A NOVEL VIDEO-BASED SMOKE DETECTION METHOD BASED ON COLOR INVARIANTS

O. Besbes

COSIM Lab., ISITCOM, Sousse University,
G.P.1 Hammam Sousse, 4011, Tunisia.

A. Benazza-Benyahia

COSIM Lab., SUP'COM, Carthage University,
Technopôle Ghazala, 2080, Tunisia.

ABSTRACT

In this paper, we address the issue of designing a smoke detector robust to illumination variations. Our contribution consists in resorting to color invariants as salient smoke features. More precisely, the proposed detector employs consecutively of an illumination invariant color representation, a photometric gain based background subtraction, a chrominance detection and a smoke identification based on two invariant color descriptors. The experimental results show that the proposed method can effectively detect smoke with robustness to illumination changes and noises, frequently encountered in wildfire video-surveillance environments.

Index Terms— Wildfire smoke detection, color invariants, photometric gain, chrominance detection, invariant color descriptors.

1. INTRODUCTION

In the last decade years, smoke detection in video surveillance has attracted research attention as smoke plays a crucial role in early fire detection. Indeed, smoke appears usually at early stage of fire and can be observed from a great distance. A lot of Video Smoke Detection (VSD) methods have been proposed, and a comprehensive review of them is presented in [1]. Video-based smoke detection remains nevertheless a challenging problem because of two main reasons. Mainly, smoke is hard to model due to its highly variable visual appearance. Furthermore, lighting conditions, environmental changes, dynamic background, and unstable cameras affect substantially the performance of smoke detection. That is why the research on video-based smoke detection focuses mostly on the reduction of false alarms, while saving the computation time.

Most existing methods exploit various visual characteristics of smoke such as color, texture, shape and motion features. For instance, in [2], candidate smoke regions were extracted by combining an energy analysis from wavelet coefficients, a background subtraction and, a chrominance detection. In [3], a statistical analysis is carried out using the idea that the smoke shows grayish color with different illumination. In [4], candidate smoke regions were extracted following three consecutive steps, namely a block-based frame difference, a rule-based chrominance detection, and an accumulative motion model. Later, in [5] smoke was detected based on the analysis of color and texture features of moving regions, extracted by means of a background subtraction procedure. The temporal behavior of smoke was captured by a mixture of Gaussians modeling the wavelet energy in order to classify smoke patterns. More recently, in [6], a back-propagation neural network was used to detect smoke patterns, which were described by a concatenation of the histograms of local binary pattern and local binary pattern variance pyramids. Differently, in [7] a support vector machine (SVM) classifier was applied on spatio-temporal correlation descriptors to identify smoke. Furthermore, Wang *et al.* in [8] combined the color information with a modified center symmetric local ternary pattern to differentiate between smoke and non-smoke regions. More lately, in [9] an wavelet transformation was implemented based on the RGB contrast-image to distinguish smoke from other low frequency signals and, the existence of smoke was determined by analyzing the shape and energy variations. Park *et al.* in [10] presented a wildfire smoke detection method based on a spatio-temporal bag-of-features and a random forest classifier. More recently, histograms of oriented gradients (HOG) and histograms of optical flow (HOF) were computed in [11] to take into account both appearance and motion information. Initially, candidate smoke regions were identified using background subtraction and HSV-color analysis. Subsequently, three SVM classifiers were combined to identify smoke regions based on respectively spatio-temporal energy analysis, bag of HOG-HOF features, and dynamic texture analysis.

In this paper, we focus on color cue for distinguishing smoke regions because it provides powerful information as shown in the aforementioned methods. In that case, the main challenge consists in preserving robustness with respect to photometric variations as are common in outdoor scenes. These variations can affect either locally a scene such as shadows, shadings and specularities, or globally as light intensity changes and light color changes. Consequently, the performance of smoke detection can be greatly affected if the descriptors used are not robust to such changes. To the best of our knowledge, this issue has not yet been addressed in the context of smoke detection. In this paper, we propose a novel VSD approach which exploits some photometric invariants to achieve robustness against both local and global illumination changes. This paper is organized as follows. First, the proposed VSD method is presented in Sec. 2. Then, experimental results are provided in Sec. 3. Finally, conclusions are drawn in Sec. 4.

2. SMOKE DETECTION METHODOLOGY

Our novel color invariant-based smoke detection method relies on the following steps. First, the RGB background image is estimated. Then, both current frame and background image are converted towards an illumination invariant color representation. After that, an effective block-based background/foreground segmentation under challenging conditions is performed by means of photometric gains. Afterwards, non-smoke blocks are filtered out by exploiting the photometric properties of smoke. Finally, two invariant color descriptors are exploited to identify smoke regions.

2.1. Adaptive background estimation

We adopt a robust adaptive background model to deal with the local variations of illumination over time. Let F_n denote the RGB-valued frame of the sequence at time n and, F_n^c be its channel $c \in \{R, G, B\}$. The background image B_0 is firstly initialized by the first frame F_0. Then, it is updated over time, at each pixel (x,y), according to a linear combination of previous background and frame [12]:

$$B_{n+1}(x,y)=\begin{cases} \alpha B_n(x,y)+(1-\alpha)F_n(x,y) & \text{if } M_n(x,y)=0,\\ B_n(x,y) & \text{otherwise,}\end{cases} \tag{1}$$

where $\alpha \in [0,1]$ is a time constant parameter reflecting the sensitivity of the update to the variations and, M_n is a binary mask such as $M_n(x,y)=1$ for pixels with changes in illumination or generated by moving objects. A naive way to derive M_n is the thresholding of the difference between two consecutive frames. However, the high level of noise in outdoor sequences induces unacceptable inaccuracies. Therefore, we calculate the block difference as the sums of differences of pixels within a block:

$$b_n(i,j)=\begin{cases} 1 & \text{if } \sum\limits_{c\in\{R,G,B\}}\left|\sum\limits_{(x,y)\in b_{ij}} F_n^c(x,y)-F_{n-1}^c(x,y)\right|>T_b,\\ 0 & \text{otherwise,}\end{cases} \tag{2}$$

where T_b is a predetermined threshold, and block b_{ij} is on the i^{th} row and j^{th} column in the video. Then, we derive easily the binary mask $M_n(x,y)=b_n(i,j)$, $\forall\,(x,y)\in b_{ij}$. Thereby, pixels whose color values have changed between the two consecutive frames are detected. Note that M_n is not suitable to segment moving regions, but it is just requisite to estimate the background model. Concerning the learning rate α, it depends on the percentage of moving pixels and, it can be set as $\alpha=1-\frac{\sum_{(x,y)}M_n(x,y)}{N}$, where N is the area of M_n.

2.2. Illumination invariant color representation

Color is an effective cue for discriminating smoke regions [3, 4]. However, color information in recorded RGB videos depends on the illumination condition under which the scene is viewed. In particular, large variations in lighting conditions are constantly encountered during a day in outdoor environments. These changes can be modeled by the well-known diagonal-offset model [13] as follows:

$$\begin{bmatrix} R^r\\ G^r\\ B^r\end{bmatrix}=\begin{bmatrix}\beta_1 & 0 & 0\\ 0 & \beta_2 & 0\\ 0 & 0 & \beta_3\end{bmatrix}\begin{bmatrix} R^u\\ G^u\\ B^u\end{bmatrix}+\begin{bmatrix} o_1\\ o_2\\ o_3\end{bmatrix} \tag{3}$$

where $[R^u,G^u,B^u]$ denotes an image taken under an unknown light source which is mapped to a transformed image $[R^r,G^r,B^r]$ taken under the reference light, called canonical illuminant. The mapping is parameterized by the light color changes $\{\beta_1,\beta_2,\beta_3\}$ and, the shift parameters $\{o_1,o_2,o_3\}$. Motivated by the studies in color constancy [14], we propose to convert both the frame F_n and its corresponding background B_n from the RGB color space towards an invariant color space to discard the influence of *global* illumination changes. More precisely, the RGB images are normalized independently;

$$\begin{bmatrix} R'\\ G'\\ B'\end{bmatrix}=\begin{bmatrix}\frac{R-\mu_R}{\sigma_R}\\ \frac{G-\mu_G}{\sigma_G}\\ \frac{B-\mu_B}{\sigma_B}\end{bmatrix}, \tag{4}$$

with the mean μ_c and the standard deviation σ_c of the distribution in channel c computed over the whole image. This yields, for every channel, a distribution with zero mean and unit variance. Thanks to this normalization, the *transformed* RGB color space ensures invariance to light color changes and shifts [15]. Indeed, we can check it through the diagonal-offset model. For the sake of clarity, let us consider for instance the red channel:

$$\mu_R=\frac{1}{N}\sum_{(x,y)}(\beta_1R^u(x,y)+o_1)=\beta_1\mu_R^u+o_1, \tag{5}$$

$$\sigma_R=\sqrt{\frac{1}{N}\sum_{(x,y)}(\beta_1R^u(x,y)+o_1-\beta_1\mu_R^u-o_1)^2}=\beta_1\sigma_R^u, \tag{6}$$

$$\frac{R-\mu_R}{\sigma_R}=\frac{\beta_1R^u+o_1-\beta_1\mu_R^u-o_1}{\beta_1\sigma_R^u}=\frac{R^u-\mu_R^u}{\sigma_R^u}. \tag{7}$$

Note that the light color change $\{\beta_1,\beta_2,\beta_3\}$ and shift $\{o_1,o_2,o_3\}$ parameters are canceled out. In the following, B_n and F_n denote respectively the *transformed* background and the *transformed* frame at time n into the illumination invariant color representation. These images are also scaled to $[0,255]$. It is important to point out that the frame is normalized with the means and the standard deviations computed over the corresponding background image to discount the influence of its additional moving objects. Fig. 1 illustrates the effect of this normalization on images acquired under different illumination conditions.

Fig. 1. The illumination invariant color representations (b) and (d) of respectively the RGB images (a) and (c).

2.3. Moving region detection

Once the background model B_n has been estimated, the moving regions are extracted from the corresponding frame F_n using the background subtraction technique. Let $D_n=|F_n-B_n|$ be the absolute difference of the background image and current frame in the illumination invariant color space. Rather than thresholding the difference D_n, we propose to rely on the photometric gain in order to detect moving regions in different color similarity situations and under *local* illumination changes. For each channel $c\in\{R',G',B'\}$, an adapted version of the photometric gain [12] is computed as follows:

$$\Lambda_n^c(x,y)=\begin{cases} 1-\frac{\min\left[F_n^c(x,y),B_n^c(x,y)\right]}{\max\left[F_n^c(x,y),B_n^c(x,y)\right]\times D_n^c(x,y)} & \text{if } D_n^c(x,y)\neq 0,\\ 0 & \text{otherwise.}\end{cases} \tag{8}$$

On the one hand, we note that $\Lambda_n^c(x,y)$ takes values close to 1 each time the moving pixel is different from the corresponding background pixel and, coherently with a high probability that the pixel could be marked as foreground regarding channel c. On the other hand, $\Lambda_n^c(x,y)$ is close to 0 if the pixel belongs to the background. Thereby, moving regions can be identified thanks to the resulting overall gain, defined as $\Lambda_n(x,y)=\prod\limits_{c\in\{R',G',B'\}}\Lambda_n^c(x,y)$. Moreover, we perform a blockwise decision so as to achieve further robustness against noise and local illumination changes:

$$r_n(i,j)=\begin{cases} 1 & \text{if } \frac{1}{|b_{ij}|}\sum\limits_{(x,y)\in b_{ij}}\Lambda_n(x,y)>T_r,\\ 0 & \text{otherwise.}\end{cases} \tag{9}$$

A moving block has an average gain greater than a predetermined threshold T_r (set to 0.5). Finally, we derive the moving detection binary mask $R_n(x,y) = r_n(i,j)\ ,\ \forall\,(x,y) \in b_{ij}$. Thus, moving pixels with high photometric gains are reliably detected. Fig. 2 illustrates the effectiveness of photometric gains in background/foreground segmentation. Moving regions are segmented out accurately even in color similarity situations. Indeed, our background subtraction method is able to detect moving smoke regions in front of white background. In addition, false positive detected pixels, mostly due to noise and local illumination changes, are considerably reduced. This demonstrates clearly the robustness of our strategy face to noise and local illumination changes.

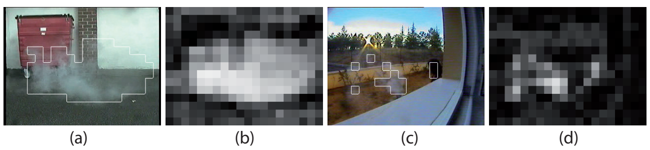

Fig. 2. (a) and (c) the detected moving regions by using respectively the photometric gains (b) and (d).

2.4. Chrominance detection

In order to filter out non-smoke colored moving pixels, we make use of the following photometric properties of smoke [1]:

$$C = \max(R', G', B') - \min(R', G', B')\,, \tag{10}$$

$$I = \frac{R' + G' + B'}{3}, \tag{11}$$

$$S = \sqrt{\frac{1}{2}(R' - G')^2 + \frac{1}{6}(R' + G' - 2B')^2}, \tag{12}$$

$$\textbf{Rule} \ :\ (C < T_1) \text{ and } (T_2 < I < T_3) \text{ and } (S < T_4)\,, \tag{13}$$

where the values of the thresholds were experimentally determined using a number of training images[1]. Smoke's color is black, grayish or white, we can then detect smoke colored pixels by thresholding the chroma C, the intensity I and the saturation S. For a smoke pixel, R', G' and B' values are very close to each other, i.e. $(C < T_1)$, its intensity ranges from T_2 to T_3, and its saturation is lower than T_4. These thresholds are set respectively to 35, 80, 190 and 25, as justified by Fig. 3. According to the rules defined in (13), the chrominance detection is performed to validate each moving block. If at least 10% of the moving block's pixels are smoke colored then the block is considered as a candidate smoke region.

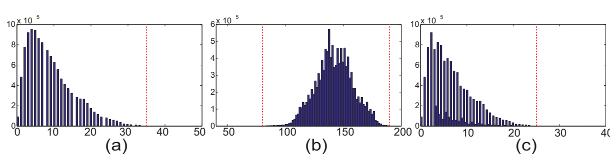

Fig. 3. The photometric properties of smoke according to (a) chroma C, (b) intensity I and (c) saturation S.

[1]The database (http://wildfire.fesb.hr/) contains a selection of wildfire smoke images, manually segmented in 3-classes defined as: smoke, maybe smoke and non-smoke.

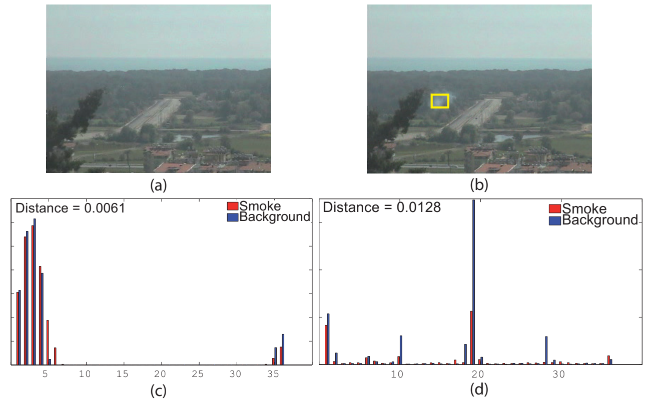

Fig. 4. Comparison of invariant color descriptors of regions with and without smoke: (a) background image, (b) frame with smoke, (c) the robust hue descriptors, and (c) the hue-based oriented gradient histograms.

2.5. Invariant color descriptors

Regions under semi-transparent smoke change their intensity (i.e they become lighter or darker) but keep preserved their chromacity. For instance, a green pixel covered by an white smoke becomes light-green, which is lighter than green but has the same chromacity. In order to exploit this color property, we chose the HSI color space since it provides a natural separation between chromacity and intensity. Both intensity I and saturation S, defined respectively in (11) and (12), are exploited to detect candidate smoke regions according to the rule (13), as explained in the subsection 2.4. As regards the *hue* cue, we notice that a smoke on background does not change its hue H defined as:

$$H = \arctan\left(\frac{\sqrt{3}\,(R' - G')}{(R' + G' - 2B')}\right). \tag{14}$$

In addition, smoke often lowers the saturation S of the background. A color descriptor that combines hue and saturation cues is thus able to discriminate powerfully the smoke regions. The desired color descriptor should be computationally efficient and, offers a certain amount of photometric invariance while maintaining high discrimination power. For these reasons, we make use of the *robust hue descriptor* [16]. This color descriptor is a 36-dimensional weighted histogram $\mathcal{H}^h$ where the bin index is determined by hue and the weight by saturation. The hue descriptor is invariant to local illumination changes, namely shadows, shadings and specularities.

Furthermore, regions under smoke retain most of their texture. Herein, we exploit the *hue-based oriented gradient histogram* since a smoke on background does not change its gradient orientation, while its gradient magnitude becomes lower. This descriptor [17] is likewise a weighted histogram $\mathcal{H}^g$ where a gradient orientation bin index is weighted by gradient magnitude. It is worth to note that gradients are calculated from the hue in order to obtain invariance to shadow, shading and specular edges [18].

For each candidate block, we compute the invariant color descriptors for both background image and current frame. A candidate block is classified as smoke if for both descriptors the Chi-squared distance (15) between frame and background histograms is lesser than predetermined thresholds:

$$\chi^2(\mathcal{H}_1, \mathcal{H}_2) = \frac{1}{2} \sum_k \frac{(\mathcal{H}_1[k] - \mathcal{H}_2[k])^2}{\mathcal{H}_1[k] + \mathcal{H}_2[k]}, \quad (15)$$

$$\textbf{Rule} : \left(\chi^2\left(\mathcal{H}_F^h, \mathcal{H}_B^h\right) < T_h\right) \text{ and } \left(\chi^2(\mathcal{H}_F^g, \mathcal{H}_B^g) < T_g\right). \quad (16)$$

Fig. 4 shows a comparison on these descriptors with and without smoke, it is clear that smoke smooths them and thus their bins become lower. Hence, we assess the similarity between a smoke region and the background reference in chromaticity and texture by means of respectively the robust hue descriptor and the hue-based oriented gradient histogram descriptor. These descriptors, as shown in the next section, allow a good trade-off between photometric invariance and discriminating power in the context of video-based smoke detection.

Table 1. Smoke Detection performance in Bilkent videos.

Smoke videos	**TNR**	**TPR**	**TPR of [11]**
sBehindtheFence	100	94.72	94.44
sBtFence2	100	99.08	98.71
sEmptyR1	100	98.08	73.08
sEmptyR2	100	89.55	88.60
sMoky	100	86.23	99.78
sWasteBasket	92.60	99.89	99.29
sWindow	100	94.30	88.52
Total average	**98.94**	**94.55**	**91.77**

3. EXPERIMENTAL RESULTS

We have tested our method on real smoke videos from Bilkent University[2]. These videos include outdoor and indoor environments under various illumination conditions. The used block's size is 16×16. Fig. 5 demonstrates the performance of the proposed method for the different video sequences. We evaluate the performance of our method, as shown in Table 1, by computing true positive rate TPR and true negative rate TNR which are, respectively, defined by:

$$\text{TPR} = \frac{\text{Number of TP frames}}{\text{Number of TP frames} + \text{Number of FN frames}}, \quad (17)$$

$$\text{TNR} = \frac{\text{Number of TN frames}}{\text{Number of TN frames} + \text{Number of FP frames}}. \quad (18)$$

These quantitative results are compared with those obtained by the recent smoke detection method [11]. The proposed method outperforms the method in [11], with a TPR's average of 94.55%. Besides, we obtain an a TPR's average of 98.94%. This shows the robustness of the proposed method since false alarms are considerably reduced. It is important to note that these promising results are achieved by exploiting the color cue and employing the photometric invariance. This indicates that color is a powerful cue for discriminating smoke, especially when it is coupled with the aspect of photometric invariance.

4. CONCLUSION

In this paper, a novel smoke detection method based on color invariants is proposed. It relies on an illumination invariant color representation, photometric gains to segment out moving regions, a chrominance detection to filter out non-smoke colored regions, and two invariant color descriptors to identify reliably smoke regions. Experimental results, carried out under various challenging conditions, demonstrate the robustness of our method to illumination changes very often encountered in wildfire video-surveillance environments, as well as to the noise corrupting constantly acquired videos. In future works, we will incorporate additional discriminant informations such as texture, shape and motion features to improve further the smoke detection.

[2]http://signal.ee.bilkent.edu.tr/VisiFire/Demo/SampleClips.html

Fig. 5. Experimental results of the proposed method.

5. REFERENCES

[1] A. E. Çetin, K. Dimitropoulos, B. Gouverneur, N. Grammalidis, O. Günay, Y. H. Habiboglu, B. U. Töreyin, and S. Verstockt, "Video fire detection - review," *Digit. Signal Process.*, vol. 23, no. 6, pp. 1827–1843, 2013.

[2] B. U. Toreyin, Y. Dedeoglu, and A. E. Cetin, "Contour based smoke detection in video using wavelets," in *EUSIPCO*, Florence, Italy, September 2006, pp. 1461–1464.

[3] T. Celik, H. Ozkaramanli, and H. Demirel, "Fire and smoke de-

tection without sensors: Image processing based approach," in *EUSIPCO*, Poznan, Poland, September 2007, pp. 1794–1798.

[4] F. Yuan, "A fast accumulative motion orientation model based on integral image for video smoke detection," *Pattern Recognition Letter*, vol. 29, no. 7, pp. 925–932, May 2008.

[5] S. Calderara, P. Piccinini, and R. Cucchiara, "Smoke detection in video surveillance: A mog model in the wavelet domain," in *ICVS*, Santorini, Greece, May 2008, pp. 119–128.

[6] F. Yuan, "Video-based smoke detection with histogram sequence of lbp and lbpv pyramids," *Fire Safety Journal*, vol. 46, no. 3, pp. 132–139, April 2011.

[7] Y. H. Habiboglu, O. Günay, and A. E. Çetin, "Covariance matrix-based fire and flame detection method in video," *MVA*, vol. 23, no. 6, pp. 1103–1113, 2012.

[8] Y. Wang, T. W. Chua, R. Chang, and N. T. Pham, "Real-time smoke detection using texture and color features," in *IAPR ICPR*, Tsukuba, Japan, November 2012, pp. 1727–1730.

[9] J. Q. Chen, Y. W. Wang, Y. H. Tian, and T. J. Huang, "Wavelet based smoke detection method with rgb contrast-image and shape constrain," in *IEEE VCIP*, Kuching, Malaysia, November 2013, pp. 1–6.

[10] J. Park, B. Ko, J. Y. Nam, and S. Kwak, "Wildfire smoke detection using spatiotemporal bag-of-features of smoke," in *IEE WACV*, Clearwater Beach, Fl, USA, January 2013, pp. 200–205.

[11] P. Barmpoutis, K. Dimitropoulos, and N. Grammalidis, "Smoke detection using spatio-temporal analysis, motion modeling and dynamic texture recognition," in *EUSIPCO*, Lisbon, Portugal, September 2014, pp. 1078–1082.

[12] A. Leone and C. Distante, "Shadow detection for moving objects based on texture analysis," *Pattern Recognition*, vol. 40, no. 4, pp. 1222–1233, April 2007.

[13] G. D. Finlayson, S. D. Hordley, and R. Xu, "Convex programming colour constancy with a diagonal-offset model," in *IEEE ICIP*, Genova, Italy, September 2005, vol. 3, pp. III–948–51.

[14] D. A. Forsyth, "A novel algorithm for color constancy," *IJCV*, vol. 5, no. 1, pp. 5–36, September 1990.

[15] K. E. A. van de Sande, T. Gevers, and C. G. M. Snoek, "Evaluating color descriptors for object and scene recognition," *IEEE Trans. PAMI*, vol. 32, no. 9, pp. 1582–1596, September 2010.

[16] J. V. Weijer and C. Schmid, "Coloring local feature extraction," in *ECCV*, Graz, Austria, May 2006, pp. 334–348.

[17] P. Dollar, Z. Tu, P. Perona, and S. Belongie, "Integral channel features," in *BMVC*, London, UK, September 2009, pp. 1–11.

[18] J. V. Weijer, T. Gevers, and J. M. Geusebroek, "Edge and corner detection by photometric quasi-invariants," *IEEE Trans. PAMI*, vol. 27, no. 4, pp. 625–630, 2005.