Foreground detection enhancement using Pearson correlation filtering

Rafael Marcos Luque-Baena, Miguel A. Molina-Cabello, Ezequiel López-Rubio, and Enrique Domínguez

Department of Computer Languages and Computer Science. University of Málaga.
Bulevar Louis Pasteur, 35. 29071 Málaga. Spain.
{rmluque,miguelangel,ezeqlr,enriqued}@lcc.uma.es

Abstract. Foreground detection algorithms are commonly employed as an initial module in video processing pipelines for automated surveillance. The resulting masks produced by these algorithms are usually postprocessed in order to improve their quality. In this work, a postprocessing filter based on the Pearson correlation among the pixels in a neighborhood of the pixel at hand is proposed. The flow of information among pixels is controlled by the correlation that exists among them. This way, the filtering performance is enhanced with respect to some state of the art proposals, as demonstrated with a selection of benchmark videos.

Keywords: Foreground detection, postprocessing, Pearson correlation, morphological operators, background modeling

1 Introduction

Extracting objects of interest from a video or image is a very important task in computer vision applications. Numerous features can be extracted from the foreground to develop classifying and recognizing processes in subsequent steps. For instance, autonomous visual systems must be able to recognize relevant objects and its movements in order to maintain an internal representation of the environment and understand the scene. Most of surveillance systems are only interested on the moving objects, so the aim of these segmentation algorithms[10, 6] consists in separating the foreground pixels from the background pixels.

In the literature, many foreground detection algorithms [9, 7, 2, 8] have been proposed, whose internal parameters (thresholds, sizes of regions...) and postprocessing techniques are fixed to obtain meaningful results depending on the application.

The result produced by foreground object detection algorithms usually contains noise and it is not suitable to carry out high-level processing like object tracking or the analysis of the behavior of these objects. So that, it is necessary to execute a filtering of the binary mask obtained in the segmentation step. Some of the possible causes that can produce this kind of noise are listed below:

- Camera noise. It is produced by the quality of the acquisition of the images from the camera. Sometimes, a pixel from an image presents a color tone, and the same pixel in the next frame (without any movement in the scenario) exhibits a different one. In this kind of noise, the noise produced by the video downsampling applied in the hardware device can be included.
- Reflection noise. The movement of a spotlight, for example, the sun, produces that some background parts reflect the light and the result of the foreground object detection algorithm is affected by this effect, considering those zones as foreground.
- Noise in the background objects. Several parts of the objects have the same color (or tone in the case of grayscale images) as the background behind them. This similarity produces that some algorithms do not detect these pixels as foreground objects, so they are not correctly detected.
- Shadows and abrupt illumination changes. Most algorithms detect the projected shadows of the objects as foreground. The illumination changes (for example, turn on a light in a room) also produce that the algorithms fail in the detection of the foreground objects.

All of these commented failures can not be solved by the objects in motion segmentation pixel-level algorithms. Thus, it is necessary to develop postprocessing techniques in order to improve the quality of the final segmentation.

2 Model Approach

In this section, the foreground object detection algorithm (2.1), considered as our baseline method, is described. After that, the morphological operators are depicted (subsection 2.2), since they are the most applied technique in order to remove the noise. And finally we also propose the employ of the Pearson correlation to reduce the noise (subsection 2.3).

2.1 Foreground object detection algorithm

A foreground object detection algorithm provides a mask for each frame of a sequence where it presents the same size than the input frame and each pixel has a value in the range $[0, 1]$ that represents the likelihood (in order to manage the uncertainty) of belonging to the foreground. Thus, pixels from this mask with a value close to 1 (white pixels) represent the foreground objects, while pixels with a value close to 0 (black pixels) are considered as background.

In order to quantify the improvement of the application of a postprocessing method to the result produced by the algorithms of this kind, we need to compare the output of the algorithm and the result provided by the postprocessing method.

In this case, we have considered as our baseline method the algorithm described in [5], and noted as AE. This proposal uses mixtures of uniform distributions and multivariate Gaussians with full covariance matrices, and it is indicated
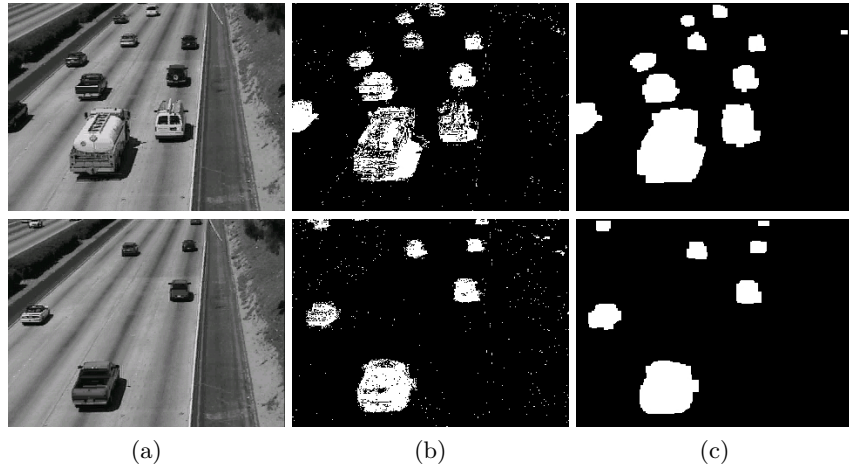
Fig. 1. Result after the noise removal. (a) Frames corresponding to a traffic sequence. (b) Segmentation by Gaussian distributions. (c) Noise removal by morphological operators.

to detect foreground objects in complex context, like videos which exhibit dynamic backgrounds or shadow appearances. In addition, due to the employ of a stochastic approximation, the computational complexity of this algorithm is low, so it is a suitable method for real time applications.

2.2 Morphological operators

Morphological operators like erosion or dilation are applied to the segmentation of the foreground objects in order to remove spurious pixels [1, 3], that cause the three first items referred in Section 1. The aim of the application of this kind of process is to remove those pixels that do not belong to the foreground (NFN, *non-foreground noise*) and to delete those that they are detected as background (NBN, *non-background noise*) in the closest zones and the interior of the objects that actually belong to the foreground.

The erosion process erodes a unit over the external limits of the objects. The dilation is the opposite process, expanding the limits of the foreground objects. The decision about the order and the quantity of filters to be applied is quite significant. The order of the operators affects the quality and the quantity affects the quality and time complexity of the algorithm.

For example, if we apply the dilation and then the erosion process, we can not remove isolated unit pixels (NFN), since the dilation operator increases its limits with one pixel and the erosion will remove the added pixels, thereby keeping the spurious original pixels. On the other hand, this order will remove some NBN noise from the inside of the objects (gaps). In the case of the application of this operations in inverse order, so that, dilation after erosion, it will remove isolated

unit pixels that do not belong to the foreground (NFN), but the filling of the existing gaps inside the objects (NBN) can not be carried out. Thus, the sequence of the application of the morphological operators depends on the peculiarities of the analyzed scene.

2.3 Pearson Correlation

In this subsection we propose the employ of information from the 8 neighbors of a given pixel $\mathbf{x}$ in order to remove the noise. Nevertheless, we can not consider all neighbors in the same proportion due to several adjoining pixels are not related. For example, this can be produced in the border of a road: the outside pixels (where the vehicles do not drive along it) are practically independent from those that belong to the interior of the road (where the vehicles usually circulate) in spite of their proximity.

Thus, we need a quantitative measure of the correlation of pixels pairs. Our selected measure is the Pearson correlation [11] between two random variables, that they will be the likelihood to belong to the foreground, $P_{Fore,\mathbf{x}}$ and $P_{Fore,\mathbf{y}}$, corresponding to each pair of pixels 8 neighbors $\mathbf{x}$ and $\mathbf{y}$:

$$\rho_{\mathbf{x},\mathbf{y}} = \frac{\phi_{\mathbf{x},\mathbf{y}}}{\sqrt{\nu_{\mathbf{x}}}\sqrt{\nu_{\mathbf{y}}}} \tag{1}$$

$$\phi_{\mathbf{x},\mathbf{y}} = \mathrm{cov}\left(P_{Fore,\mathbf{x}}, P_{Fore,\mathbf{y}}\right) =$$
$$E\left[\left(P_{Fore,\mathbf{x}} - E\left[P_{Fore,\mathbf{x}}\right]\right)\left(P_{Fore,\mathbf{y}} - E\left[P_{Fore,\mathbf{y}}\right]\right)\right] \tag{2}$$

$$\nu_{\mathbf{x}} = \mathrm{var}\left(P_{Fore,\mathbf{x}}\right) = E\left[\left(P_{Fore,\mathbf{x}} - E\left[P_{Fore,\mathbf{x}}\right]\right)^2\right] \tag{3}$$

$$\nu_{\mathbf{y}} = \mathrm{var}\left(P_{Fore,\mathbf{y}}\right) = E\left[\left(P_{Fore,\mathbf{y}} - E\left[P_{Fore,\mathbf{y}}\right]\right)^2\right] \tag{4}$$

where we have that:
$$E\left[P_{Fore,\mathbf{x}}\right] = \pi_{Fore,\mathbf{x}} \tag{5}$$

$$E\left[P_{Fore,\mathbf{y}}\right] = \pi_{Fore,\mathbf{y}} \tag{6}$$

and π_{Fore} is the likelihood to belong to the foreground that is updating throughout the time.

Please, note that the properties of the Pearson correlation imply that:

$$\rho_{\mathbf{x},\mathbf{y}} \in [-1, 1] \tag{7}$$

$$\rho_{\mathbf{x},\mathbf{y}} = \rho_{\mathbf{y},\mathbf{x}} \tag{8}$$

where the last equation saves half of the required calculations.

The values of $\rho_{\mathbf{x},\mathbf{y}}$ are high and positive if and only if the pixels $\mathbf{x}$ and $\mathbf{y}$ are usually assigned to the same class, i.e. both pixels belong to the background or the foreground. On the other hand, if the pixels are quite independent, then we will have $\rho_{\mathbf{x},\mathbf{y}} = 0$; we remember that independence implies no correlation,

but the reverse implication is not true, i.e. if it does not exist a correlation between variables then this does not imply that these variables are independent. In relation to the negative correlations, it is expected that they will be infrequent, and theoretically they are related to pairs of pixels that are usually assigned to opposite classes. Nevertheless, in practice, the negative correlations are obtained due to the noise in the estimations of the input data.

The correlations $\rho_{\mathbf{x},\mathbf{y}}$ allow us to obtain a free-noise version of $P_{Fore,\mathbf{x}}(\mathbf{t})$, combining it with the information of the 8 neighbors of $\mathbf{x}$:

$$\widetilde{P}_{Fore,\mathbf{x}}(\mathbf{t}) = \mathrm{ramp}\left(\frac{1}{9}\sum_{\mathbf{y}\in Neigh(\mathbf{x})}\rho_{\mathbf{x},\mathbf{y}}P_{Fore,\mathbf{y}}(\mathbf{t})\right) \quad (9)$$

where $Neigh(\mathbf{x})$ contains the pixel $\mathbf{x}$ and its 8 neighbors, and

$$\rho_{\mathbf{x},\mathbf{x}} = 1 \quad (10)$$

$$\mathrm{ramp}(z) = \begin{cases} z & \text{iff } z \geq 0 \\ 0 & \text{in other case} \end{cases} \quad (11)$$

ramp function is used in the equation (9) in order to fix the variable in the range $[0,1]$ of belonging likelihood to the foreground, when a excess of negative correlations related to the noise is presented. Empirically, it is observed that the argument of the trunc function in the equation (9) is negative in less than 0.5% of the cases, since $\rho_{\mathbf{x},\mathbf{y}}$ are nearly always positive.

The $\widetilde{P}_{Fore,\mathbf{x}}(\mathbf{t})$ values correspond to a probability of belonging to the foreground class, which can be considered a measure of uncertainty. Thus, a pixel whose $\widetilde{P}_{Fore,\mathbf{x}}(\mathbf{t})$ value is close to one will be part of a foreground object with great certainty, while if its value is close to 0.5 it is not possible to deduce anything (it has the same probability of belonging to the foreground as to the background).

3 Experimental results

3.1 Parameter selection

In order to appreciate what kind of filtering is the most convenient, we are going to make a comparison between the two proposed alternatives, the application of morphological operators and the Pearson correlation. For this last case, we have to consider that the the baseline algorithm (AE) has to provide the likelihood that a pixel belongs to the foreground class. We will note it as *Basic* and different learning rate options are considered: $\varepsilon_0 = \{1e^{-4}, 5e^{-4}, 0.001, 0.005, 0.01, 0.05, 0.1\}$. In the case of the morphological operators, we will distinguish in the order application. We note as $OMDE$ the application of the morphological operators in the *dilation - erosion* order, and $OMED$ is related to the *erosion - dilation* order, due to the highly significant influence of a change in the order. We will have

the number of applied iterations as a parameter, where $\#\mathcal{ED} = \#\mathcal{DE} = \{1, 2, 3\}$. For example, a value of $\#\mathcal{DE} = 2$ indicates that two iterations of the dilation operator are carried out, followed by two iterations of the erosion operator.

3.2 Sequences

We have used several sequences in order to make the comparison. This repository of indoor and outdoor videos provides a wide range of the existing variability in real scenes. We have selected the sequences named Video2 (V2) and Video4 (V4) from a synthetic dataset (the sequences are composed by motion objects generated by software with real background scenes) designed for the algorithm competition of foreground objects detection methods created by the International Conference VSSN'06 [1].

Several sequences have been chosen from the dataset[2] developed by Li et al. [4], namely: water surface (WS), moving escalators in a subway station (SS), campus with plentiful vegetation moving continuously (CAM), meeting room with moving curtain (MR) and public fountain throwing water (FT).

We also have taken three videos from the IPPR (Image Processing and Pattern Recognition)[3] contest held in Taiwan in 2006, whose names are IPPRData1 (IP1), IPPRData2 (IP2) and IPPRData3 (IP3). The two first sequences are indoor videos where a corridor is observed from different points of view, and the third video is a highway recorded from an elevated position.

Finally, we have included in our test dataset two sequences from CAVIAR dataset[4]: a sequence where people are walking on the corridor (OC) and an outdoor level crossing scene (LC).

3.3 Qualitative and quantitative results

From a qualitative point of view, a comparison of the four commented options can be observed in Figures 2 and 3, one of them without the application of any postprocessing method (*Basic AE*) and the remaining options by applying the different alternatives described previously (*Pearson*, *OMED* and *OMDE*). The results are obtained with the best tuned configuration that we have tested. It is obvious that the postprocessing method improves and removes the existing noise (false positives) in scenes with a high variability in the background. It can be observed in the sequence *CAM*, in the two first columns of the Figure 2, and *V4* and *WS*, two last columns of the Figure 3. In these videos, the *dilation - erosion* strategy in morphological operators (*OMDE*) is not the best appropriate because the application of the dilation operator in the first step maintains the noise. On the other hand, *Pearson* and *OMED* largely remove the spurious pixels, without any considerable penalization in the real objects in motion. Nevertheless, note

---

[1] `http://mmc36.informatik.uni-augsburg.de/VSSN06_OSAC`

[2] `http://perception.i2r.a-star.edu.sg/bk_model/bk_index.html`

[3] `http://media.ee.ntu.edu.tw/Archer_contest/`

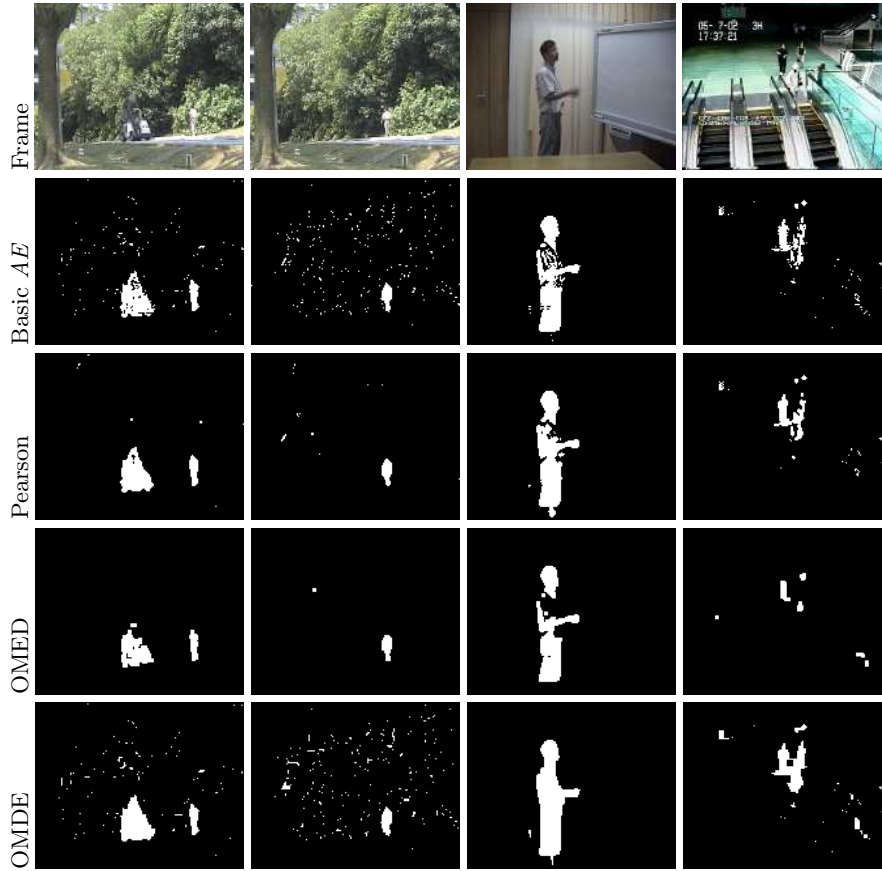[4] `http://homepages.inf.ed.ac.uk/rbf/CAVIARDATA1/`

Fig. 2. Results over complex scenes with stable objects in motion. Two first columns exhibit the exterior sequence *CAM* with vegetation continuously moving on it (frames 1372 and 1392). Third and forth columns show the analysis of the sequences *MR*, meeting room, and *SS*, escalators, over the frames 3242 and 4558, respectively.

that *Pearson* maintains the shape of the object better than *OMED* in spite of filtering the noise of the background in a lower grade.

Nevertheless, in scenes where the appearance of false negatives is prevailing, the employ of *OMED* is worse than the basic output of the segmentation algorithm, as it can be observed in the third and forth columns in the Figure 2. In this case, the postprocessing method with *OMED* produces more false negatives than the original ones, while the output with *OMDE* is the optimal in comparison with the remaining alternatives.

In order to compare the performance of each proposed method from a quantitative point of view, we have chosen several well-known measures [4]. The spatial accuracy (AC), the precision (PR) and the recall (RC) are considered in this work in make the comparison. A value in the range $[0, 1]$, where higher
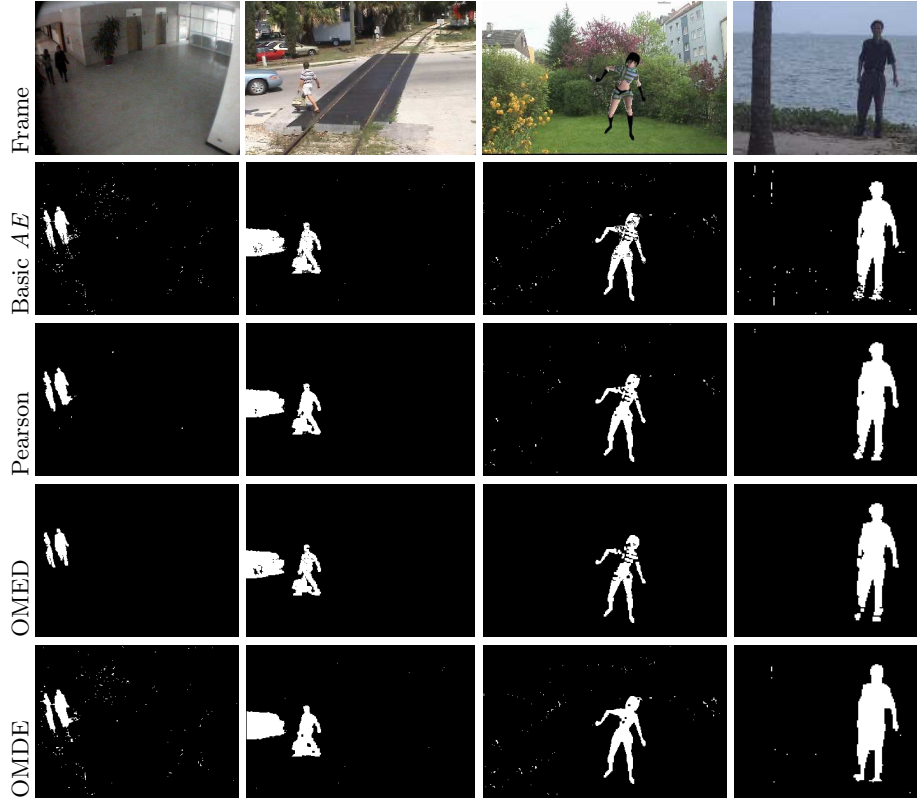
Fig. 3. Another set of experimental results. An interior scene of a hall (*IP1*) is observed in the first column (frame 116) while a scene of a level crossing without barriers is shown in the second column (*LC*, frame 389). Third and forth columns analyze the sequences *V4*, synthetically generated, and *WS*, where the movement of the waves of the sea, with frames 815 and 1624, respectively.

is better, is provided by each measure. We also consider True positives (TP), True negatives (TN), False negatives (FN) and false positives (FP) rates. A good overall evaluation of the performance of a given method is offered by AC, while PR must be considered against RC. Each measure can be defined as follow:

$$AC = \frac{TP}{TP + FN + FP} \tag{12}$$

$$RC = \frac{TP}{TP + FN} \tag{13}$$

$$PR = \frac{TP}{TP + FP} \tag{14}$$

Table 1. Quantitative assessment using the measure AC. The mean and the standard deviation are shown after the analysis of each set of data, where best result is highlighted in **bold**. The last row indicates the average performance of each method, where the best method obtains one point, the second two points, and so on. The fewer score the better the postprocessing on average. Rows with a star indicate that the results are statistically significant.

	Basic	Pearson	OMED	OMDE
Campus (CAM)	0.594±0.168	0.725±0.093	**0.735±0.075**	0.607±0.182
*Meeting Room (MR)	0.775±0.060	0.803±0.058	0.752±0.072	**0.866±0.036**
Subway Station (SS)	0.458±0.124	0.462±0.128	0.424±0.140	**0.511±0.144**
Fountain (FT)	0.407±0.150	0.435±0.143	0.463±0.086	**0.480±0.194**
IPPRData1 ($IP1$)	0.536±0.169	0.620±0.170	**0.632±0.183**	0.539±0.181
*IPPRData2 ($IP2$)	0.451±0.134	0.469±0.140	**0.628±0.133**	0.459±0.137
IPPRData3 ($IP3$)	0.606±0.174	**0.620±0.172**	0.603±0.194	0.619±0.171
*Level Crossing (LC)	0.878±0.035	**0.897±0.030**	0.879±0.035	0.883±0.033
Corridor (OC)	0.709±0.037	0.712±0.038	**0.723±0.039**	0.706±0.038
Video2 ($V2$)	**0.920±0.019**	0.919±0.023	0.910±0.024	0.916±0.027
Video4 ($V4$)	0.682±0.111	0.722±0.112	**0.723±0.115**	0.714±0.128
WaterSurface (WS)	0.870±0.023	**0.902±0.025**	0.890±0.026	0.899±0.022
Puntuación	41	**22**	29	28

In general, and for all the tested sequences, we can indicate that *Pearson* is the most suitable method together with *OMED* for scenes with false positives, while those with false negatives is not as effective as *OMDE*, but it is quite competitive. This information is shown in Tables 1, 2 and 3.

In the first of them (Table 1) we can assume the comments from a visual point of view of the qualitative comparison from Figures 2 and 3. We could consider *OMED* as the best suitable method because it is the best in five of the twelve analyzed sequences, while *Pearson* and *OMDE* are better in three videos each one. But, because of the fact that *OMED* is the worst in quality terms in several sequences (see the performance for the *MR*, *IP3* and *V2* scenes) we have incorporated a ranking index which scores each method in ascending order of its performance. Thus, the winner method for a sequence will obtain one point, the second two points, the third three points and the forth four points. The last row of the Table 1 represents the sum of these scores for all the sequences. From these results, we can conclude that *Pearson* is the best method on average because, although it is not always the best, it is the second best method in most of the sequences (see the ranking of the methods in Table 3). Therefore, this approach could be viewed as the most stable method regardless of the type of sequence. It is interesting to observe that, except one case and with low margin (sequence *V2*), the *Pearson* correlation filtering always improves the result produced by the *Basic* segmentation algorithm.

Another non-trivial advantage of the postprocessing with *Pearson* is that it is not necessary to adjust any kind of parameters, except those required by the segmentation algorithm (ε_0 for AE in this experiments). As it is observed in

Table 2. Best configuration for each method and each sequence.

Seq.	Metd.	Accuracy	*Precision*	*Recall*	Parameters
CAM	Basic	0.594±0.168	0.65±0.19	0.88±0.06	ε_0=0.01
	Pearson	0.725±0.093	0.79±0.11	0.90±0.05	ε_0=0.005
	OMED	**0.735±0.076**	0.85±0.09	0.85±0.08	ε_0=0.001, #$\mathcal{ED}$=1
	OMDE	0.607±0.182	0.64±0.20	0.93±0.05	ε_0=0.01, #$\mathcal{DE}$=1
MR	Basic	0.775±0.060	0.96±0.03	0.80±0.06	ε_0=0.0005
	Pearson	0.803±0.058	0.96±0.03	0.83±0.06	ε_0=0.0005
	OMED	0.752±0.072	0.98±0.02	0.76±0.08	ε_0=0.0005, #$\mathcal{ED}$=1
	OMDE	**0.866±0.036**	0.95±0.03	0.91±0.02	ε_0=0.001, #$\mathcal{DE}$=3
SS	Basic	0.458±0.124	0.70±0.18	0.59±0.16	ε_0=0.01
	Pearson	0.462±0.128	0.72±0.19	0.59±0.17	ε_0=0.01
	OMED	0.424±0.140	0.68±0.17	0.54±0.17	ε_0=0.1, #$\mathcal{ED}$=1
	OMDE	**0.511±0.144**	0.65±0.19	0.72±0.14	ε_0=0.01, #$\mathcal{DE}$=2
FT	Basic	0.407±0.150	0.52±0.18	0.64±0.11	ε_0=0.005
	Pearson	0.435±0.143	0.52±0.17	0.72±0.10	ε_0=0.001
	OMED	0.463±0.086	0.71±0.08	0.57±0.10	ε_0=0.0001, #$\mathcal{ED}$=1
	OMDE	**0.480±0.194**	0.51±0.20	0.87±0.10	ε_0=0.01, #$\mathcal{DE}$=2
IP1	Basic	0.536±0.169	0.66±0.18	0.73±0.17	ε_0=0.01
	Pearson	0.620±0.170	0.81±0.11	0.72±0.19	ε_0=0.005
	OMED	**0.632±0.183**	0.93±0.06	0.67±0.20	ε_0=0.01, #$\mathcal{ED}$=1
	OMDE	0.539±0.181	0.63±0.19	0.78±0.17	ε_0=0.01, #$\mathcal{DE}$=1
IP2	Basic	0.451±0.134	0.53±0.18	0.76±0.08	ε_0=0.01
	Pearson	0.469±0.140	0.56±0.19	0.77±0.08	ε_0=0.005
	OMED	**0.628±0.133**	0.77±0.15	0.78±0.12	ε_0=0.01, #$\mathcal{ED}$=2
	OMDE	0.459±0.137	0.52±0.17	0.81±0.08	ε_0=0.01, #$\mathcal{DE}$=1
IP3	Basic	0.606±0.174	0.70±0.17	0.81±0.14	ε_0=0.05
	Pearson	**0.619±0.172**	0.71±0.17	0.83±0.14	ε_0=0.05
	OMED	0.603±0.194	0.79±0.14	0.72±0.20	ε_0=0.05, #$\mathcal{ED}$=1
	OMDE	0.619±0.171	0.68±0.17	0.87±0.11	ε_0=0.05, #$\mathcal{DE}$=1
LC	Basic	0.879±0.035	0.91±0.03	0.96±0.01	ε_0=0.01
	Pearson	**0.897±0.030**	0.91±0.03	0.98±0.01	ε_0=0.01
	OMED	0.879±0.034	0.93±0.02	0.94±0.02	ε_0=0.01, #$\mathcal{ED}$=1
	OMDE	0.883±0.033	0.89±0.03	0.99±0.01	ε_0=0.01, #$\mathcal{DE}$=1
OC	Basic	0.709±0.037	0.75±0.02	0.92±0.04	ε_0=0.01
	Pearson	0.712±0.038	0.75±0.02	0.92±0.04	ε_0=0.01
	OMED	**0.723±0.039**	0.81±0.02	0.87±0.04	ε_0=0.01, #$\mathcal{ED}$=3
	OMDE	0.706±0.038	0.74±0.02	0.94±0.04	ε_0=0.01, #$\mathcal{DE}$=1
V2	Basic	**0.920±0.019**	0.94±0.01	0.97±0.02	ε_0=0.001
	Pearson	0.919±0.023	0.95±0.01	0.97±0.02	ε_0=0.0005
	OMED	0.910±0.024	0.94±0.01	0.96±0.02	ε_0=0.0001, #$\mathcal{ED}$=1
	OMDE	0.916±0.027	0.94±0.01	0.97±0.03	ε_0=0.005, #$\mathcal{DE}$=1
V4	Basic	0.682±0.111	0.75±0.13	0.88±0.03	ε_0=0.005
	Pearson	0.722±0.112	0.79±0.13	0.90±0.03	ε_0=0.005
	OMED	**0.723±0.115**	0.82±0.15	0.87±0.03	ε_0=0.001, #$\mathcal{ED}$=1
	OMDE	0.714±0.128	0.74±0.14	0.96±0.02	ε_0=0.005, #$\mathcal{DE}$=1
WS	Basic	0.870±0.023	0.95±0.02	0.91±0.02	ε_0=0.001
	Pearson	**0.902±0.025**	0.97±0.01	0.93±0.02	ε_0=0.0005
	OMED	0.890±0.026	0.98±0.01	0.91±0.02	ε_0=0.0001, #$\mathcal{ED}$=1
	OMDE	0.899±0.022	0.96±0.02	0.93±0.02	ε_0=0.005, #$\mathcal{DE}$=2

Table 3. *P-values* after the application of the *Student's t test* between the best method and the remaining competitors. If all the *p-values* of the same sequence are lower than 0.05, the result of the best technique is regarded as significant, whose name in the table is formatted in **bold** style.

Seq.	Best Metd.	Competitor 1		Competitor 2		Competitor 3	
CAM	OMED	Pearson	0.7079	OMDE	6.10e-003	Basic	1.43e-003
MR	**OMDE**	Pearson	**2.06e-004**	Basic	**1.10e-006**	OMED	**2.05e-007**
SS	OMDE	Pearson	0.2647	Basic	0.2172	OMED	0.0606
FT	OMDE	OMED	0.7237	Pearson	0.4092	Basic	0.1884
IP1	OMED	Pearson	0.5246	OMDE	2.93e-006	Basic	5.38e-007
IP2	**OMED**	Pearson	**0.00e+000**	OMDE	**0.00e+000**	Basic	**0.00e+000**
IP3	Pearson	OMDE	0.9538	Basic	0.4230	OMED	0.3761
LC	**Pearson**	OMDE	**6.25e-005**	OMED	**1.00e-006**	Basic	**3.75e-007**
OC	OMED	Pearson	0.4916	Basic	0.3821	OMDE	0.2957
V2	Basic	Pearson	0.3784	OMDE	4.96e-003	OMED	8.45e-013
V4	OMED	Pearson	0.9188	OMDE	0.3522	Basic	5.65e-006
WS	Pearson	OMDE	0.6302	OMED	0.1377	Basic	1.37e-004

Table 2, the number of iterations of *erosion - dilation* or *dilation - erosion* for the morphological operators is not always the same, so it is needed to estimate the optimal value to avoid performing the search exhaustively.

Finally, in Table 3 a statistical significance test employing the *Student's t test* is carried out. A two-tailed test between the best method and the remaining alternatives is computed, where the result is considered statistically significant if the *p-value* is lower than 0.05 for all the comparisons in the same sequence.

Consequently, the obtained results are quite similar, since the improvement of the method is significant in only three of the tested sequences. Nevertheless, if we observe, for example, the *p-values* of the *CAM* sequence, we can see that *OMED* is not significantly better than *Pearson*, but, in fact, it is regarding *OMDE* and the proposal without any postprocessing method (*Basic*). In addition, it can be observed that *Pearson* is the second method (*Competitor 1* in Table 3) in practically all the sequences where it is not the winner.

4 Conclusion

We can conclude that the postprocessing method, in general, is necessary because it substantially improves the result obtained by the segmentation algorithms. In addition, it has been observed that *Pearson* is the best method on average for all the tested sequences, although with no statistically significant evidence.

Acknowledgments

References

1. Burdick, H.E.: Digital imaging - theory and applications. McGraw-Hill (1997)
2. Gamarra, M., Zurek, E., San-juan, H., Eng, S., Norte, U.: A Study of Image Analysis Algorithms for Segmentation , Feature Extraction and Classification of Cells. Journal of Information Systems Engineering & Management 2(4) (2017)
3. Kastrinaki, V., Zervakis, M., Kalaitzakis, K.: A survey of video processing techniques for traffic applications. Image and Vision Computing 21(4), 359 – 381 (2003)
4. Li, L., Huang, W., Gu, I.Y.H., Tian, Q.: Statistical modeling of complex backgrounds for foreground object detection. IEEE Transactions on Image Processing 13(11), 1459–1472 (2004)
5. López-Rubio, E., Luque-Baena, R.M.: Stochastic approximation for background modelling. Computer Vision and Image Understanding 115(6), 735–749 (2011)
6. Lopez-Rubio, E., Luque-Baena, R.M., Dominguez, E.: Foreground detection in video sequences with probabilistic self-organizing maps. International Journal of Neural Systems 21(3), 225 – 246 (2011)
7. Masood, S., Sharif, M., Masood, A., Yasmin, M., Raza, M.: A Survey on Medical Image Segmentation. Current Medical Imaging Reviews 11(1), 3–14 (2015)
8. Nilakant, R., Menon, H.P., Vikram, K.: A Survey on Advanced Segmentation Techniques for Brain MRI Image Segmentation. International Journal on Advanced Science, Engineering and Information Technology 7(4), 1448–1456 (2017)
9. Palomo, E.J., Domínguez, E., Luque-Baena, R.M., Muñoz, J.: Image Compression and Video Segmentation Using Hierarchical Self-Organization. Neural Processing Letters 37(1), 69–87 (2013)
10. R.M. Luque, E. Domínguez, E.P., Muñoz, J.: A neural network approach for video object segmentation in traffic surveillance. In: Springer (ed.) Lecture Notes in Computer Science. vol. 5112, pp. 151 – 158 (2008)
11. Rodgers, J.L., Nicewander, W.A.: Thirteen ways to look at the correlation coefficient. The American Statistician 42(1), 59–66 (1988)