IMPROVING BLOTCH DETECTION IN OLD FILMS BY A PREPROCESSING STEP BASED ON OUTLIER STATISTICAL TEST

Heyfa Ammar-Badri and Amel Benazza-Benyahia

URISA, SUP'COM, Université de Carthage
Cité Technologique des Communications de Tunis, El Ghazala, Ariana, 2083, Tunisia
phone: + (216) 857000 ext. 1027, fax: + (216) 856829, email: benazza.amel@supcom.rnu.tn

ABSTRACT

This work deals with the detection of blotches which are specific degradations commonly affecting old movies. The novelty of our approach consists in integrating a preprocessing step into any conventional blotch detector operating in the spatial domain. This preprocessing step aims at outputting candidate regions that may be degraded. Its originality relies on the fact that it is based on a statistical outlier detection test coupled with a compensation motion method. Experimental results carried out on real old movies, show that this preprocessing highly improves the detection performances of the conventional Simplified Rank Ordered Differences (SROD) detector and the detector based on the Auto-Regressive (AR) model.

1. INTRODUCTION

Films have the important role of tracing human civilizations and historical events among the years. They represent an invaluable data inherited from generation to generation. Unfortunately, an important amount of old films has been lost due to bad storage conditions and handlings. Many efforts were firstly devoted by archivists and researchers to save this heritage by manually restoring some degraded films. Therefore, the huge amount of data to be restored makes the manual restoration procedure infeasible. This is the reason why an automatic and digital restoration carried out on digitized version of old films is advocated. Generally, a generic restoration method is usually composed of two steps : contaminated areas are detected before to be corrected. In our work, we are interested in detecting special artifacts called blotches. These artifacts are caused by imperfect archiving conditions and essentially originate from the loss of gelatin and the presence of dust and dirt on the surface of the film. Blotches randomly occur in the sequence as dark or bright spots with arbitrary shape, size and brightness. Given the randomness of blotches, they hardly appear in successive frames with the same location, shape and size. Consequently, they cause temporal discontinuities in the sequence. This property is exploited in almost all proposed blotch detection methods as [1, 2]. In this paper, we propose to design a preprocessing step which detects candidate regions that may be blotched. In this way, the preprocessing facilitate the detection task of any blotch detector operating in the spatial domain. The originality of our approach relies on the fact that blotches are considered as local illumination variations between successive frames. To the best of our knowledge, this property has never been exploited by the reported blotch detectors. Furthermore, an additional novelty consists in resorting to statistical outlier tests to detect the candidate areas. This paper is organized as follows. In Section 2, we give a brief review of the main blotch detection methods. Then, the new method we propose is detailed in Section 3. Finally, in Section 4, we provide the performances of our method on both artificial and real degraded sequences and in Section 5, some conclusions are drawn.

2. A BRIEF STATE OF ART

Most of blotch detectors and that operate in the spatial domain consider blotches as discontinuities between the degraded frame $I^{(k)}$ of size $L \times L$ and its previous and subsequent frames $I^{(k-1)}$ and $I^{(k+1)}$ respectively. Hence, as shown by the block-diagram depicted in Figure 1, a motion estimation and compensation step are firstly performed to reduce discontinuities due to object motions, and subsequently, motion compensated frames $I_{\rm mc}^{(k-1)}$ and $I_{\rm mc}^{(k+1)}$ are generated. Secondly, corrupted pixels are found by comparing the differences between their intensities and those of their homologous in $I_{\rm mc}^{(k-1)}$ and $I_{\rm mc}^{(k+1)}$ according to a given threshold. For instance, Nadenau and Mitra [3] have proposed to compare the Rank Ordered Differences (ROD) between the current pixel intensity value $I^{(k)}(\mathbf{x})$ and those of the six reference spatio-temporal neighbors $P_1, \ldots P_6$ shown in Figure 2.

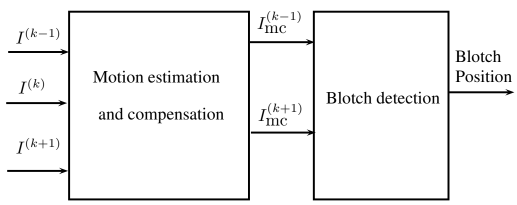


Figure 1: Block-diagram of conventional blotch detector.

Let z_m be the intensity value of the reference pixels ordered by rank $z_1 < z_2 < z_3 < z_4 < z_5 < z_6$. Then, the three order differences $ROD(\mathbf{x}, l)$ with $l = 1, 2, 3$ are computed as follows:

$$ROD(\mathbf{x}, l) = \begin{cases} z_l - I^{(k)}(\mathbf{x}) \text{ if } I^{(k)}(\mathbf{x}) \leq (z_3 + z_4)/2 \\ I^{(k)}(\mathbf{x}) - z_{7-l} \text{ if } I^{(k)}(\mathbf{x}) > (z_3 + z_4)/2. \end{cases} \quad (1)$$

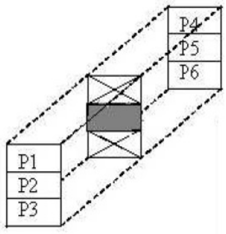


Figure 2: Reference pixels P_k used by the ROD detector.

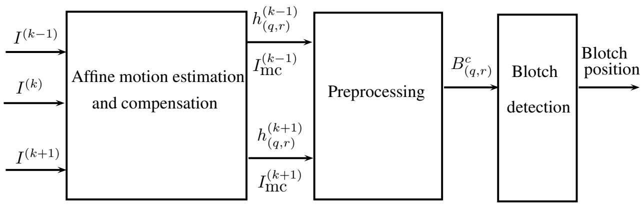


Figure 3: Block-diagram of the proposed detector.

The pixel $\mathbf{x}$ is considered as blotched if at least one of the three ordered differences $ROD(\mathbf{x}, l)$ exceeds a given threshold. Later, Van Roosmalen [2] has simplified and improved this method by considering the Simplified Rank Ordered Differences $SROD$ defined by:

$$SROD(\mathbf{x}) = \max\{0, z_1 - I^{(k)}(\mathbf{x}), I^{(k)}(\mathbf{x}) - z_6\}. \quad (2)$$

Pixels whose $SROD$ values are greater than a given threshold are judged as contaminated. Another detector based on an Auto-Regressive (AR) model was proposed in [1]. The principle is to assume that uncorrupted frames follow an AR model of order P:

$$I^{(k)}(\mathbf{x}) = \hat{I}^{(k)}(\mathbf{x}) + e^{(k-1)}(\mathbf{x}) + e^{(k+1)}(\mathbf{x}) \quad (3)$$

where

$$\begin{aligned}\hat{I}^{(k)}(\mathbf{x}) &= \textstyle\sum_{l=1}^{P} a_l^{(k-1)} I_{\mathrm{mc}}^{(k-1)}(\mathbf{x} + \mathbf{q}_l^{(k-1)}) \\ &+ \textstyle\sum_{l=1}^{P} a_l^{(k+1)} I_{\mathrm{mc}}^{(k+1)}(\mathbf{x} + \mathbf{q}_l^{(k+1)})\end{aligned} \quad (4)$$

where the $a_l^{(k-1)}$ and $a_l^{(k+1)}$ $(l = 1, \ldots, P)$ are the $2P$ AR model coefficients estimated from a support of pixels belonging to $I_{\mathrm{mc}}^{(k-1)}$ and $I_{\mathrm{mc}}^{(k+1)}$, and the $\mathbf{q}_l$ denotes the relative positions of these reference pixels with respect to the underlying pixel. A pixel at the location $\mathbf{x}$ is judged as corrupted if $(e^{(k-1)}(\mathbf{x}))^2 > T$ and $(e^{(k+1)}(\mathbf{x}))^2 > T$ where T is a given threshold.

The main drawback of these methods is the high false alarms rate they generate. This may be due to the unreliability of the motion estimation in presence of blotches. In this paper, our rationale is to consider a blotch as a local brightness variation between the degraded frame $I^{(k)}$ and its previous and subsequent frames. Consequently, we adopt an affine Motion Estimation (ME) technique which was found to be robust to the illumination variations [5] and we propose to exploit the information provided by the ME step in order to define potential candidate of bloched regions.

3. PROPOSED APPROACH

In order to improve the performances of the conventional detectors depicted in Figure 1, we add a preprocessing step driven by the affine ME step to select candidate blocks before the blotch detection is carried out. The block diagram of the proposed detector is depicted in Figure 3 and it is detailed in the sequel.

3.1 Robust motion estimation

By considering that a blotch is a local temporal illumination variation within the sequence, it is necessary to resort to a motion model that handles the temporal brightness variations. In this respect, the affine model [5] is widely used:

$$I^{(k)}(\mathbf{x}) = h^{(k-1)}(\mathbf{x}) I^{(k-1)}(\mathbf{x} - \mathbf{d}^{(k-1)}(\mathbf{x})) + b^{(k-1)}(\mathbf{x}) \quad (5)$$

where $h^{(k-1)}(\mathbf{x})$ is the illumination prediction coefficient, and $b^{(k-1)}(\mathbf{x})$ is an additive noise. $\mathbf{d}^{(k-1)}$ is the displacement vector of pixel $\mathbf{x}$ between $k-1$ and k. Generally, this model is considered to be locally valid. More precisely, it is used to split $I^{(k)}$ into disjoint blocks $B_{(q,r)}$ of size $\ell \times \ell$ where (q, r) denotes the spatial position (q, r) of the leftmost top pixel in the block. The motion vector and the illumination prediction coefficient are assumed to be constant for all the pixels within $B_{(q,r)}$. Furthermore, a search area $\mathcal{S}$ of size $(\ell + 2p) \times (\ell + 2p)$ is defined in $I^{(k-1)}$ where p is the maximum displacement allowed horizontally and vertically. Then, the block within $\mathcal{S}$ that matches $B_{(q,r)}$ according to a given matching criterion $\mathcal{C}$ is searched for. Very often, the mean square error is retained for $\mathcal{C}$:

$$\begin{aligned}&\mathcal{C}(\mathbf{d}, h_{(q,r)}^{(k-1)}, b_{(q,r)}^{(k-1)}) = \\ &\textstyle\sum_{\mathbf{x} \in B_{(q,r)}} (I^{(k)}(\mathbf{x}) - h_{(q,r)}^{(k-1)} I^{(k-1)}(\mathbf{x} - \mathbf{d}) - b_{(q,r)}^{(k-1)})^2.\end{aligned} \quad (6)$$

In this case, the multiplicative and additive coefficients are firstly estimated for each displacement $\mathbf{d}$ (or equivalently each candidate block in $\mathcal{S}$):

$$h_{(q,r)}^{(k-1)}(\mathbf{d}) = \frac{\mathrm{Cov}[I^{(k)}(\mathbf{x}) I^{(k-1)}(\mathbf{x} - \mathbf{d})]}{\sigma^2_{I^{(k-1)}(\mathbf{x}-\mathbf{d})_{\mathbf{x} \in B_{(q,r)}}}}$$

$$b_{(q,r)}^{(k-1)}(\mathbf{d}) = m_{I^{(k)}(\mathbf{x})_{\mathbf{x} \in B_{(q,r)}}} - h_{(q,r)}^{(k-1)} m_{I^{(k-1)}(\mathbf{x}-\mathbf{d})_{\mathbf{x} \in B_{(q,r)}}}$$

where σ^2 is the variance of $I^{(k-1)}$, Cov denotes the covariance function, and $m_{I^{(k)}(\mathbf{x})_{\mathbf{x} \in B_{(q,r)}}}$ is the mean value of the pixels within the block $B_{(q,r)}$. The most suitable motion vector $\mathbf{d}_{(q,r)}^{(k-1)}$ associated to the block $B_{(q,r)}$ is then easily deduced by:

$$\mathbf{d}_{(q,r)}^{(k-1)} = \arg\min_{\mathbf{d} \in \mathcal{S}} \mathcal{C}(\mathbf{d}, h_{(q,r)}^{(k-1)}(\mathbf{d}), b_{(q,r)}^{(k-1)}(\mathbf{d})) \quad (7)$$

Consequently, it is straightforward to generate the motion compensated frame $I_{mc}^{(k-1)}$. A similar procedure with the reference frame $I^{(k+1)}$ allows also to compute a motion vector $\mathbf{d}_{(q,r)}^{(k+1)}$ and a prediction coefficient $h_{(q,r)}^{(k+1)}$ and the motion compensated frame $I_{mc}^{(k+1)}$.

3.2 Proposed preprocessing

For all the blocks $B_{(q,r)}$ in $I^{(k)}$, we obtain a set $\mathcal{H}^{(k-1)} = \{h_{(q,r)}^{(k-1)}\}_{(q,r)}$. Corrupted blocks have illumination coefficients that take isolated values among the distribution of all coefficients of the remaining blocks since there is no exact match between corrupted block and the reference one. An example of the coefficient distribution is given in Figure 4 where the corrupted block is shown in red. This has motivated us to consider that the values

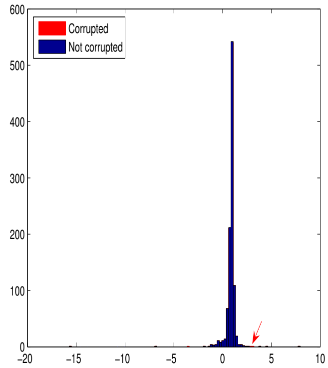


Figure 4: An example of the distribution of $\{h_{(q,r)}^{(k-1)}\}_{(q,r)}$ values in the cases of non blotched and blotched image

associated to contaminated blocks correspond to outliers of the set $\mathcal{H}^{(k-1)}$. Consequently, a detection of candidate blotched blocks reduces to detect outlier values of $\mathcal{H}^{(k-1)}$. In this respect, we resort to a statistical test to find these multiple outliers. More precisely, we retain the statistical Minimum Covariance Determinant (MCD) test for its efficiency and relatively low complexity [4]. Likewise the majority of the statistical tests, the MCD test requires that the underlying data within the sample set are normally distributed. Therefore, we have to proceed to a Gaussianization of the set $\mathcal{H}^{(k-1)}$ by applying the Box and Cox power transform [6] to map $\mathcal{H}^{(k-1)}$ into a set $\tilde{\mathcal{H}}^{(k-1)}$ of realizations of a Gaussian random variable

$$\forall (q,r) \quad \tilde{h}_{(q,r)}^{(k-1)} = \frac{(h_{(q,r)}^{(k-1)})^{\lambda} - 1}{\lambda} \tag{8}$$

where λ is a parameter which can be estimated according to the maximum likelihood criterion. Note that the Box and Cox transform requires the positivity of the observations and consequently, $\mathcal{H}^{(k-1)}$ is initially translated to obtain positive values prior to apply the transform.

The key idea of the MCD test is to search for the most concentrated subset $\tilde{H}_c^{(k-1)}$ of size $r = \lfloor (1-\epsilon)N \rfloor$ among some predefined r-subsets where N is the size of $\tilde{\mathcal{H}}^{(k-1)}$, ϵ corresponds to the smallest amount of outlier contamination that can have an arbitrarily large effect on the mean and variance estimates. The subset $\tilde{H}_c^{(k-1)}$ is determined by firstly, ordering the observations in increasing order: $\tilde{h}_{(1)}^{(k-1)} \leq \tilde{h}_{(2)}^{(k-1)} \leq \ldots \leq \tilde{h}_{(N)}^{(k-1)}$. Secondly, for $i = 1, \ldots, N-r+1$, contiguous subsets $R_i^{(k-1)} = \{\tilde{h}_{(i)}^{(k-1)}, \ldots, \tilde{h}_{(r+i-1)}^{(k-1)}\}$ of size r are defined. The estimators $\hat{\mu}_i$ and $\hat{\sigma}_i^2$ of respectively the mean and the variance of each subset $R_i^{(k-1)}$ are computed:

$$\hat{\mu}_i = \frac{1}{r}\sum_{n=i}^{r+i-1} \tilde{h}_{(n)}^{(k-1)}, \quad \hat{\sigma}_i^2 = \frac{1}{r}\sum_{n=i}^{r+i-1} (\tilde{h}_{(n)}^{(k-1)} - \hat{\mu}_i)^2. \tag{9}$$

Let i^* be the index of the subset having the lowest variance. The robust Euclidean distance $d(\tilde{h}_n^{(k-1)}; \hat{\mu}_{i^*}, \hat{\sigma}_{i^*})$ allows to measure how far away any observation $\tilde{h}_n^{(k-1)}$ from the center $\hat{\mu}_{i^*}$ of $\tilde{R}_{i^*}$ relative to its size:

$$\forall n = 1, \ldots, N, \quad d(\tilde{h}_n^{(k-1)}; \hat{\mu}_{i^*}, \hat{\sigma}_{i^*}) = |\tilde{h}_n^{(k-1)} - \hat{\mu}_{i^*}| / \hat{\sigma}_{i^*}. \tag{10}$$

Given that the squared distance d^2 has a χ_1^2 distribution with one freedom degree, an observation $\tilde{h}_n^{(k-1)}$ is considered as an outlier if $d^2(\tilde{h}_n^{(k-1)}; \hat{\mu}_{i^*}, \hat{\sigma}_{i^*})$ exceeds a threshold easily derived from a two-tailed test for a given confidence value ε.

The same procedure is applied to the set $\mathcal{H}^{(k+1)}$ in order to detect suspicious blocks in the forward direction. Finally, a block $B_{(q,r)}^c$ is judged as a blotched candidate if both, the corresponding $\tilde{h}_{(q,r)}^{(k-1)}$ and $\tilde{h}_{(q,r)}^{(k+1)}$ are detected as outliers. Note that these two conditions should be simultaneously satisfied because of the temporal discontinuity of the blotches.

This preprocessing step is a blotch detector by itself. Therefore, it detects contaminated blocks which may be parts of bloched areas whose shape are not necessarily rectangular. For this reason, a refinement step is necessary and consists of applying any heuristic detector such as described in Section 2, on the candidate blocks.

4. EXPERIMENTAL RESULTS

The proposed preprocessing step should be evaluated in order to firstly adjust the optimal values of the parameters on which it depends, and secondly, measure its contribution within the final blotch detector shown in Figure 3. Thus, two rounds of experiments are performed. The first round consists in the evaluation of the preprocessing step according to its own parameters: ε is varying from 0.5 to 1 by step of 0.04, for different block sizes ℓ and different values of p. The second round of experiments aims to compare the performances of the whole detector coupled with the *SROD* detector and the detector based on the 3D AR model. As suggested by Roosmalen in [2], the retained support for the 3D AR based detector is depicted in Figure 5.

Tests are carried out on a set of 20 frames extracted from the "La Bataille du pacifique" sequence where 9 frames are artificially degraded. The artificial degradation consists of adding real blotches by using ground truth (GT) dirt maps obtained thanks to a special

infrared-film scanner. These maps typically show dirt as darker areas set against a lighter background. The binary ground truth masks are generated through a thresholding of the related infrared images. The threshold values are manually set in order to find the binary GT mask $M_{\mathrm{g}}^{(k)}$ as close as possible to the human perception of these defects. A set of 5 real degraded frames extracted from the "Gallipoli" sequence are also used in the experiments. The GT masks showing blotches positions affecting the real sequence are also obtained manually according to the human perception of the degradations. For all the experiments, the detection performances are measured in terms of good detection rate $P_c = \frac{|M^{(k)} \bigcap M_{\mathrm{g}}^{(k)}|}{|M_{\mathrm{g}}^{(k)}|}$, false alarm rate $P_f = \frac{|M^{(k)} \bigcap \overline{M_{\mathrm{g}}^{(k)}}|}{|\overline{M_{\mathrm{g}}^{(k)}}|}$ and error rate $P_e = P_1(1 - P_c) + (1 - P_1)P_f$, where $P_1 = \frac{|M_{\mathrm{g}}^{(k)}|}{L^2}$ and $|\cdot|$ denotes the number of ones contained in the considered sets. Receiver Operator Characteristic (ROC) curves are used for the evaluation. It is worth noting that since the experiments are applied to an extract of the sequence, several values of P_f and P_c are obtained. ROC curves depicted in Figure 6 plot the mean value of P_c among all values of P_c obtained for all the frames by varying ε, versus the mean value of P_f obtained for all P_f related to all the frames. This figure shows good performances of the proposed preprocessing step, particularly when $\ell = 8$ and $p = 5$. The best results are obtained when $\varepsilon = 0.88$ for the "La Bataille du pacifique " extract, and $\varepsilon = 0.8$ for the "Gallipoli" sequence. We also plot in Figure 7 the mean value of the error rates P_e for each value of ε.
For the remaining tests, we choose $\varepsilon = 0.8$ as it is less tolerant to outliers, and as false positives will be omitted by the refinement step of the detector.

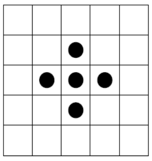


Figure 5: Pixels used as support from reference frames for the 3D AR detector.

To emphasize the importance of the proposed preprocessing, the second round of tests aims at integrating it prior to the *SROD* detector, and then prior to the 3D AR model based detector. Figure 8 shows the obtained results on two artificially degraded frames respectively extracted from the "Gallipoli " and "Tierce " sequences. We can easily note the substantial gain drawn from the proposed preprocessing for both retained detectors. Thanks to the block candidate selection step, the false alarm rates decrease at the same rate of correct detections.
A qualitative evaluation is performed on a real degraded frame extracted from the "Gallipoli" sequence (frame #195). Figure 4 allows a visual inspection and confirms the efficiency of the proposed preprocessing step compared to the classical *SROD* detector (we have used the same value of threshold for both detectors). It is worth pointing out that there is also a real illumination variation between the degraded frame and its subsequent one which explains the high false alarms rate obtained by the *SROD* detector.

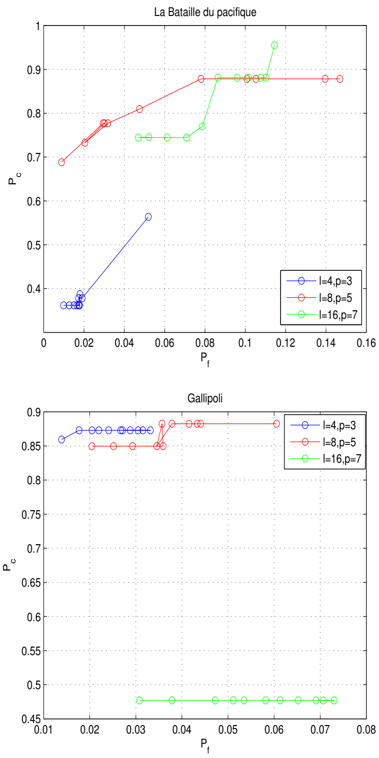


Figure 6: ROC curves evaluating the preprocessing step by varying ε for different values of (ℓ, p). Top: "La Bataille du pacifique" extract. Bottom: the "Gallipoli " extract.

5. CONCLUSION

We have proposed in this paper an efficient preprocessing step of any blotch detector operating in the spatial domain. Its novelty relies on the hypothesis that a blotch could be viewed as a local temporal illumination variation. Experimental results have shown the efficiency of the preprocessing step in reducing false alarms rate, and the robustness of the proposed approach to the illumination variation within the sequence of state-of-art bloch detector.

REFERENCES

[1] A. C. Kokaram, *Motion picture restoration: digital algorithms for artefact suppression in degraded motion picture film and video.* Springer Verlag, 1998.

[2] P. M. B. Van Roosmalen, "Restoration of archived film and video, " *PhD Thesis*, Delft University of Technology, The Netherlands. 1999.

[3] M. J. Nadenau, S. K. Mitra, "Blotch and scratch detection in image sequences based on rank ordered differences," in *Internat. Workshop on Time Vary-*

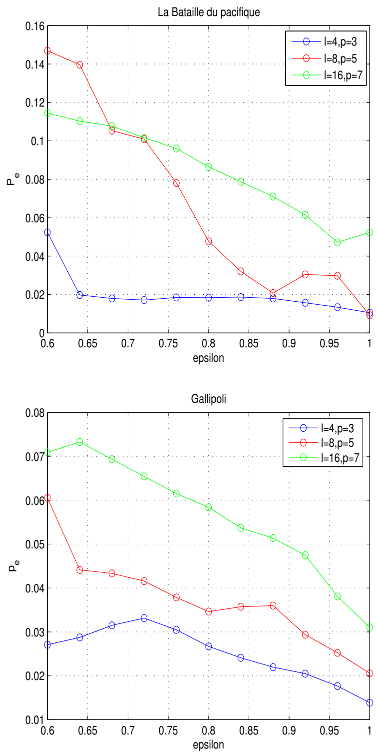


Figure 7: Error rates of the preprocessing step by varying ε, for different values of (ℓ, p). Top: "La Bataille du pacifique " extract. Bottom: the "Gallipoli" extract.

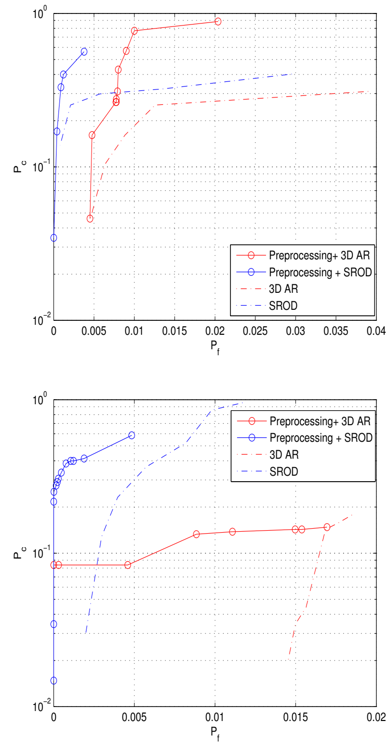


Figure 8: ROC curves related to artificially degraded frames extracted from the "Gallipoli" (Top) and "Tierce" (Bottom) sequences.

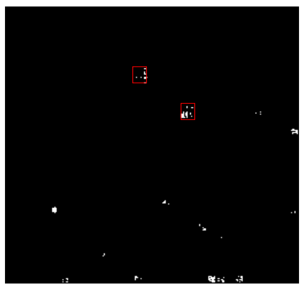
ing Image Processing and Moving Object Recognition, Firenze, Italy, pp. 1–7, 1996.

[4] P. J. Rousseeuw and K. V. Driessen, "A fast algorithm for the minimum covariance determinant estimator," in *American Statistical Association and the American Society for Quality*, vol. 41, no. 3, pp. 212–223, 1999.

[5] H. Nicolas and C. Labit, "Motion and illumination variation estimation using a hierarchy of models: Application to image sequence coding," *Journal of Visual Communication and Image Representation*, vol. 6, pp. 303–316, Dec. 1995.

[6] E. P. G. Box, D. R. Cox "An analysis of transformations," *Journal of Royal Statistical Society, Series B (Methodological)*, vol. 26, no. 2, pp. 211–252, 1964.

[7] J. Villasenor, B. Belzer, J. Liao, "Wavelet filter evaluation for image compression," in *IEEE Trans. on on Image Processing*, vol. 4, no. 8, pp. 1053–1060, Aug. 1995



Mask obtained by the proposed detector

Mask obtained by the SROD detector

Figure 9: Evaluation on a real degraded frame. white pixels are considered as blotches