

Trainable blotch detection on high resolution archive films minimizing the human interaction

Attila Licsár · Tamás Szirányi · László Czúni

Received: 14 July 2006 / Accepted: 30 March 2007 / Published online: 9 November 2007
© Springer-Verlag 2007

Abstract Film archives are continuously in need of automatic restoration tools to accelerate the correction of film artifacts and to decrease the costs. Blotches are a common type of film degradation and their correction needs a lot of manual interaction in traditional systems due to high false detection rates and the huge amount of data of high resolution images. Blotch detectors need reliable motion estimation to avoid the false detection of uncorrupted regions. In case of erroneous detection, usually an operator has to remove the false alarms manually, which significantly decreases the efficiency of the restoration process. To reduce manual intervention, we developed a two-step false alarm reduction technique including pixel- and object-based methods as post-processing. The proposed pixel-based algorithm compensates motion, decreasing false alarms at low computational cost, while the following object based method further reduces the residual false alarms by machine learning techniques. We introduced a new quality metric for detection methods by measuring the required amount of manual work after the automatic detection. In our novel evaluation technique, the ground truth is collected from digitized archive sequences where defective pixel positions are detected in an interactive process.

A. Licsár (✉) · L. Czúni
Department of Image Processing and Neurocomputing,
University of Pannonia, Egyetem str. 10,
8200 Veszprém, Hungary
e-mail: licsara@almos.vein.hu

L. Czúni
e-mail: czuni@almos.vein.hu

T. Szirányi
Distributed Events Analysis Research Group,
Computer and Automation Research Institute,
Hungarian Academy of Sciences, Kende u. 13-17,
1111 Budapest, Hungary
e-mail: sziranyi@sztaki.hu

Keywords Digital film restoration · Blotch detection · Object classification · Motion estimation

1 Introduction

In film archives around the world there are huge amounts of archive films to be restored. Archive films suffer from several degradations such as blotches, scratches, image vibration, flickering, fading, discoloring, etc. Digital restoration methods should provide semi-automatic processing with minimal human intervention at low cost and high quality, resulting in efficient restoration of the film heritage. Film archive professionals emphasize high quality results without introducing new artifacts and prefer changes only where inevitable. These aspects are more important than a highly automated workflow [1].

Blotches are usually caused by dirt, damage of the film surface and chemical or biological degradation processes such as the effects of mold. Hence, these one-frame defects can be modeled as temporal intensity discontinuities not originating from different object motion such as non-rigid motion, (dis)occlusion, motion blur, or changes in the lighting. For example, a bright, fast moving object in front of a dark background could cause intensity discontinuities resulting in false alarm (FA) detections (uncorrupt pixels recognized as defected regions).

Experiences in the restoration of archive films [2] show that the correction of blotches needs high computational complexity and significantly more manual work than the correction of other artifacts. This is caused by the facts that (1) blotches are one-frame defects that appear at random positions on consecutive frames and have arbitrary shape, size and varying range of intensity; (2) there is a huge amount of data to be processed, e.g. motion estimation (ME) and

Table 1 Comparison of different blotch detection methods

Author	Detector	Support	Frames	ME is in	ME	Post-processing	Objective evaluation on
Tenze [14]	Morph	Spatial	1	–	–	–	–
Naranjo [11]	Morph	Hybrid	3	–	–	–	–
Joyeux [13]	Morph	Spatial	1	–	–	Morphology	Artificial blotches
Nadenau [7]	ROD	Temporal	3	Pre-processing	BM	Morphology	Artificial blotches
Kokaram [5]	SDI,MRF,AR	Temporal	3	Pre-processing	BM,GM	Morphology	Artificial blotches
Buisson [18]	mod. SDI	Temporal	3	Pre-processing	BM	Morphology	Artificial blotches
Gangal [6]	ROD	Temporal	5	Pre-processing	BM	Morphology	Artificial blotches
Ghaderi [4]	SDIa	Hybrid	3	Pre-processing	BM	Morphology	Artificial blotches
Schallauer [19]	mod. ROD	Temporal	3	Pre-processing	BM	Morphology	–
Roosmalen [16]	SROD	Temporal	3, 5	Pre-processing	BM	HT,CRG	Artificial blotches
Tenze [12]	ESTD	Hybrid	3	Post-processing	BM	ME	Artificial blotches
Our proposed	SROD	Temporal	3	Post-processing	GM	ME+OC	Real blotches

BM: block matching with pixel accuracy; *GM*: gradient based method with sub-pixel accuracy; *ME*: motion estimation; *OC*: object classification; *Real/artificial blotches*: objective performance evaluation of blotch detection is utilized on real blotches collected from archive film or on artificially generated blotches; *HT*: hysteresis thresholding; *CRG*: constrained region growing; *SDI*: spike detection index; *ROD*: rank ordered differences; *MRF*: markov random field; *ESTD*: enhanced spatio-temporal detector; *AR*: auto-regressive; *Morphology*: morphological filter based

compensation (MC) at high resolution of 2 K (2,000 pixels per line), 4 K or 6 K images; (3) artifact detection is an ill-posed problem because the detection algorithms need to know the motion around the missing areas, while this information is not reliable in the defected regions; (4) high risk of false alarm detection. Accordingly, our goals are to speed up the restoration work without significantly decreasing the quality and to minimize the necessity of the human intervention, which is a bottleneck of the restoration work.

A typical restoration framework of one-frame defects is the following (see Table 1). (1) Pre-processing (including the neighboring images), i.e. ME and MC; (2) applying a blotch detector to estimate the temporal discontinuities; (3) post-processing of detection results to increase accuracy; (4) reconstruction of the corrupt image regions. In [3], the problems of detection, reconstruction and motion estimation are handled in one Bayesian framework. The drawbacks of the above two approaches are twofold: (1) computing ME over the whole frame is computationally expensive; (2) the erroneous ME could cause high FA rate. If FA cases are not handled carefully then the reconstruction step could cause a loss of original image details and could generate additional annoying artifacts not acceptable by archivists [1]. In practice, usually, after the automatic detection step an operator, significantly decreasing the efficiency of the restoration process, removes the false alarms manually. In this paper, we focus on the detection process and post-processing methods to avoid the above problems and to reduce the amount of this manual intervention. In the proposed solution, the first step involves the detection of the temporal discontinuities and the following post-processing steps are intended to filter out only the FA results.

The computational cost of our proposed hierarchical gradient-based ME method is relatively low because it decreases the amount of data to be processed without the reduction of the performance. We also introduce a solution where we analyze the spatial features of the blotch candidates for final classification. Since a static blotch model cannot be defined for the various types of sequences and blotches, a trainable classification method detects and removes the FA results.

In the statistical evaluation of the performance of the detection methods, usually some predefined artificial blotch model [4–7] is considered for both of the shape and the surface in the generation of the test data. In our case, this approach cannot be applied because the proposed object-based classification method analyzes the image intensity features, which would be effected by the artificial blotch generation model. Other evaluation technique works on real archive films where the locations of defects are determined by infrared scanning [8,9]. The disadvantage of this approach is that surface defects can be utilized only on color films. Accordingly, we applied a novel test bed generation process that involves digitized archive sequences and the areas of reference blotch masks are marked by an interactive process.

We also have proposed a new quality metric, for the performance evaluation of our proposed methods, based on the estimation of the amount of the required manual work after the automatic processes. This important step of the restoration process is also considered in our experiments.

The idea of the blotch classification was first introduced in our previous work [10] that is now completed with improved region of interest (ROI) based motion estimation, blotch detection, feature extraction, classification, training and evaluation methods.

Our approach has the following novelties:

1. A new restoration framework is introduced with a two-step post-processing method where the first stage removes FA results by temporal analysis (utilizing ME and MC) and the second step removes the residual false detections by spatial analysis and classification of image features of blotches.
2. In the blotch detection process, the hierarchical gradient-based ME method is computationally optimized to minimize the computational consumption.
3. Image features of blotches are analyzed in a trainable framework to reduce the FA detections that were not removed due to the unreliable ME.
4. The performance is evaluated on real archive images, where the ground truth set of the defected pixels is collected by a semi-automatic process. This approach gives objective performance evaluation with different kinds of blotches and it can be utilized on both color and black& white films.
5. A new quality metric is introduced in our evaluation process for the image restoration in real situations considering the human factor.

2 Overview of blotch detectors and pre/post-processing methods

Blotch detector algorithms can be divided into three main groups: (1) spatial detection by analysis of the contrast or local maxima/minima; (2) spatio-temporal methods based on the detection of temporal discontinuities; (3) hybrid methods which are the mixture of the previous two groups [4, 11, 12].

The first group includes fast morphological filters [13, 14] to detect impulsive distortions but with limited capability to detect large blotches with low contrast. These methods usually fail for most of the dynamic highlights (nearly saturated areas, like gleaming).

Algorithms in the second group are based on the detection of temporal discontinuities among several frames like the 3D auto-regressive (AR) models [5], spike detection index (SDI) [15], Markov random field (MRF) [5], rank ordered differences (ROD) [7], simplified ROD (SROD) [16], enhanced spatio-temporal detector (ESTD) [12], spatio-temporal multilevel median filter (MMF filters) such as ML3D or ML3Dex [15], soft morphological filters (SMF) [17]. There are algorithms with three [15, 16] or five-frame [6] temporal support where discontinuities are estimated between two or more frame pairs. However, increasing the temporal aperture does not necessarily increase the performance, but it does increase the computational cost. According to the comparative evaluation of MRF, SDI and ROD methods in [7], ROD resulted in the best trade-off between accuracy and complexity. The

performance of the SROD detector is similar to the performance of the ROD detector for high CD rates and it is similar or better than the ROD detector [16] at low false alarm rates.

A typical pre-processing step of spatio-temporal methods is the bi-directional MC between actual and neighboring frames to reduce false alarms originating from motions. ME is usually done by block matching [7, 16, 18] with pixel accuracy or by gradient-based motion estimation techniques (GM) [15] with sub-pixel accuracy. Usually, ME methods apply only translational motion models, which can fail in the presence of complex motion. One of the possible solutions can be the embedding of a more sophisticated motion model but this can result in a very complex and expensive ME process on high resolution images. In order to decrease the computational complexity, block matching methods are applied with multi-scale processing [7, 16, 19] or with heuristic (e.g., logarithmic) searching [6, 18]. The drawbacks are that they do not guarantee optimal detection and the estimation gives only pixel accuracy. Another way to reduce the computational cost of the ME is the pre-selection of the pixels of the possible defected regions and thus the calculation of the optical flow is not necessary on the whole image [12]. In this approach, a spatio-temporal method (ESTD) detects artifacts and then a ME is utilized in the regions labeled as blotches. However, this approach was not compared experimentally to the original method, where the ME is done in the pre-processing step, and only the detector was evaluated.

Often the post-processing of detector results is needed to refine the partially detected blotches using dilation, improved constrained region growing or hysteresis thresholding [16]. There are other processing methods to detect false alarm objects, originating from image noise, by summarizing the detector response of each position of a blotch [16]. The drawback is that the detector response is affected by the erroneous ME and MC so in the worst case it can even increase the false alarm rate. On the contrary, our post-processing approach analyses only the spatial features of blotch candidates so it is able to filter out most of the falsely detected areas originating from erroneous ME. All of the methods in Table 1, except for the method of Nadenau [7], show detection results on real images but objective performance evaluation is computed only on artificially generated blotches. In Table 1, we summarized the main features of all the discussed algorithms.

In our paper, the correction of blotches (inpainting) is out of our focus but it should be investigated in relation with the false alarm problem. In case of erroneous ME, the locally bright or dark areas can be detected as corrupted regions (FA detections) because the registration of their positions between neighboring frames is not correct. If the reconstruction of FA results utilizes spatially surrounding information [13, 14] these regions can be degraded. Spatio-temporal interpolation methods [6, 16, 18, 19] fuse both spatial and temporal surrounding information after motion compensation.

Since the erroneous motion vectors causes FA detections of uncorrupted regions these vectors results in falsely motion compensation and interpolation in the reconstruction step. Several methods try to avoid this effect by the corrections of motion vectors [16, 19] in the area of the detected regions assuming that the blotch artifact disturbed the motion estimation in that region. However, this step only spatially interpolates the motion vectors from the surrounding motion vectors that do not guarantee good results. We follow the similar approach as in [19].

3 The fundamentals of the proposed blotch detection

In this section we summarize the steps of our blotch detection method that involves flicker correction, blotch detector, ROI-based motion estimation and an object classification phase. We selected a state-of-the-art technique SROD [16] as blotch detector. The main steps of our algorithm are the following (see Fig. 1):

1. *Flicker correction.* The later applied ME is based on image gradients which is sensitive to intensity fluctuations. Therefore, intensity compensation is needed in case of flickering sequences.
2. *Blotch detector.* We detect temporal discontinuities on three adjacent frames resulting in list of candidate areas referred as ROI. We modified the original SROD method by setting the radius to zero ($r = 0$, see Eq. 1) and thus maximized the correct detection rate.
3. *ME and MC.* The detected ROI mask involves false alarms typically due to object motion. In our solution, the ME and MC steps are calculated only on this mask. This approach keeps the cost of motion estimation low.
4. *Blotch detector on the motion compensated mask areas.* Now the radius parameter is set to one ($r = 1$, see Eq. 1) that is the standard SROD detector.
5. *Object classification.* The connected components of the output of the second blotch detector are considered as objects. An image feature based classification method detects and removes falsely detected objects.

The basic concept of the above procedure is to run the blotch detection with high CD rate and then reduce the FA rate in two consecutive post-processing steps (step 4 and 5). The CD rate (also known as true positive) is measured as the fraction of correctly detected pixels out of the total number of the corrupted pixels while the false alarm rate is measured as the fraction of incorrectly marked pixels out of all uncorrupted pixels.

The problem of strong flickering of archive films cannot be neglected since it may have serious unwanted effects on motion estimation. An efficient and robust flicker correc-

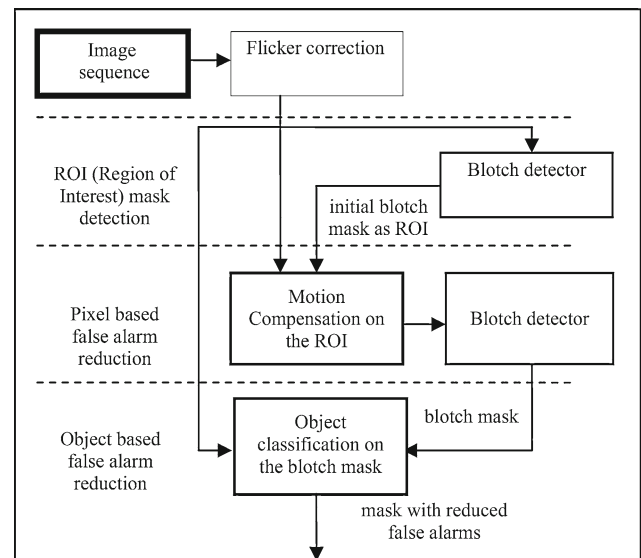


Fig. 1 Processing flow of the proposed blotch detection method

tion method is introduced in [16, 18], where the noise model involves multiplicative and additive components. The deflickering is computed before blotch detection. Fast camera motion can also corrupt results so global motion compensation is utilized by phase correlation method [20].

In steps 2 and 4, the temporal discontinuities are detected by a modified SROD detector:

$$\text{SROD}(\bar{z}, r) = \begin{cases} \min(p_r(\bar{z}) - I(\bar{z})) & \text{if } \min(p_r(\bar{z}) - I(\bar{z})) > 0 \\ I(\bar{z}) - \max(p_r(\bar{z})) & \text{if } I(\bar{z}) - \max(p_r(\bar{z})) > 0 \\ 0 & \text{otherwise} \end{cases}, \quad (1)$$

$$\text{mask}_{\text{SROD}}(\bar{z}) = \begin{cases} 1 & \text{if } \text{SROD}(\bar{z}, r) > T \\ 0 & \text{otherwise} \end{cases}, \quad (2)$$

where the SROD function (Eq. 1) analyzes the minimal difference of luminance values corresponding to position \bar{z} between the actual and the neighboring frames. Function $I(\bar{z})$ expresses the pixel luminance on the actual image while $p_r(\bar{z})$ is a set of reference values on the preceding and consecutive images. This set is collected from the horizontal and vertical neighborhood of \bar{z} given by radius r . Consequently, if $r > 0$ then samples are collected in the neighborhood of \bar{z} defined by a cross shape. The blotch mask is segmented by a threshold function (Eq. 2), where parameter T is called the sensitivity of the detector. In the original SROD method r is set to 1. In our case, if r is 0 then spatial neighborhood information is not used because the detector utilizes only the \bar{z} position on the neighboring frames giving higher detection and false alarm rates.

4 False alarm reduction with ROI based motion estimation

We reduced FA results and high computational cost with ROI-based motion estimation. The result of this step is called SROD-MC-SROD in the remaining part of the paper. A similar detection-MC-detection approach can be found in [12] with ESTD detector and BM motion estimation method. However, the statistical evaluation and the performance analysis were examined without MC. The disadvantage of this technique is that objects with large motion can be falsely detected as blotches. Therefore, we improved this technique with a hierarchical gradient-based motion estimation technique with a ROI-based estimation and compensation of motion. In our method, a modified SROD detector is applied to increase the CD and to decrease the FA rates of this approach. We further reduced the cost of ME step by utilizing the advantages of both SROD-MC-SROD and modified hierarchical gradient based motion estimation method.

In our three-frame blotch detection process the bi-directional dense optical-flow fields are computed by the hierarchical gradient based motion estimation of Bergen [21]. We generate MC images and refine the optical flow simultaneously in a top-down approach in both directions ($n = i - 1, n = i + 1$). We start at the highest level of the pyramid, compute the optical flow, upscale the optical flow by two, and then generate MC images at a layer below. At this lower level we refine the optical flow between the MC images ($n = i - 1, n = i + 1$) and the central image ($n = i$). At the bottom of the pyramid we get two MC images which will be the inputs of the SROD detector.

We use a five level Gaussian pyramid [22] (in our pyramid structure the level 4 is the original image and level 0 is the top level, furthermore the resolution factor is 2 between adjacent levels) and the spatial support of the image gradient is a 9×9 window (this later value was determined experimentally on noisy 2 K archive images, see Sect. 6.2). The size of the window is a trade off between the maximum range of the ME and the detection sensitivity. The motion vectors are calculated by eigenvalue analysis, where the ratio of the greatest and the least eigenvalue is investigated [21]. If this ratio is relatively small (in our case under 100) then the motion is well constrained along both eigenvectors otherwise only the eigenvector corresponding to the higher eigenvalue is utilized.

In our solution we achieved speedup in two ways (see Fig. 2):

1. Motion is estimated only over a mask at each level of the pyramid. Each mask is a resized version of the original ROI but morphologically enlarged by a 6-step dilation.
2. We stop the refinement of the ME before reaching the finest level. Instead, we just upscale the optical flow by interpolation to generate the final motion compensated

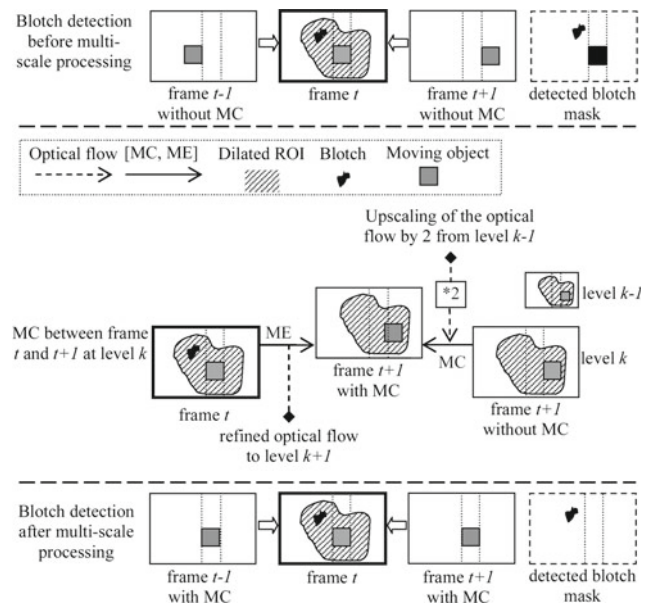


Fig. 2 False alarm reduction with region of interest (ROI) based motion estimation (ME) and compensation (MC) (only forward direction and one level of the pyramid are shown)

neighboring images. Since the gradient-based estimation gives sub-pixel motion values the upscaling still results in sub-pixel precision giving satisfactory results as experienced in our tests (see Sect. 6). Making the estimation on the finest level would be slower without significantly increasing the precision due to the high frequency noise of archive images (grain) at full resolution.

Finally, the SROD detector, with setting $r = 1$, is applied on the motion compensated images. A further benefit of the above method is that the computational cost basically depends on the amount of motion and on the amount of artifacts.

5 False alarm reduction with trainable blotch classification

After the pixel-based detection the connected components of the obtained mask are considered as objects. The image features of these areas are extracted for the classification as real blotches (true positive) or non-blotch (false positive or FA). We investigate the following image features as input parameters for the classification:

- the maximum of horizontal and vertical intensity gradient on the blotch area,
- the internal (inside the blotch area) and external (outside the blotch area) variances that are calculated by pixel intensities of the internal/external area of the blotch,

- the internal and external contrast (difference of minimal and maximal intensity),
- the difference of internal and external mean values,
- the difference of internal and external variances,
- the perimeter.

The external area of the object is determined by the bounding box of the defect, enlarged by a 2-step dilation, excluding pixel positions of the detected blotch. Since it is very difficult to construct an explicit blotch model for a given film and using general models might be useless considering the large variety of archive films we propose an adaptive classification by training [10].

The classification, by the above parameters, is carried out by a feed-forward neural network (NN) [23] or by a support vector machine (SVM) [24]. Both methods have eight input parameters and return with two possible outputs: true blotch or false blotch. The NN method includes two hidden layers with 12 and 10 neurons, determined experimentally, and the transfer function is selected as a sigmoid function. The training algorithm is the well-known back-propagation [25] where the learning factor, which controls the speed of learning, is set to 0.6. The SVM-based training applies C-support vector classification method [24] where the kernel is a radial basis function (RBF):

$$K(x_i, x_j) = e^{-\gamma \|x_i - x_j\|^2}. \quad (3)$$

Here x are the training samples and the γ parameter was determined as 0.2 by our experiments.

Training samples are generated in a two-step method where first the above proposed automatic blotch detection method finds blotch candidates, then the operator qualifies the results as real blotches or false alarm results. Naturally, this approach requires high sensitivity from the automatic detection mechanism, since we do not want the operator to waste his time by the careful manual selection of undetected blotch areas. To accelerate the manual process the operator should only correct the displayed label of a given candidate or he can select more than one object framing an image region. After the teaching process the blotches are classified automatically with one of the trained mechanisms and if the correct recognition rate, calculated on the training set, is not satisfactory, the operator can expand the training set with new samples or rebuild the whole set. The typical scenario of blotch correction is the following:

- (1) automatic blotch detection on a given sequence,
- (2) visualization of the blotch candidates and manual collecting of blotch and non-blotch samples (false alarms) for training purposes,
- (3) automatic classification of detected blotches on the whole image sequence to reduce the necessity of the manual intervention,
- (4) visualization of the classification result (see Fig. 6) and manual correction of the remaining false alarm detections where the operator is able to modify the result of the previous classification step,
- (5) correction only of the verified list of blotches.

6 Performance analysis

6.1 Generation of the ground truth data from real sequences

Usually, the performance of detection methods is statistically evaluated on artificially generated test sequences. Randomly selected image regions are replaced with a rendered blotch defined by a simple model, which is in general a homogenous surface with sharp contour.

To be more realistic we made tests on digitized archive films where the ground truth data set consisted of defective pixel positions, selected by an interactive method. This is an important issue since in the post-processing phase the intensity information over these positions would be essentially determined by an artificial blotch model.

In our previous work [10], the silhouette of defects were manually painted for each pixel of the image. This process was time consuming so we developed a technique that speeds up this process. In the ground truth generation process, an SROD detector is applied, without MC, at high sensitivity. The threshold parameter T (see Eq. 1) was set to 10% of the intensity range (in our case $T = 25$ in eight-bit quantization). The detected candidates are visually marked on the displayed image. In the next interactive process, an expert verifies each candidate and labels the true blotches. These objects build up the ground truth set. If the result is still not satisfactory, manually painted masks can be also drawn. In the next steps, the holes are filled with binary morphology to remove the noise of the detector. In the ground truth preparation process other artifacts such as scratches or semi-transparent blotches are omitted because their detection is out of our focus. The manual selection of the small (pixel size) distortions is not realizable owing to the large grain noise of high resolution archive images. For this reason, before building the ground truth, small impulsive distortions of perimeter less than 10 pixels (at 2 K resolution) are removed by a rank order spatial filtering method [26].

In this paper, the test sequences are from the first Hungarian color film “Mattie the Goose-Boy” (1949) scanned at 2 K resolution. The extracted five test sequences include 140 frames with the corresponding blotch masks. This ground truth database includes a variety of video content such as fast local and global motion, motion blur, occlusion, small random patterns, etc.

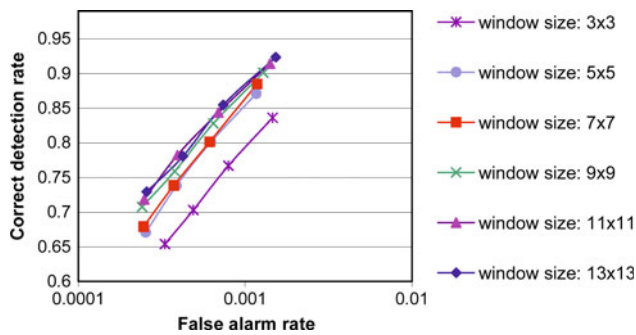


Fig. 3 The performance of the proposed method (SROD-MC-SROD) with different size of spatial support of the image gradient by ROC analysis

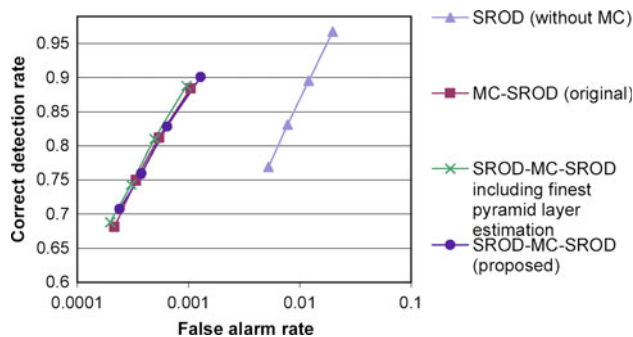


Fig. 4 The performance of the proposed method compared to the original techniques by ROC analysis

6.2 Performance evaluation of blotch detection

We examined the performance of our blotch detection method regarding the window size of the spatial support of the image gradient (Fig. 3) and we compared the accuracy (Fig. 4) and the time consumption (Fig. 5) of our proposed method (SROD-MC-SROD) with the SROD detector and the motion compensated SROD (MC-SROD). The objective evaluation is done by the receiver operating characteristic (ROC) curve summarizing the correct detection and false alarm rates measured at different sensitivity parameter settings (see parameter T in Eq. 2). This parameter was set between 30 and 60 in steps of 10.

Figure 3 describes ROC curves calculated by the proposed SROD-MC-SROD method with different size of spatial support of the image gradient. It can be observed that the support window of size 3×3 , 5×5 and 7×7 yields lower accuracy than the cases of 9×9 , 11×11 and 13×13 which give similar performance. In further experiments the size is fixed at 9×9 . In Fig. 4, it can be seen that ROC curves of the SROD-MC-SROD and that of the MC-SROD are approximately identical while the time consumption of our method is significantly (with 50%) reduced (see Fig. 5). We also show the result of the SROD detector without MC to illustrate that the test sequences had significant motion.

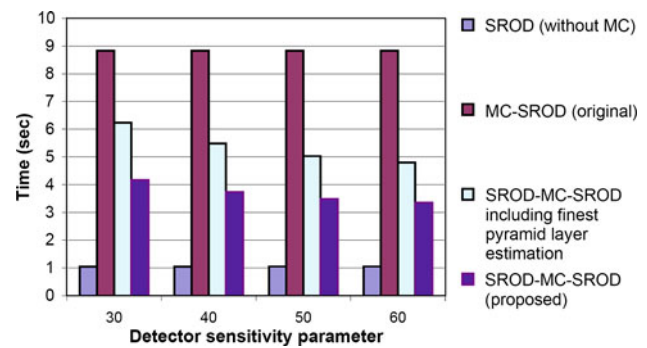


Fig. 5 Time consumption of the proposed method compared to the original techniques (on a 3.4 GHz PC)

On a 3.4 GHz Intel PC the processing time of a 2 K image is less than 4.2 s with the proposed method while the original MC-SROD is 8.8 s. We illustrate that the effect of simply upscaling motion information to the finest level of the image pyramid results in similar CD and FA rates than estimation on the whole pyramid (see Sect. 4).

The speedup depends on the amount of motion and on the FA rate affected by the sensitivity parameter. For comparison we also give the computational time of two 3-frame bi-directional motion estimation methods:

- recursive block matching at 2 K on 700 MHz processor is 60 s [18],
- a “log D ” block matching at CIF resolution (352×288 pixels) on 2.4 GHz processor is 25 s [6].

6.3 Performance of the trainable blotch classification method

The aim of object classification is to decrease the FA rate without decreasing the CD rate. In this section, we analyze the effects of classification considering the FA rate and the CD rate. The values of FA and CD are measured at several sensitivity settings, as given in the previous section, and the average measured values are compared.

Two training sets were composed in our experiments and in all cases the training and the test sets were different. Training and test sequences were collected from the same scene. Training Set I was collected from five different sequences. The first set involves 168 positive and 168 negative samples. The CD rate on the training set was 94.4% with the NN and 93% with the SVM method. Training Set II is created from Set I with including additional training samples from the actually processed image sequence: an operator collected 20 positive and 20 negative samples from the actually processed sequence.

It has been proved that with one hidden layer a NN can describe any continuous while with two hidden layers it can

Table 2 Average false alarm (FA) and correct detection (CD) rates (%) measured in case of several architectures of neural network

Number of nodes in the two hidden layers ^a	SROD-MC-SROD with NN	
	FA	CD
8;6	0.010	75.1
10;8	0.010	75.2
12;10	0.010	74.9
14;12	0.009	74.5
16;14	0.009	75.2

^a 12;10 indicates that the two hidden layers involve 12 and 10 neurons

estimate any function [27]. To examine the sensitivity of the NN, we have tested several architectures with different number of nodes in the two hidden layers (Table 2) where training Set II was utilized.

It can be seen (Table 2) that different architectures of neural network produce similar accuracy. To ensure the generality of the NN, we utilize two hidden layers with 12 and 10 neurons in further experiments.

We compared the performance of the SROD-MC-SROD method with NN and SVM-based post-processing (see SROD-MC-SROD-NN and SROD-MC-SROD-SVM methods in Table 3). The changes in the performance of both methods are included also in Table 3 as the average of ratios of FA and CD rates (e.g., FA^B/FA^A) over five test sequences. The blotch classification methods significantly reduced the FA rate by 80–85% while the CD rate decreased by only 6–12%, which is in accordance with the general requirement of archivists. It can be observed that the NN gives slightly higher CD and FA rate compared to the SVM method. In case of Training Set II we experienced that both of the CD and FA rates have increased but this change is not significant.

The average processing time of feature extraction and classification was less than 1.7 s/frame at the highest sensitivity and about 0.7 s/frame at the lowest detector sensitivity with both methods on a 3.4 GHz Intel PC.

In Fig. 6 rectangular boxes and circles indicate the false and true blotches classified where the size of frames are proportional to the estimated blotch size. Examples show several situations where erroneous motion estimation results in false alarm. Sequence “A” and “B” (Fig. 6) shows examples for non-rigid motion (frame#17 and#19 on sequence “A”), occlusion/disocclusion and motion blur (e.g., stick of the man on frame#22 and buttons of clothes on sequence “B”), or gleaming (gun of the soldier on frame#15 on sequence “A”).

6.4 Objective measurement of human intervention

One of our aims was the reduction of the necessity of human intervention after the fully automatic processing. At this step the detected candidates and the results of the classification are also displayed for the operator. To accelerate the manual process the expert should only correct the displayed label of candidates. The operator has to verify all of the candidates and has to correct (label) some of them. Accordingly, we estimated the amount of the manual work by measuring the number of the wrongly classified blotches before and after the proposed blotch classification. We automatically measured how many manual reclassifications (from true blotch to false blotch and vice versa) would be necessary to achieve the best matching with the ground truth database resulting in the highest CD and lowest FA rates. The time consumption of the manual correction step is proportional to the number of blotches to be displayed and verified and to the number of the blotches to be reclassified by the operator. To determine these values we summarize the FA and false negative (FN) detection of blotch objects where classification methods were trained by the samples of Set II. These values are summarized in Table 4 and give an indication of the level of manual intervention at different detector sensitivity values that determines the maximal CD rate of the SROD-MC-SROD method. Considering the grain-like image noise our algorithms disregards candidates with perimeter less than 10 pixels. It is important that ROC curves give pixel based statistics while the present measure evaluates object based statistics.

Table 3 Average false alarm (FA) and correct detection (CD) rates (%) after NN-and SVM-based object classification with two training sets

Training sets	SROD-MC-SROD		SROD-MC-SROD –NN		SROD-MC-SROD –SVM	
	FA ^A	CD ^A	FA ^B	CD ^B	FA ^C	CD ^C
Average rates (%) over five test sequences						
Set I.	0.08	80.0	0.009	72.7	0.007	70.0
Set II.	0.08	80.0	0.010	74.9	0.008	72.5
Average of ratios of rates over five test sequences in % ^a						
Set I.	FA^B/FA^A		CD^B/CD^A		FA^C/FA^A	
Set II.	17.9		91.1		15.1	
Set I.	CD^C/CD^A		FA^C/FA^A		CD^C/CD^A	
Set II.	87.8		16.0		90.7	

^a $FA^B/FA^A = 17.9\%$ means that the FA^B rate decreases by 82.1% compared to FA^A rate

Fig. 6 Classification examples of the proposed post-processing on two image sequences: original image sequence (*top*) and its zoomed image parts (*bottom*). False alarms of the initial detections are marked with *rectangles* while correct detections are indicated by *circles*



Table 4 The amount of the manual intervention (correction/frame) that reduces false alarm (FA) and false negative (FN) detections

<i>T</i>	SROD-MC-SROD		SROD-MC-SROD with NN		SROD-MC-SROD with SVM	
	No. of objects to be verified : No. of objects to be manually corrected					
	CD (%)	FA	FN	FA	FN	FA
30	90	39.6 : 34.4	32.4 : 0.6	7.2 : 2.5	34 : 1.2	5.6 : 1.6
40	83	18.7 : 14.3	12.9 : 0.4	5.8 : 1.7	13.9 : 0.8	4.8 : 1.2
50	77	11.6 : 7.9	6.8 : 0.2	4.8 : 1.3	7.4 : 0.5	4.2 : 1
60	70	7.9 : 4.8	3.9 : 0.1	4.0 : 0.9	4.3 : 0.2	3.6 : 0.7

T: parameter of the detector sensitivity; *CD*: correct detection rate in %; *Manual intervention*: the operator manually accepts or rejects the candidates with the state of real blotch or non-blotch; *FA*: the candidate is wrongly classified as real blotch; *FN*: the candidate is wrongly classified as non-blotch

It can be observed that without blotch classification the operator needs to verify more than 39.6 blotches/frame from which he should correct 34.4 blotches/frame in order to mini-

mize the FA rate. After the automatic classification the operator should only verify 7.2 (5.6) blotches and select 2.5 (1.6) blotches on one frame with the NN (SVM) method at the

highest sensitivity. We experienced that the operator should make much less correction steps to maximize the CD rate with the NN classification method (see FN in Table 4). Therefore, the operator needs negligible manual work to achieve a higher CD rate. Otherwise, if the main criterion is the low false alarm rate without any human intervention (fully automatic mode) the SVM-based method is favorable due to the lower false alarm rate (see FA in Table 4). The reduction of the manual interventions is between 50 and 86 % if the main criterion is the minimal FA rate. Since the SROD-MC-SROD method does not utilize any classification the detection result involves only candidates with the state of real-blotch and so FN (the candidate is wrongly classified as non-blotch) parameters cannot be measured in the Table 4.

7 Conclusion

In this paper, we discussed the problems of blotch detection which is a key issue in the correction of one-frame defects. Important aspects were computational efficiency at high resolution and low false alarm rate with minimal human interaction.

In order to achieve these goals, we developed a two-step detection method with SROD detector to keep the computational cost of the motion estimation low and to reduce the amount of FA detections by object-based classification as post-processing. Other techniques compare methods on real sequences only in subjective tests or they utilize non-archive and undamaged sequences where the blotch artifacts and the ground truth are artificially generated.

We proposed a new evaluation technique of detection methods where the test sequences are digitized archive images at 2 K resolution and the ground truth data set is generated by a semi-automatic process. We also introduced a new quality metric for the film restoration in real situation. It evaluates the performance of our post-processing method and it is based on the estimation of the required manual work to remove false classifications of candidate areas. The main advantage of our method is that it decreases false detections where the motion estimation fails in the first step due to the fast or complex motion of objects (non-rigid motion, (dis)occlusion of moving objects, etc.) or low image quality of archive films. Our experiments tested on archive sequences proved that the proposed method significantly speeds up the motion estimation step without quality loss compared to other SROD methods and our post-processing method significantly reduces the FA rate and the necessity of the manual work.

References

1. Delaney, B., Hoomans, B.: Prestospace user requirements feedback meeting in London. In: *An Integrated Solution for Audio-Visual Preservation and Access* (2004)
2. Czúni, L., Hanis, A., Kovács, L., Kráncz, B., Licsár, A., Szirányi, T., Kas, I., Kovács, Gy., Manno, S.: Digital motion picture restoration system for film archives (DIMORF). *SMPTE Motion Imaging J.* **113**, 170–176 (2004)
3. Kokaram, A.C.: On missing data treatment for degraded video and film archives: a survey and a new Bayesian approach. *IEEE Trans. Image Process.* **13**(3), 397–415 (2004)
4. Ghaderi, M., Kasaei, S.: Novel post-processing methods used in detection of blotches in image sequences. *AEÜ Int. J. Electron. Commun.* (58), 58–64 (2004)
5. Kokaram, A.C., Morris, R., Fitzgerald, W., Rayner, P.: Detection of missing data in image sequences. *IEEE Trans. Image Process.* **4**(11), 1496–1508 (1995)
6. Gangal, A., Kayıkçıoğlu, T., Dizdaroglu, B.: An improved motion-compensated restoration method for damage color motion picture film. *Signal Proc. Image Commun.* (19), 353–368 (2004)
7. Nadenau, M.J., Mitra, S.K.: Blotch and scratch detection in image sequences based on rank ordered differences. In: *Time-Varying Image Processing and Moving Object Recognition*, pp. 27–35. Elsevier, Amsterdam (1997)
8. Tilie, S., Laborelli, L.: Blotch detection for digital archives restoration based on the fusion of spatial and temporal detectors. In: *The 9th International Conference on Information Fusion, Florence, Italy*, pp. 1–8 (2006)
9. Ren, J., Vlachos, T.: Segmentation-assisted dirt detection for the restoration of archived films. *Proceedings of BMVC*, vol. 1, Oxford, England, pp. 359–368 (2005)
10. Licsár, A., Czúni, L., Szirányi, T.: Trainable post-processing method to reduce false alarms in the detection of small blotches of archive films. *IEEE International Conference on Image Processing (ICIP)*, Italy, pp. 562–565 (2005)
11. Naranjo, V., Albiol, A., Mossi, J.M., Albiol, A.: Morphological Lambda Reconstruction Applied to Restoration of Blotches in Old Films. *Proceedings of the 4th IASTED International Conference on Visualisation, Imaging and Image Processing, Spain* (2004)
12. Tenze, L., Ramponi, G., Carrato, S.: Robust detection and correction of blotches in old films using spatio-temporal information. *Proceedings of SPIE International Symp. Electronic, San Jose*, pp. 348–357 (2002)
13. Joyeux, L., Boukir, S., Besserer, B., Buisson, O.: Reconstruction of degraded image sequences. Application to Film Restoration. *Image and Vision Computing (MAVIS)* **19**(8), 503–516 (2001)
14. Tenze, L., Ramponi, G., Carrato, S.: Blotches Correction and Contrast Enhancement for Old Film Pictures. *Proceedings of the International Conference on Image Processing, ICIP, Canada*, pp. 660–663 (2000)
15. Kokaram, A.C.: *Motion Picture Restoration: Digital Algorithms for Artefact Suppression in Degraded Motion Picture Film and Video*. Springer, Heidelberg (1998)
16. Lagendijk, R.L., van Roosmalen, P.M.B., Biemond, J.: *Video enhancement and restoration. Handbook of Image & Video Processing*, Academic Press, San Diego, pp. 227–242 (2000)
17. Hamid, M.S., Harvey, N.R., Marshall, S.: Genetic algorithm optimization of multidimensional grayscale soft morphological filters with applications in film archive restoration. *IEEE Trans. Circuits Syst. Video Technol.* **13**(5), 406–416 (2003)
18. Buisson, O., Boukir, S., Besserer, B.: *Motion Compensated Film Restoration. Machine Vision and Applications*, Vol 13, Issue 4, pp. 201–212 Springer, Heidelberg (2003)
19. Schallauer, P., Pinz, A., Haas, W.: Automatic restoration algorithms for 35 mm film. *Videre J. Comput. Vis. Res.* **1**(3), 60–85 (1999)
20. Licsár, A., Czúni, L., Szirányi, T.: Adaptive Stabilization of Vibration on Archive Films. *Lecture Notes in C.S. LNCS 2756* pp. 230–237. Springer, Heidelberg (2003)

21. Bergen, J.R., Burt, P.J., Hingorani, R., Peleg, S.: A three-frame algorithm for estimating two-component image motion. *IEEE Trans. Pattern Anal. Mach. Intell.* **14**(9), 886–896 (1992)
22. Burt, P.J., Adelson, E.H.: The Laplacian pyramid as a compact image code. *IEEE Trans. Commun.* **31**(4), 532–540 (1983)
23. Fukushima, K.: Cognitron: a self-organizing multilayered neural network. *Biol. Cybern.* **20**, 121–136 (1975)
24. Cortes, C., Vapnik, V.: Support-vector network. *Mach. Learn.* **20**, 273–297 (1995)
25. Rumelhart, D., Hinton, G., Williams, R.: Learning internal representations by error propagation. *Parallel Distributed Processing*, MIT Press, Cambridge, pp. 318–362 (1986)
26. Czúni, L., Császár, G., Hanis, A., Kovács, L., Licsár, A., Szirányi, T.: Semi automatic digital motion picture restoration system with learning capabilities. In: *Learning for Adaptable Visual Systems (LAVS)*, Cambridge, UK, 2004
27. Cybenko, G.: Approximation by Superposition of a Sigmoidal Function. *Math. Control Signals Syst.* **2**, 303–314 (1989)