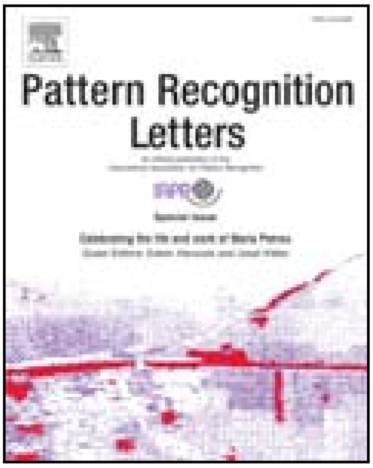
Journal Pre-proof



Moving Object Detection Under Different Weather Conditions Using Full-Spectrum Light Sources

Rania Rebai Boukhriss, Emna Fendri, Mohamed Hammami

PII: S0167-8655(19)30318-6
DOI: https://doi.org/10.1016/j.patrec.2019.11.004
Reference: PATREC 7688

To appear in: *Pattern Recognition Letters*

Received date: 20 February 2019
Revised date: 29 October 2019
Accepted date: 6 November 2019

Please cite this article as: Rania Rebai Boukhriss, Emna Fendri, Mohamed Hammami, Moving Object Detection Under Different Weather Conditions Using Full-Spectrum Light Sources, *Pattern Recognition Letters* (2019), doi: https://doi.org/10.1016/j.patrec.2019.11.004

Highlights

- Moving object detection based on background modeling incorporating the principle of inter-frame difference
- A new switching strategy for moving object detection using full spectrum lights sources (Infrared and Visible spectra)
- Moving object detection in low illumination and bad weather conditions (i.e. fog, snow, rain, darkness, etc.)

Pattern Recognition Letters
journal homepage: www.elsevier.com

Moving Object Detection Under Different Weather Conditions Using Full-Spectrum Light Sources

Rania Rebai Boukhriss[a,**], Emna Fendri[b], Mohamed Hammami[b]

[a]MIRACL-ISIMS, Sfax University, Sakiet Ezzit, Sfax, Tunisia
[b]MIRACL-FS, Sfax University, Road Sokra, Sfax, Tunisia

ABSTRACT

The moving object detection always remains an active field of research given the variety of challenges related to this topic. In fact, most of the challenges related to the low illumination and weather conditions (fog, snow, rain, etc.) remain unresolved and require more developments. In this paper, our intrinsic objective is to overcome these challenges using an effective moving object detection method. Unlike most works in the literature that use one of the two infrared or visible spectra independently, we proposed a Moving Object Detection method based on background modeling using the Full-Spectrum Light Sources (*FSLS-MOD*). To better ensure the adaptability and independence of the moving object speeds and sizes, the principle of the inter-frame differences' methods is introduced in the background modeling stage. Furthermore, we applied a new strategy to switch between the spectra allowing us to benefit from the advantages of each spectrum and carry out a better moving object detection even in bad weather conditions. An experimental study by quantitative and qualitative evaluations proved the robustness and effectiveness of our proposed method of moving object detection using the switching strategy between full-spectrum light sources under different illuminations and weather conditions.
Keywords: Full-spectrum Light Sources; Weather Conditions Classification; Moving Object Detection; Thermal Infrared Camera

1. Introduction

The ability to extract information regarding moving objects is a crucial factor of intelligent video surveillance systems performance. Although the moving object detection was very developed in the literature, it still remains an active area of research and the object of many studies. According to previous research, existing approaches for motion detection in video surveillance systems can be classified in four categories: inter-frame differences Cheng and Wang (2014); Haritaoglu et al. (2000), background modeling Chen and Liu (2014); Bouwmans (2014), optical flow Ribeiro et al. (2018); Qi and An (2011) and hybrid methods Kushwaha et al. (2017); Yin et al. (2016). The diversity of research is related to the complexity of the observed scenes which present a variety of challenges. These challenges are mainly related to the problems of background, moving objects or acquisition materials. In the literature, many works have tackled these challenges and several hand-crafted solutions have been proposed to overcome most of them Lim and Keles (2018); Qin et al. (2016). In addition, recent deep learning based methods, especially the convolutional neural networks (*CNN*), were proposed for robust targets detection Lin et al. (2018a,b); Wang et al. (2018). However, the challenges related to the bad weather conditions (such as fog, snow, rain, darkness, etc.) remain unresolved and require more substantial research to find efficient solutions. In video surveillance, we can rely on either a visible spectrum or an infrared one. However, the methods based only on the visible spectrum suffers from such limitations as failure to face shadows, camouflage, night or poor visibility conditions caused by weather conditions such as fog, rain and snow. Furthermore, an infrared-based system may find it difficult to handle some information in certain situations. For example, during a hot sunny day, it will highlight almost the entire image, so it will provide a lot of hot areas or objects. Thus, the use of a visible sensor with an infrared

**Corresponding author:
e-mail: rania.rebai@hotmail.fr (Rania Rebai Boukhriss)

sensor makes the vision systems more robust and enables them to function under various lighting and climatic conditions, during both day and night, in summer as well as in winter Conaire et al. (2005); Yan et al. (2018). Since our basic objective is to achieve an accurate moving object detection during the day (morning, afternoon and night) for particular hot objects such as people and vehicles, we opted for use the full-spectrum light sources. In this paper, we proposed a moving object detection method with a new switching strategy between the spectra light sources allowing us to take advantage of each of them, without the need to correlate the spectra nor to generate a fused spectrum. Our switching strategy ensures an accurate moving object detection despite the challenges related to the low illumination and bad weather conditions. In fact, the visible spectrum is used on sunny days under normal weather conditions, while the infrared spectrum will be used in bad weather conditions such as fog, rain, snow, etc. or at night. In addition, for the moving object detection we have adapted a method developped previously in our research team Hammami et al. (2013). In fact, our moving object detection method is based on background modeling and finds its originality in the integration of the principle of the methods based on inter-frame differences in the stage of the background modeling and in the way in which this integration was carried out, making it more adaptable and independent of the moving object speeds and sizes. The remainder of this paper is organized as follows: Section 2 introduces our proposed methods of switching between the full-spectrum light sources and the moving object detection. The experimental results of our work are displayed in Section 3. Finally, our conclusion and future work are stated in Section 4.

2. Proposed method for moving object detection using full-spectrum light sources (*FSLS-MOD*)

In this paper, we suggested a moving object detection method using the full-spectrum light sources, which follows an appropriate temporal behavior. In fact, we introduced a new switching strategy between the spectra light sources allowing us to profit from the advantages of each spectrum. The process of our proposed method consists of two main steps: The switching between full-spectrum lighting sources and the moving object detection as shown in Fig.1. The first step is based on weather condition classification of visible images in one of two states: *Normal State* that is sunny days under normal weather conditions or *Abnormal State*, during bad weather conditions such as rain, snow or fog. In fact, the visible spectrum is used in a *Normal State*, while the infrared spectrum is applied in an *Abnormal State* or at night. Relying on this classification result, our switching method decides which spectrum is to be used in the next step. Therefore, the moving object detection method will be applied in the InfraRed (IR) or in the VISible (VIS) spectra. As previously mentioned, the proposed method adopts a background modeling method incorporating the principle of inter-frame difference in the background modeling stage.

2.1. *Full-Spectrum Light Sources Switching*

As our main objective is to perform correct moving object detection during the whole day whatever the weather condition is,

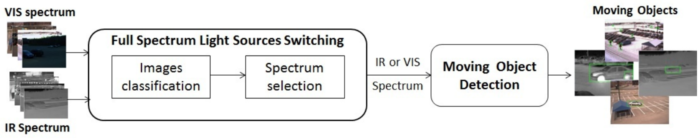


Fig. 1. Process of our *FSLS-MOD* method.

we proposed a new strategy to switch between the spectra light sources which would allow us to benefit from the advantages of each of them. This strategy is based on *Weather Conditions Classification Prediction Model* (*WCC-PM*) which is used to classify the images of the VIS spectrum in *Normal State* (*N.S.*) or *Abnormal State* (*A.S.*). The *N.S.* indicates that the VIS images were taken on a sunny day at normal weather conditions. As for *A.S.*, it denotes that the VIS images were taken at night or during bad weather condition. The *WCC-PM* is generated by following a process of knowledge extraction from the data through a supervised learning technique based on decision trees and using a large and representative training data set in our previous works Boukhriss et al. (2015). The decision result of our *WCC-PM* allows us to decide which spectrum to use for the moving object detection. In fact, if the VIS image is in *N.S.*, we use the VIS spectrum; otherwise, we use the IR spectrum.

2.2. *Moving Object Detection*

Our study of the state of the art highlights that the background modeling-based methods are well suited to our objective. Moreover, given the performance of the inter-frame differences based methods, our method relies on the integration of their principle in the modeling of the background. The proposed method for moving object detection rests on 4 main stages which are depicted in the following sub-sections.

2.2.1. *Background Model Initialization*

In the initialization step, we used three frames (F^t, F^{t-1}, F^{t-2}) to build the initial background model (B^i). In fact, we obtain the initial model resting essentially on the frame at a specific moment during which the non-moving pixels values are kept in the model whereas the values of the background pixels hidden by the moving pixels of this frame are approached relying on the variations among the pixel values of the other two frames (F^{t-1},F^{t-2}), using equation 1.

$$B^i = \left|F^t - \left|F^{t-1} - F^{t-2}\right|\right| \tag{1}$$

2.2.2. *Background Model Update*

In this step, we proposed a selective technique to update the background model to the changes in background pixels. In fact, only the pixels classified as background pixels with significant change, are added to the model. In our method, an analysis of the spatio-temporal entropies of the pixels is applied to select these pixels in an accurate way. In addition, we proposed a dynamic matrix to ensure this selection. In our method, the selective updating of the background model is carried out in three steps, namely: (*i*) pixel state card construction; (*ii*) dynamic matrix update based on the pixel state card; and (*iii*) background pixels update.

***(i)* Pixel state card construction.** The pixel state card allows us to restrict the updating of the dynamic matrix to the background pixels, which reduces the calculation time. This card is obtained by inter-frame differences based on a spatio-temporal analysis of the entropy. In fact, the entropy of each pixel is calculated from a spatio-temporal analysis of its neighborhoods. Classically, the entropy of each pixel is calculated from a spatio-temporal analysis of its neighboring pixels based on the distribution of gray levels in a window. Such a calculation method does not make it possible to distinguish between the entropy of a moving pixel and that of a noise (*i.e.* entropy of a pixel contour of an object of the background). We, therefore, opted for the calculation of the entropy of each pixel based on a labeling technique Chang and Cheng (2007).

***(ii)* Dynamic matrix update.** The dynamic matrix is used to represent the short-term history of the background pixel states. First, the dynamic matrix $\varepsilon^t_{x,y}$ is initialized by equation 2, using the mask of moving pixels $M^t_{x,y}$. Then, the update of the dynamic matrix is performed by highlighting the background pixels at time t and that are moving at $t-1$ (*cf.* Equation 3).

$$\begin{cases} \varepsilon^t_{x,y} = 0 & : \quad M^t_{x,y} = 1 \\ \varepsilon^t_{x,y} = 1 & : \quad M^t_{x,y} = 0 \end{cases} \tag{2}$$

$$\begin{cases} if \quad \varepsilon^t_{x,y} = 1 \ and \ M^{t-1}_{x,y} = 1 \ then & M^t_{x,y} = 0 \\ else & M^t_{x,y} = 1 \end{cases} \tag{3}$$

***(iii)* Background pixels update.** Two update levels are envisaged: frame level or pixel level. The background update at the frame level is performed when a likelihood test on the dynamic matrix is verified. This test implies that the input frame corresponds to a background without moving objects. It is a way to intercept a frame representative of typical values of background pixels. If this likelihood test is not verified, the background update relies on the dynamic matrix to limit the linear update to a limited number of background pixels.

2.2.3. Subtraction/ Comparison

Following the background model update, a pixel classification step into background pixel or foreground pixel is performed to get a binary map for each pixel. For this reason, each new frame is subtracted from the background model built using an adaptive threshold allowing us to group the connected moving pixels in blobs and refine their shapes. The adaptive threshold is based on a local decision threshold calculated by the method proposed in Otsu (1979).

2.2.4. Post-processing

This step is carried out to refine the detection results by removing uninteresting moving areas and eliminating holes and noise from the moving regions.

2.3. Temporal Behavior

Our method of switching between full-spectrum light sources finds its originality not only in the reduction of the detection time, but also in the way to reach the objective via an appropriate temporal behavior (*cf.* Figure 2). Firstly, the speed of our approach is ensured by the use of one of the two spectra IR or VIS, unlike other fusion methods suggested in the literature which rely on a complementary treatment of the two spectra. In fact, low-level pixel-based fusion techniques require a synchronization step and the fusion of the two infrared and visible spectra into a single spectrum. Low-level region-based fusion techniques, however, perform a "double" detection since the detection of the moving objects is achieved, first, in each spectrum before subsequently merging the obtained results. Secondly, we have proceeded in two different ways at night and during the day. Indeed, during the night and especially in an outdoor scene, the detection of moving objects is better using the IR spectrum. However, during the day a switch between the two spectra is interesting especially for winter days that have alternating weather conditions and brightness variations. For this reason, and while avoiding a classification of all the images of a video stream, we proposed to proceed with a periodic treatment as shown in Figure 2. We begin with the use of the selected spectrum (S_0) resulting from the classification of the first images. This spectrum is used for the detection of moving objects during a period P at the end of which a new classification (C) and selection (S) of the adequate spectrum are performed. The duration of the period P is empirically defined according to the season and the geographical location of the scene to be monitored.

In the following section, several quantitative and qualitative experiments were identified and detailed to validate and demonstrate the robustness of our proposed method.

3. Experimental results

In order to evaluate the performance of the proposed method of moving object detection using full-spectrum light sources, three series of experiments were conducted. In the first one, we assessed the performance of our moving object detection method when facing different challenging situations using only one of the two spectra IR or VIS. In the second one, we evaluated the performance of our proposed method of moving object detection using our switching strategy between the full-spectrum light sources under different weather conditions, rather than using each of the two spectra independently. In the third one, a comparison of our method with three recent and well-known methods St-Laurent et al. (2013); Mouats and Aouf (2014); Mangale and Khambete (2016, 2018) of moving object detection using the full-spectrum light sources was established. These methods are based on background modeling for their moving object detection methods and differ in the way the IR and VIS spectra data are merged. Before presenting the results of these series of experiments, we introduced our dataset and the used evaluation metrics.

3.1. Datasets description

In a first step, to evaluate the performance of the moving object detection method while facing different challenges (*i.e.* moving background, ghost phenomenon, shadows, typical thermal artifacts, etc.) we have used a challenging dataset called

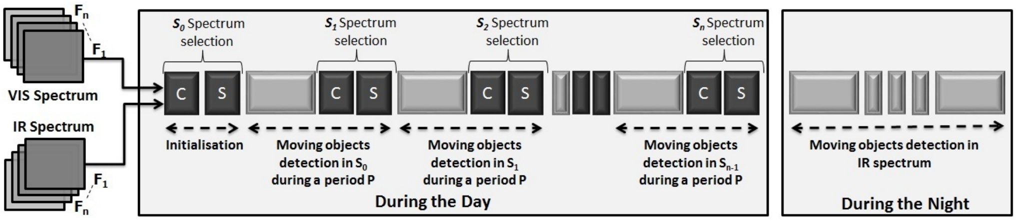


Fig. 2. Temporal behavior of our full-spectrum light sources switching method.

CD.net 2014[1]. This dataset covers a wide range of detection challenges and is representative of typical indoor and outdoor visual data Wang et al. (2014). In a second step, to assess the performance of the proposed method for moving object detection exploiting full-spectrum light sources, we used two popular databases with typical weather conditions. The first dataset was extracted from the OTCBVS Benchmark Dataset, namely, "*OSU Color-Thermal Database*" Davis and Sharma (2007). This dataset consists of thermal/color video pair sequences recorded on the Ohio State University campus during the months of February and March 2005. These sequences, involving 320*240 pixel-sized images, were captured on a cloudy day, with fairly constant illumination and soft/diffuse shadows. A manual segmentation of the moving object regions was performed on 42 images both in VIS and IR sequences for the quantitative evaluation. The results of the manual segmentation of each pair of IR and VIS images were merged using the logical AND to ensure on accurate segmentation of moving regions. The second dataset is the "*INO Video Analytics dataset*" which is made available with its ground-truth frames by St-Laurent et al. (2007). This dataset consists of ten thermal/color video sequences. As our aim was to overcome the challenges related to the low illumination and bad weather conditions, we selected eight pair thermal/color video sequences which are the most representative sequences recorded in various locations and covering different weather conditions. Table 1 provides a brief summary of the video sequences used in our experiments allowing a rich test field to validate our proposed method.

Table 1. Details of the OTCBVS and INO datasets

Datasets	IR/VIS Sequences	# Frames	Weather Conditions	Moving Objects	Description	Challenges
OTCBVS	OSU Color/ Thermal	17082	NS	Person	Daytime scene with monotonous illumination condition	Shadows(VIS) Halos(IR)
INO	Coat Deposit (CD)	551	NS	Car & Person	Sunny day	Size and Speed
	Main Entrance (ME)	2030	NS	Car & Person	Sunny afternoon	Size and Speed
	Parking Snow (PS)	2941	NS	Car & Person	Daytime on a cloudy day	Occlusion
	Close-Person (CP)	240	NS	Person	Sunny day	Camouflage
	Group Fight (GF)	1482	NS	Car & Person	Daytime scene	Occlusion
	Multiple Deposit (MD)	2400	NS	Car & Person	Daytime scene	large illumination variations
	Parking Evening (PE)	820	AS	Car & Person	Evening scene	Low Contrast of Moving Objects
	Visitor Parking (VP)	472	AS	Car & Person	Cloudy and rainy day	Low Contrast of Moving Objects

[1] http://jacarini.dinf.usherbrooke.ca/

3.2. Evaluation metrics

For the first and second series of experiments, aiming at evaluating the effectiveness of the proposed method, we relied on the confusion matrix. From this confusion matrix, *Recall* (*R*), *Precision* (*P*) and *F_measure* (*F*) were measured using the set of ground truth images, to investigate the performance of our method. The *Recall* rate was calculated to know the fraction of moving object pixels that were correctly detected (*cf.* Eq. 4). The *Precision* rate stands for the rate of the correct classification (*cf.* Eq. 5). The *F_measure* Liu and Zsu (2009), which is the harmonic mean of *Recall* and *Precision* was used. In fact, a higher *F_measure* value corresponds to a higher value of *Recall* and *Precision* (*cf.* Eq. 6).

$$Recall = \frac{TP}{TP+FN} \tag{4}$$

$$Precision = \frac{TP}{TP+FP} \tag{5}$$

$$F\_measure = \frac{2*P*R}{P+R} \tag{6}$$

Where, (*TP*) is the number of *True Positives*, (*FP*) of *False Positives* and (*FN*) of *False Negatives*.

The third series of experiments consists of comparing our method with four recent and well-known methods from the literature namely those of St-Laurent et al. (2013); Mouats and Aouf (2014); Mangale and Khambete (2016, 2018). The authors of Mouats and Aouf (2014); Mangale and Khambete (2016, 2018) used the same evaluation metrics whereas the authors of St-Laurent et al. (2013) computed the following metrics: the *Jaccard coefficient* (*J*) (*cf.* eq. 7), the *Detection Rate* (*DR*) (*cf.* eq. 8) and the *False Alarm Rate* (*FAR*) (*cf.* eq. 9), to evaluate their results of moving object detection. To compare the results of our method with those of St-Laurent et al. (2013), the metrics *J*, *DR* and *FAR* were calculated.

$$J = \frac{TP}{TP+FP+FN} \tag{7}$$

$$DR = \frac{TP}{TP+FN} \tag{8}$$

$$FAR = \frac{FP}{TP+FP} \tag{9}$$

3.3. Results and Discussion

In this section, the results of the three series of experiments were detailed and discussed.

3.3.1. *Results of the Moving Object Detection (MOD) method*

To evaluate the performance of our *MOD* method, we have compared the obtained results on the *CD.net 2014* dataset with recent methods of Sajid and Cheung (2017), Jiang and Lu (2018), Paul et al. (2017), Braham et al. (2017) and Wang et al. (2018). In fact, we present in Table 2 a category wise comparison, in terms of average of F-measure. Our method has recorded comparable rates and even outperforms the results of some recent methods.

Table 2. Category-Wise Comparison (Average of *F-measure*) on the *CD.net 2014* Dataset

Methods	Baseline	Int. Mov. Obj.	Dynamic BG	Shadow	Thermal
Our *MOD*	0,912	**0,799**	0,616	0,885	0,819
Sajid and Cheung (2017)	0,888	0,763	0,783	0,776	0,789
Jiang and Lu (2018)	0.941	0.739	0.744	0.869	0.796
Paul et al. (2017)	**0,966**	–	0,709	–	–
Braham et al. (2017)	0.960	0.787	**0.949**	**0.924**	**0.822**
Wang et al. (2018)	0,932	0,694	0,686	0,897	0,745

Although our method does not record the best rates in all the categories compared to the other works, it has a competitive computation time. In fact, to estimate the temporal complexity of our method, we have reported the average computation time (in seconds) per frame for moving object detection and this for each sequence category of *CD.net 2014* dataset. Table 3 shows the average computational overhead calculated over a fixed buffer size of 15 frames for each video category. Furthermore, we have compared our average time consumption per frame against other moving object detection methods of the literature Paul et al. (2017) Pojala et al. (2011) Zheng et al. (2006) Jodoin et al. (2007). In fact, the results presented in Table 4 show that our *MOD* method records the best computation time rates.

Table 3. Computational overhead (in sec) per input frame for different video categories

Video Category	*Baseline*	*Int. Obj. Mot.*	*Dynamic BG*	*Shadow*	*Thermal*	Average Time
Time (s)	1,221	1,312	3,274	2,941	1,56362	2,06

Table 4. Computational comparison of our Moving Object Detection (*MOD*) with other methods

Methods	*Our MOD*	Paul et al. (2017)	Pojala et al. (2011)	Zheng et al. (2006)	Jodoin et al. (2007)
Average Time consumption (s)	**2,06**	2,78	3,65	2,46	2,97

The importance of our framework is linked to the complexity of the observed scenes with a variety of challenges. Our moving object detection method has proven its potential to overcome a lot of challenges. Figure 3 shows some qualitative results that prove the strengths and the weaknesses of our method for moving object detection (*MOD*). In fact, the first three lines show the performance of our method to detect moving objects with different sizes and different speeds (vehicle, cyclist and pedestrian). It is obvious that these lines met the challenges of the background movement and the shadows casted by moving objects. This worth reminding that the elimination of cast shadows was achieved relying on our previous work Jarraya et al. (2012). Likewise, our method performs well in IR spectrum and overcomes many artifacts presented in IR videos such as camouflage effect, ghost, heat reflection on floors and windows and heat stamps (e.g., bright spots left on a seat after a person gets up and leaves) (*cf.* line 4 in Figure 3). Nevertheless, our *MOD* method has failed to eliminate the strong background motion (parasite) as it is the case in the sequences of the dynamic background category (*cf.* line 5 in Figure 3). Finally, as our background model is regularly updated to adapt to different changes in the scene, any abandoned object in the scene, after a few frames, becomes as an element of the background model. However, in some sequences of *CD.net2014* dataset the abandoned object sometimes remains detected in the Ground-Truth (*GT*) images which have affected our detection results.

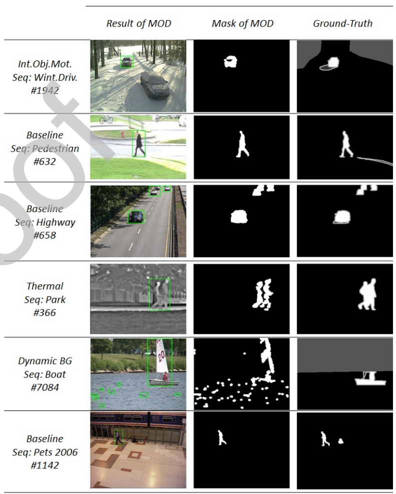


Fig. 3. Qualitative Results of *MOD* on *CD.net2014* Dataset.

3.3.2. *Independent spectra vs. switching between full-spectrum light sources under different weather conditions*

The main purpose of this series of experiments was to evaluate the robustness of the proposed method for moving object detection. The originality of the proposed method lies in its ability to perform the task either on the IR or on the VIS spectra according to the level of illumination and the state of the weather conditions. In fact, we set forward a new strategy to switch between the spectra light sources based on weather conditions classification prediction model. This series of experiments targeted two objectives. Firstly, we aimed to prove the robustness of our moving object detection method on the IR spectrum and VIS spectrum independently. Secondly, we attempted to assess the effectiveness of our switching method between the full-spectrum light sources. To this end, we started by presenting the classification rates of our prediction model *WCC-PM* on each sequence of the test dataset. Then, we proved the efficiency of our switching method against using each spectrum independently under different weather conditions.

We began by assessing quantitatively our moving object detection method in each spectrum independently. Indeed the results presented in table 5 highlight the high accuracy recorded

Table 5. Moving object detection results, in terms of *Recall* (*R*), *Precision* (*P*) and *F_measure* (*F*), in the IR and VIS spectra on each sequence of the test dataset

		OTCBVS	CD	ME	PS	CP	MD	GF	PE	VP
IR Spectrum	R	99,23	82,89	54,65	72,42	70,65	62,59	80,07	81,75	82,04
	P	77,69	95,11	89,55	74,42	75,88	62,77	81,91	92,22	72,34
	F	87,15	88,58	67,88	73,4	73,17	62,68	**80,98**	**86,67**	**76,88**
VIS Spectrum	R	93,75	94,39	74,22	83,72	78,17	75,79	68,21	75,7	85,77
	P	99,1	92,72	89,89	96,65	92,95	82,26	98,92	47,73	49,65
	F	**96,35**	**93,55**	**81,31**	**89,72**	**84,92**	**78,89**	80,75	58,55	62,9

Table 6. *Correct Classification Rates* (*CCR*) of our *WCC-PM* on each sequence of the test dataset in VIS spectrum

	OTCBVS	*CD*	*ME*	*PS*	*CP*	*MD*	*GF*	*PE*	*VP*
CCR %	100	99.41	100	98,88	100	99,75	97,91	100	99,36

in all sequences, proving the performance and the efficiency of our method. The *PE* and *VP* sequences display low precision values in the VIS spectrum, owing to the low contrast and visibility of the moving objects. These sequences were recorded in the evening with a low illumination and in bad weather conditions, as well.

The originality of the proposed method resides in our switching strategy between the full-spectrum lights sources. Our method is based on a weather conditions classification step of visible images. From this perspective, Table 6 displays the *Correct Classification Rate* (*CCR*) of our prediction model *WCC-PM* on each sequence of the test dataset whose images are classified as in *A.S.* or in *N.S.* Indeed, the *OTCBVS*, *CD*, *ME*, *PS*, *CP*, *MD* and *GF* sequences images are classified as in *N.S.* respectively with the following *CCR* percentages: 100%, 99.41%, 100%, 98.88%, 100%, 99.75%, 97.91%. On the other hand, the *PE* and *VP* sequences were classified as in *A.S.* with a *CCR* of 100% and 99.36%. These findings reflect the robustness and feasibility of our prediction model in terms of the correct classification of VIS images.

To confirm the effectiveness of our switching method, we compared the results of the detection of moving objects obtained using the IR spectrum, the VIS spectrum and the switching between full-spectrum light sources. The results are summarized in Table 7, in terms of average of *Recall*, *Precision* and *F-measure* of all the sequences of our datasets. These results prove the effectiveness of our method of moving object detection based on the switching between full-spectrum light sources. In fact, the recorded rates by our switching strategy exceeded those obtained using each spectrum separately. For example, the F-measure increased from 77,487% (using only the IR spectrum) and from 80,992% (using only the VIS spectrum) to 86,41% using our switching between the full-spectrum light sources.

Table 7. Moving object detection results, in terms of *Recall*, *Precision* and *F-measure* according to the spectrum used

	IR Spectrum	VIS Spectrum	IR-VIS switching method
R	76,254	81,656	**82,14**
P	80,21	83,154	**90,6**
F	77,487	80,992	**86,17**

3.3.3. Comparison with state-of-the-art methods

In this series of experiments, we established a comparison between our work and four recent works in the literature, namely, the work of St-Laurent et al. (2013), Mouats and Aouf (2014), Mangale and Khambete (2016) and Mangale and Khambete (2018). These methods are similar to ours as they propose detection methods based on background modeling using the full-spectrum light sources. The comparative study was performed using the same metrics and the same sequence frames presented by each work. Since they use neither the same metrics nor the same sequences, we compared our results with each of them independently. In a first step, we compared our results to those of St-Laurent et al. (2013) which proposed a fusion of thermal and color images for moving object detection in an outdoor environment. Their method is based on a background modeling approach and on a pixel-level fusion of VIS and IR spectra. In fact, they fused the data from both sensors in a single hybrid codebook in which each pixel is represented by *L CodeWords* (*CW*) that are the Luma, Chroma orange, Chroma green of the *YCoCg* color space, Thermal values and some parameters to match the pixels *CW*. In order to maximize the detection accuracy, they proposed to update the thermal detection threshold at every frame based on a periodically updated standard deviation (temporal sensor noise), and on a weighted decay of the intensity variation measured on previous consecutive thermal frames. However, to adjust this threshold the user manually sets the values of the four parameters in each video sequence, which is the major drawback of their proposed method. Table 8 displays the quantitative comparison results between our method and the work of St-Laurent et al. (2013) in terms of *DR*, *FAR* and *J* metrics on *CP*, *GF*, *PS* and *MD* sequences. The fact that their detection results somehow achieve higher rates than ours, does not diminish the intrinsic added value of our proposed method in any way. In fact, this superiority is basically due to the use of a set of optimized parameters fixed by the authors for each video sequence, which is not the case in our method.

Table 8. Comparison of the performance of the proposed method with the work of St-Laurent et al. (2013)

		St-Laurent et al. (2013)	*FSLS-MOD*
CP	*DR*	**0,898**	0,782
	J	**0,7974**	0,729
	FAR	0,123	0,071
GF	*DR*	**0,935**	0,755
	J	**0,85**	0,736
	FAR	0,097	0,027
PS	*DR*	**0,947**	0,837
	J	**0,894**	0,813
	FAR	0,059	0,034
MD	*DR*	**0,845**	0,758
	J	**0,6892**	0,62
	FAR	0,211	0,177

In a second step, we compared our results with those of Mouats and Aouf (2014). These authors proposed a moving object detection method based on background modeling using the Gaussian mixture models and presented different strategies to fuse the IR and VIS spectra. In fact, we compared our results with the two best of theirs. The first one combines data from the thermal and visible cameras prior to background subtraction. The three channels of the visible camera are reinforced

Table 9. Comparison of our detection results using the full-spectrum light sources with those of Mouats and Aouf (2014); in terms of *F_measure* average on *GF*, *PS*, *PE* and *MD* sequences

	RGBT Mouats and Aouf (2014)	*Fus85IR-15VIS* Mouats and Aouf (2014)	*FSLS-MOD*
GF	0,819	**0,852**	0,85
PS	0,717	**0,9**	0,897
MD	**0,802**	0,649	0,789
PE	0,694	0,812	**0,867**

Table 10. Comparison of the obtained precision rates of our method and those of Mangale and Khambete (2016) using full-spectrum light sources

	Mangale and Khambete (2016)	*FSLS-MOD*
GF	0,793	**0,988**
PS	0,822	**0,987**
CP	**0,959**	0,947
CD	0,652	**0,916**
ME	**0,954**	0,885

with information from the IR camera. Therefore, pixels are modeled using a *4D* vector (*RGBT*) rather than the original *3D* (*RGB*) formulation. All the pixel channels are weighted equivalently meaning that no preference is attributed to any sensor. In the second image fusion, they generate a fused image using a linear combination of IR and VIS images. The background subtraction is then run on the fused image. Their best combination, named *85IR-15VIS*, is the one that integrates 85% of thermal information with 15% of visible information in the fused images. However, their foreground segmentation is very scene dependent as the detection results vary considerably from one sequence to another. Table 9 shows a comparison between our detection results and those of Mouats and Aouf (2014) in terms of *F-measure* average on *GF*, *PS*, *MD* and *PE* sequences. In fact, our method gives comparable results and even outperforms their results using one of the two low-level fusion methods (*Fus85IR-15VIS* and *RGBT*). We outperformed their methods owing to our switching strategy which allows us to benefit from the quality of the full-spectrum light sources, without having to correlate the spectra or to generate a fused spectrum.

In a third step, we compared our method to that of Mangale and Khambete (2016). The authors detected the moving objects by a background modeling method carried out on each spectrum independently; then, the foreground regions are merged by a low level fusion technique. Indeed, their fusion is achieved by a logical "OR" operation between the masks of the foreground regions. The comparison of our detection results with those of Mangale and Khambete (2016) is presented in Table 10, in terms of average precision of an image set selected by the authors in each sequence. As shown in this table, we achieved the best rates in almost all the sequences. For example, in *GF*, *PS* and *CD* sequences we reached respectively the precision rates of 0.988%, 0.987% and 0.916% compared to 0.793%, 0.822% and 0.652% recorded by Mangale and Khambete (2016), except for the *ME* sequence, where we achieved 0.885% while the authors of Mangale and Khambete (2016) reached up to to 0.954%.

Finally, in the recent work of Mangale and Khambete (2018) the authors have proposed a new approach for the detection of moving objects using a structural similarity metric (*SSIM*) and the Gaussian mixture model (*GMM*). The *SSIM* was used to compute structural similarity map between the reference mean background frame and the foreground frame of visible spectrum (VIS) and thermal infrared (IR) independently. The threshold results of *SSIM* were applied to this map to find the moving objects and then a fusion of the VIS and thermal IR modality was carried out using different pixel-level fusion methods such logical "*OR*", discrete wavelet transform, and principal components analysis. Finally, they used a temporal analysis to eliminate noise using GMM on the fused results. Table 11 shows a comparison between our *FSLS-MOD* method and their results using logical *OR* rule to fuse the IR and VIS spectra, in terms of average of *Recall*, *Precision* and *F-measure* metrics in the *GF*, *PS*, *PE*, *CP*, *ME*, *CD* and *MD* sequences. All the methods

Table 11. Comparison of the performance of our *FSLS-MOD* method with the work of Mangale and Khambete (2018)

		Mangale and Khambete (2018)	*FSLS-MOD*
GF	*R*	0,7925	0,755
	P	0,9797	0,973
	F	**0,8561**	0,851
PS	*R*	0,6339	0,837
	P	0,8543	0,967
	F	0,7243	**0,897**
PE	*R*	0,8107	0,818
	P	0,8497	0,922
	F	0,8391	**0,867**
CP	*R*	0,7934	0,782
	P	0,9718	0,93
	F	**0,8758**	0,85
ME	*R*	0,7770	0,742
	P	0,7569	0,899
	F	0,7574	**0,813**
CD	*R*	0,6267	0,944
	P	0,8390	0,927
	F	0,7240	**0,935**
MD	*R*	0,6947	0,758
	P	0,7643	0,823
	F	0,7107	**0,789**

presented above have proposed moving object detection methods based on low-level fusion methods to fuse the IR and VIS spectra. Their fusion methods require either to generate new images on which the detection method is executed St-Laurent et al. (2013); Mouats and Aouf (2014) or to apply the foreground detection method on each spectrum implying that these regions are merged by a sort of fusion method Mangale and Khambete (2016, 2018). This fusion is of a very low processing level. Indeed, in the case of bad weather conditions, the noise recorded in the VIS spectrum will be transmitted to the fused image, which degrades the quality of the images and influence the performance of the detection methods. As far as, the hot summer day is concerned, the IR image will provide a lot of hot areas or objects. As a matter of fact, the IR camera would act poorly and a low level fusion would transfer this weakness to the merged images. For this reason, our proposed switching method seems to be a better candidate for a fusion of a higher processing level which takes into account intelligent information in a more powerful way, depending on the situation and the context in which the system is running. The high accuracy recorded by our method proves the effectiveness and the feasibility of our system to run under different situations and various weather conditions.

4. Conclusion

In this paper, our central focus was upon overcoming the challenges related to the low illumination and weather conditions such as fog, snow, rain, darkness, etc. Thus, we proposed a new method of moving object detection using a switching strategy between the full-spectrum light sources so as to benefit from the advantages of each spectrum. The originality of our method stems from the fact that we performed our detection alternating the IR and VIS spectra depending on the illumination level and the weather condition state in the VIS images. Furthermore, our moving object detection method is based on background modeling incorporating the principle of inter-frame difference in the background modeling stage. An experimental investigation was performed proving the effectiveness of our moving object detection method based on a switching strategy between the full-spectrum light sources. At this stage of analysis, we would assert that our research is a step that might be taken further. Our promising results offer different prospects and open new horizons for future works to examine the semantic classification of the detected moving objects in the IR and VIS spectra in order to fulfill a constructive and fruitful contribution to the field.

References

Boukhriss, R.R., Fendri, E., Hammami, M., 2015. Intelligent Fusion of Infrared and Visible Spectrum for Video Surveillance Application. Springer International Publishing, Cham. pp. 768–777.

Bouwmans, T., 2014. Traditional and recent approaches in background modeling for foreground detection: An overview. Computer Science Review 11-12, 31 – 66.

Braham, M., Pirard, S., Van Droogenbroeck, M., 2017. Semantic background subtraction, in: 2017 IEEE International Conference on Image Processing (ICIP), pp. 4552–4556.

Chang, M., Cheng, Y., 2007. Motion detection by using entropy image and adaptive state-labeling technique, in: 2007 IEEE International Symposium on Circuits and Systems, pp. 3667–3670.

Chen, B.W., Liu, S.L., 2014. Infrared target detection based on temporal-spatial domain fusion, in: Frontiers of Energy, Materials and Information Engineering, Trans Tech Publications. pp. 1186–1189.

Cheng, Y.H., Wang, J., 2014. A motion image detection method based on the inter-frame difference method, in: Mechanical Design and Power Engineering, Trans Tech Publications. pp. 1283–1286.

Conaire, C.O., Cooke, E., O'Connor, N., Murphy, N., Smearson, A., 2005. Background modelling in infrared and visible spectrum video for people tracking, in: 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05) - Workshops, pp. 20–20.

Davis, J.W., Sharma, V., 2007. Background-subtraction using contour-based fusion of thermal and visible imagery. Computer Vision and Image Understanding 106, 162 – 182. Special issue on Advances in Vision Algorithms and Systems beyond the Visible Spectrum.

Hammami, M., Jarraya, S.K., Ben-Abdallah, H., 2013. On line background modeling for moving object segmentation in dynamic scenes. Multimedia Tools Appl. 63, 899–926.

Haritaoglu, I., Harwood, D., Davis, L.S., 2000. W4: Real-time surveillance of people and their activities. IEEE Transactions on Pattern Analysis and Machine Intelligence 22, 809–830.

Jarraya, S.K., Boukhriss, R.R., Hammami, M., Ben-Abdallah, H., 2012. Cast shadow detection based on semi-supervised learning, in: Campilho, A., Kamel, M. (Eds.), Image Analysis and Recognition, Springer Berlin Heidelberg, Berlin, Heidelberg. pp. 19–26.

Jiang, S., Lu, X., 2018. Wesambe: A weight-sample-based method for background subtraction. IEEE Transactions on Circuits and Systems for Video Technology 28, 2105–2115.

Jodoin, P., Mignotte, M., Konrad, J., 2007. Statistical background subtraction using spatial cues. IEEE Transactions on Circuits and Systems for Video Technology 17, 1758–1763.

Kushwaha, A.K.S., Srivastava, S., Srivastava, R., 2017. Multi-view human activity recognition based on silhouette and uniform rotation invariant local binary patterns", journal="multimedia systems 23, 451–467.

Lim, L.A., Keles, H.Y., 2018. Foreground segmentation using convolutional neural networks for multiscale feature encoding. Pattern Recognition Letters 112, 256 – 262.

Lin, C., Lu, J., Wang, G., Zhou, J., 2018a. Graininess-aware deep feature learning for pedestrian detection, in: Ferrari, V., Hebert, M., Sminchisescu, C., Weiss, Y. (Eds.), Computer Vision – ECCV 2018, Springer International Publishing, Cham. pp. 745–761.

Lin, C., Lu, J., Zhou, J., 2018b. Multi-grained deep feature learning for robust pedestrian detection. IEEE Transactions on Circuits and Systems for Video Technology PP, 1–14.

Liu, L., Zsu, M.T., 2009. Encyclopedia of Database Systems. Springer Publishing Company, Incorporated. 1st edition.

Mangale, S., Khambete, M., 2016. Camouflaged Target Detection and tracking using thermal infrared and visible spectrum imaging. Springer International Publishing, Cham. pp. 193–207.

Mangale, S.A., Khambete, M.B., 2018. Approach for moving object detection using visible spectrum and thermal infrared imaging. Journal of Electronic Imaging 27, 1 – 12.

Mouats, T., Aouf, N., 2014. Fusion of thermal and visible images for day/night moving objects detection, in: 2014 Sensor Signal Processing for Defence (SSPD), pp. 1–5.

Otsu, N., 1979. A threshold selection method from gray-level histograms. IEEE Transactions on Systems, Man, and Cybernetics 9, 62–66.

Paul, N., Singh, A., Midya, A., Roy, P.P., Dogra, D.P., 2017. Moving object detection using modified temporal differencing and local fuzzy thresholding. J. Supercomput. 73, 1120–1139.

Pojala, C., Somnath, S., and, 2011. Moving object detection in the presence of dynamic backgrounds using intensity and textural features. Journal of Electronic Imaging 20, 043009.

Qi, Y., An, G., 2011. Infrared moving targets detection based on optical flow estimation, in: Proceedings of 2011 International Conference on Computer Science and Network Technology, pp. 2452–2455.

Qin, M., Lu, Y., Di, H., Huang, W., 2016. A background basis selection-based foreground detection method. IEEE Transactions on Multimedia 18, 1283–1296.

Ribeiro, M., Lazzaretti, A.E., Lopes, H.S., 2018. A study of deep convolutional auto-encoders for anomaly detection in videos. Pattern Recognition Letters 105, 13 – 22. Machine Learning and Applications in Artificial Intelligence.

Sajid, H., Cheung, S.S., 2017. Universal multimode background subtraction. IEEE Transactions on Image Processing 26, 3249–3260.

St-Laurent, L., Maldague, X., Prevost, D., 2007. Combination of colour and thermal sensors for enhanced object detection, in: 2007 10th International Conference on Information Fusion, pp. 1–8.

St-Laurent, L., Prvost, D., Maldague, X., 2013. Combination of thermal and color images for accurate foreground / background segmentation in outdoor environment, in: 2013 IEEE International Conference on Image Processing, pp. 3431–3435.

Wang, K., Gou, C., Wang, F., 2018. M4CD: A robust change detection method for intelligent visual surveillance. CoRR abs/1802.04979.

Wang, Q., Gao, J., Yuan, Y., 2018. A joint convolutional neural networks and context transfer for street scenes labeling. IEEE Transactions on Intelligent Transportation Systems 19, 1457–1470.

Wang, Y., Jodoin, P., Porikli, F., Konrad, J., Benezeth, Y., Ishwar, P., 2014. Cdnet 2014: An expanded change detection benchmark dataset, in: 2014 IEEE Conference on Computer Vision and Pattern Recognition Workshops, pp. 393–400.

Yan, Y., Ren, J., Zhao, H., Sun, G., Wang, Z., Zheng, J., Marshall, S., Soraghan, J., 2018. Cognitive fusion of thermal and visible imagery for effective detection and tracking of pedestrians in videos. Cognitive Computation 10, 94–104.

Yin, J., Liu, L., Li, H., Liu, Q., 2016. The infrared moving object detection and security detection related algorithms based on w4 and frame difference. Infrared Physics & Technology 77, 302 – 315.

Zheng, J., Wang, Y., Nihan, N.L., Hallenbeck, M.E., 2006. Extracting roadway background image: Mode-based approach. Transportation research record 1944, 82–88.

Declaration of Competing Interest

All authors have participated in (a) conception and design, or analysis and interpretation of the data; (b) drafting the article or revising it critically for important intellectual content; and (c) approval of the final version.

This manuscript has not been submitted to, nor is under review at, another journal or other publishing venue.

The authors have no affiliation with any organization with a direct or indirect financial interest in the subject matter discussed in the manuscript.