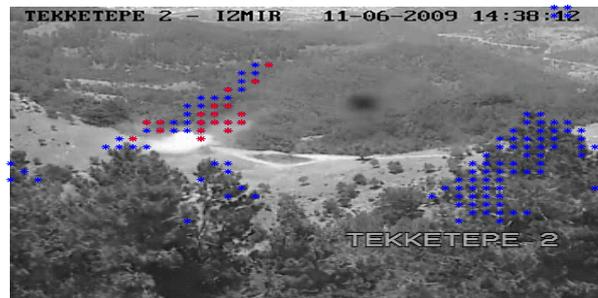
(b)

Références

- [1] H. N. Le Houérou, "Vegetation wildfires in the mediterranean basin : evolution and trends," *Ecologia Mediterranea*, vol. 13, pp. 13-24, 1987.
- [2] W. Jones, "An algorithm for fast and reliable fire detection," *8th Fire Suppression and Detection Research Application Symp.*, January 2004, Orlando.
- [3] W. Phillips III, M. Shah, and N. V. Lobo, "Flame recognition in video," *Pattern Recognition Letters*, vol. 23, no. 1-3, pp. 319-327, 2002.
- [4] T. H. Chen, P. H. Wu and Y. C. Chiou, "An early fire-detection method based on image processing," *IEEE Internat. Conf. on Image Processing*, Singapore, pp. 1707-1710, Oct 2004.
- [5] N. Grammaladis *et al.*, "A multisensor network for the protection of culturel heritage," *European Signal and Image Processing Conf.* Barcelona, Spain, 2011.



(a) Séquence V1



(b) Séquence V5

FIG. 2 – Images codées en MJPEG : blocs classés en $c = 0$ (en bleu) en $c = 1$ (en rouge) avec $(k_0, k_1) = (1, 30)$.

- [6] Z. Xiong, R. Caballero, H. Wang, A. Finn, M. A. Lelic and P. Peng, "Vidoe-based smoke detection possibilities, techniques, and challenges," *Suppression and Detection Technical Working Conf.*, Orlando, FL, USA, March 2007.
- [7] T. H. Chen, Y. H. Yin, S. F. Huang and Y. T. Ye, "The smoke detection for early fire-arming system base on video processing," *IEEE Internat. Conf. on Image Processing*, Singapore, pp. 1707-1710, Oct 2004.
- [8] B. U. Toreyin, Y. Dedeoglu, U. Gudukbay, A. E. Çetin, "Computer vision based method for real time fire and flame detection," *Pattern Recognition Letter*, no. 27, pp. 49-58, 2006.
- [9] S. Verstockt, P. Lambert, R. Van de Walle, B. Merci and B. Sette, "State of the art in vision-based fire and smoke detection," *Internat. Conf. on Automatic Fire Detection*, 2009)
- [10] T. X. Truong, J.-M. Kim, "An early smoke detection system based on motion estimation," *Internat. Forum on Strategic Technology*, pp. 437-440, Ulsan, Korea, Oct. 2010.
- [11] B. U. Toreyin, Y. Dedeoglu and A. E. A. E. Çetin, "Contour based smoke detection in video using wavelet," *European Signal and Image Processing Conf.*, Firenze, Italy, Sept. 2006.
- [12] T. Celik, H. Ozkaramanli and H. Demirel, "Fire and smoke detection witout sensors : image processing based approach," *European Signal and Image Processing Conf.*, Poznan, Poland, Sept. 2007.
- [13] A. Rafiee, R. Tavakoli, R. Dianat, S. Abbaspour and M. Jamshidi, "Fire and smoke detection using wavelet analysis and disorder characteristics," *Internat. Conf. on Computer Research and Development*, Shanghai, China, pp. 262-265, March 2011.
- [14] R. Gonzalez-Gonzalez, V. Alarcon-Aquino, R. Rosas-Romero, O. Starostenko, J. Rodriguez-Asomoza and J. M. Ramirez-Cortes, "Wavelet-based smoke detection in outdoor video sequences," *IEEE Midwest Symp. on Circuits and Systems*, Seattle, Washington, USA, pp. 383-387, August 2010.
- [15] H. Maruta, A. Nakamura and F. Kurokawa, "A novel smoke detection method using support vector machine," *IEEE TENCON*, pp. 210-215, Nov. 2010.
- [16] H. Maruta, A. Nakamura, T. Yamamichi and F. Kurokawa, "Image based smoke detection with local Hurst exponent," *IEEE TENCON*, pp.4653-4656, Nov.2010.
- [17] I. Richardson, *H.264 and MPEG-4 video compression*, Wiley, Chichester, West Sussex, England, September 2003.
- [18] F. Luthon, M. Liévin, F. Faux, "On the use of entropy power for threshold selection," *Signal Processing*, vol. 84, pp.1789-1804, 2004.
- [19] B. Mandelbrot, "Self-affine fractals and fractal dimension," *Physica Scripta*, vol. 32, pp. 257-260, 1985.
- [20] M. Petrou, P. Garcia-Sevilla. *Dealing with texture*, Wiley, Chichester, West Sussex, England, 2006.