Available online at www.sciencedirect.com
ScienceDirect
Signal Processing 87 (2007) 541–551
SIGNAL PROCESSING
www.elsevier.com/locate/sigpro

Efficient detection of temporally impulsive dirt impairments in archived films ☆

Jinchang Ren[a,b,*], Theodore Vlachos[a]

[a]Centre for Vision, Speech and Signal Processing, University of Surrey, Guildford, GU2 7XH, UK
[b]School of Computers, Northwestern Polytechnic University, Xi'an, 710072, China

Received 8 March 2006; received in revised form 19 June 2006; accepted 20 June 2006
Available online 18 July 2006

Abstract

We propose a novel approach for the detection of temporally impulsive dirt impairments in archived film sequences. Our method does not require motion compensation and uses raw differences between the current frame and each of the previous and next frames to extract a confidence signal which is used to localize and label dirt regions. A key feature of our method is the removal of false alarms by local region-growing. Unlike other work utilizing manually added dirt impairments, we test our method on real film sequences with objective ground truth obtained by infrared scanning. With confidence information extracted from color channels, dirt areas of low contrast to the corresponding gray image can be successfully detected by our method when motion-based methods fail. Comparisons with established algorithms demonstrate that our method offers more efficient, robust and accurate dirt detection with fewer false alarms for a wide range of test material.

Keywords: Archive restoration; Film dirt detection; Noise detection; Image enhancement; Image restoration; Video signal processing

1. Introduction

The automatic restoration of historic films is emerging as a key enabling technology towards improving access to cultural assets of unique value. By improving baseline picture quality and by reducing the perceptual impact of archive-related impairments film archive restoration can meet viewers' aesthetic expectations and enrich the viewing experience. Moreover, the suppression of such impairments has vital implications on the efficiency of video coding algorithms used in the television and multimedia distribution chains such as MPEG-2. Consequently, film restoration has recently attracted a lot of interest and several high-profile collaborative projects have received generous funding at an European level such as AURORA (Automatic Restoration of Original Film and Video Archives, 1995), BRAVA (Broadcast Restoration of Archives by Video Analysis, 1999) and more recently PrestoSpace (Preservation towards Storage and access Standardised Practices for Audiovisual Contents in Europe, 2004).

In their lifetime, films may suffer many kinds of damages due to environmental hazards such as humidity and dust, chemical instabilities, improper

☆This work was carried out in the framework of project PrestoSpace, supported by the European Commission FP6-IST-507336.
*Corresponding author. Centre for Vision, Speech and Signal Processing, University of Surrey, Guildford, GU2 7XH, UK.
E-mail address: j.ren@surrey.ac.uk (J. Ren).

0165-1684/$ - see front matter © 2006 Elsevier B.V. All rights reserved.
doi:10.1016/j.sigpro.2006.06.013

storage and handling practices and even poorly maintained projectors [1–7]. In this paper, we focus on the detection of impairments occasionally referred to 'dirt'. Generally, dirt is a temporally impulsive (single-frame) event, appearing mostly as dark or bright opaque spots of random size, shape and location (see Fig. 1). It is due to particles that are attached to the film or abrasions which occurred during storage or when the film passed through various transport mechanisms [1,4]. These are among the most commonly encountered impairments and consequently their successful detection is a priority issue in any archive restoration system.

In this paper, we propose a novel automatic method for dirt detection based on local region-growing. Firstly, a confidence measure is defined and attached to detected dirt areas. This allows a variable degree of treatment according to the preference and supports fine tuning under various confidence levels. Secondly, using this confidence measure, candidate dirt regions are localized using a local region-growing process. With confidence information extracted from all three color components, dirt areas of low contrast in the luminance channel can be efficiently detected. It should be noted that in such cases motion-based methods are known to fail. Thirdly, for the first time in this type of study objective ground truth (GT) of dirt is used for performance assessment of real film material. Among other things this allows the computation of Receiver Operating Characteristic (ROC) type of curves. Finally, a key element of our method is that it does not require the use of motion estimation and motion-compensated prediction. This reduces its complexity considerably and makes it a good candidate for fast implementations.

This paper is organized as follows. In Section 2, we review related work featuring in the literature while in Section 3 we describe the extraction of a confidence measure for dirt detection. In Section 4, we present the proposed dirt detection algorithm while in Section 5 we provide experimental evidence including a comparative assessment with other methods. Finally, conclusions are drawn in Section 6.

2. Literature review

Storey's work [8] was perhaps the earliest contribution to the electronic detection and concealment of film dirt in the context of a hardware-based system that was proposed. A pixel was flagged as dirt if the corresponding absolute differences between the current frame and each of the previous and next frames were high. As motion-compensated prediction requires a high degree of

Fig. 1. Examples of dirt (marked within white boxes). (a) Static text, (b) low-motion natural scene, (c) moderate-motion natural scene and (d) fast-motion natural scene.

complexity and can be unreliable when motion estimation fails, many spatial filtering techniques for dirt detection have also been proposed as alternatives. Existing methods and models with or without motion compensation are discussed below.

2.1. Dirt detection with motion compensation

Complementary to bi-directional motion compensation, temporal median filtering is a common approach in dirt detection using the current frame and its two motion-compensated neighbors [2,3,9–11]. In Schallauer et al. [2], a pixel is taken as dirt and filtered if both its absolute differences between current frame and the two compensated images exceed a first (higher) threshold while at the same time the absolute difference between the two compensated images is less than a second (lower) threshold, thus we can denote it as a double-threshold method (DTM). Kokaram [3] proposed a so-called 'Spike Detection Index' (SDI) which is based on the identification of high absolute differences between the current frame and two motion-compensated images. The extended SDI method, SDIp, additionally requires sign consensus of the two differences above. Nadenau and Mitra [9], have proposed the rank order detector (ROD), in which a total of seven pixels from three consecutive frames are compared against three thresholds. Gangal et al. [10] extended ROD to five frames to improve accuracy in heavily corrupted images or occluded blotches. Tenze et al. [11] adopted adaptive block matching in spatial–temporal filtering for blotch detection, but their method can only work efficiently when motion is slow.

Let f_n be the current frame and let the two motion-compensated neighboring frames be C_{n-} and C_{n+}. We define D_{n-} and D_{n+} as the differences between each of these two images and the current frame, i.e., $D_{n-}=f_n-C_n$ and $D_{n+}=f_n-C_n$. Using the above dirt detection according to DTM, SDI and SDIp methods are summarized below. In (1), we have $t_2>t_1$, and normally D_{SDIp} is more accurate than D_{SDI} [3].

$$D_{\mathrm{DTM}}=\begin{cases}1 & \text{if } |D_{n-}|>t_2, |D_{n+}|>t_2, |C_{n-}-C_{n+}|<t_1,\\ 0 & \text{otherwise,}\end{cases} \tag{1}$$

$$D_{\mathrm{SDI}}=\begin{cases}1 & \text{if } |D_{n-}|>t_2, |D_{n+}|>t_2,\\ 0 & \text{otherwise,}\end{cases} \tag{2}$$

$$D_{\mathrm{SDIp}}=\begin{cases}1 & \text{if } |D_{n-}|>t_2, |D_{n+}|>t_2, D_{n-}D_{n+}>0,\\ 0 & \text{otherwise.}\end{cases} \tag{3}$$

To determine dirt in ROD, three pair of pixels are extracted from C_{n-} and C_{n+} in three rows, i.e., $(i-1,j)$, (i,j) and $(i+1,j)$. These six pixels are sorted in increasing order in a list $[r_1, r_2,\ldots,r_6]$ where r_6 is the maximum. Then the median of the list is extracted as $\mathrm{med}=(r_3+r_4)/2$. Three thresholds, $t_k\in[1,3]$, are then taken to determine the dirt as follows:

$$D_{\mathrm{ROD}}=\begin{cases}1 & \text{if } \vee(e_k\geqslant t_k)=\text{true, where } t_1<t_2<t_3,\\ 0 & \text{otherwise,}\end{cases} \tag{4}$$

$$e_k=\begin{cases}f_n(i,j)-r_{7-k}, & \text{if } f_n(i,j)>\mathrm{med},\\ r_k-f_n(i,j), & \text{otherwise.}\end{cases} \tag{5}$$

Moreover, we can also find many statistical approaches in dirt detection, using auto-regressive (AR) and Markov random field (MRF) models [12–20]. Unfortunately, there are no definitive statistical models for material of this kind and consequently such methods will fail if their assumptions cannot be satisfied or if accurate and robust motion compensation cannot be achieved [14,15].

2.2. Dirt detection without motion compensation

Considering film dirt as impulsive type of noise in the spatial domain, it can be detected and recovered via filtering without motion compensation. In general, median and morphological filtering is widely used [21–27]. Subsequently, dirt detection is based on the identification of high difference values between the current frame and the filter output. In [21], Alp et al. introduced a ML3D filter for noise removal in image sequences, which utilized two-stage median filtering in three sub-windows. Kokaram [3] extended ML3D to five sub-windows and proposed the so-called ML3Dex filtering approach with which satisfactory results have been reported [4]. Hardie and Boncelet [24] proposed lower-upper-middle (LUM) filters which utilized two parameters for adjustable smoothing and sharpening of images. However, this method did not perform well in the presence of fast-moving objects. Nieminen et al. [25] presented a multi-stage median filter (MMF), which uses hierarchical median operations to reject sparkle type of impairments. Senel et al. [26] proposed a

topological median filter to extract edges in noise; however, the filtered images are of unacceptable visual quality in most cases. In Arce [27], MMF filters have further evolved as multi-stage order statistic filters (MOS). Hamid et al. [4] proposed soft morphological filtering (SMF) and used a genetic algorithm to determine the size and shape in the filters. Nevertheless, SMF seems impractical for most applications because it needs a sufficient number of representative dirt samples for training purposes in order to optimize the filters.

In standard median filtering (SMF) and LUM filtering, a window W of radius r is defined for each pixel (i, j) in the current frame f_n as

$$W(i,j,r) = \{f(i_1,j_1)\}, \quad |i_1 - i| \leqslant r, \quad |j_1 - j| \leqslant r. \tag{6}$$

Hence the total number of pixels in W is $N = (2r+1)^2$. We denote $W = \{x_1, x_2, \ldots, x_N\}$, and the rank-ordered set is given by

$$x_{(1)} \leqslant x_{(2)} \leqslant \cdots \leqslant x_{(N)}. \tag{7}$$

The central pixel in the original current frame and filtered image are denoted as x' and y', respectively. In SMF, we simply have $y' = x_{(N_0)}$, where $N_0 = (N+1)/2$.

In LUM, two parameters, k and l, are introduced for smoothing and sharpening, respectively. Typically it holds that $1 \leqslant k \leqslant l \leqslant N_0$. Then, the filtered output is defined as

$$y' = \begin{cases} x^{\mathrm{L}}, & \text{if } x' \leqslant (x^{\mathrm{L}} + x^{\mathrm{U}})/2, \\ x^{\mathrm{U}}, & \text{otherwise}, \end{cases} \tag{8}$$

where x^{L} and x^{U} are the corresponding outputs of the smoothing and sharpening processes given by

$$x^{\mathrm{L}} = \text{median}\{x_{(k)}, x', x_{(l)}\}, \tag{9}$$

$$x^{\mathrm{U}} = \text{median}\{x_{(N-k+1)}, x', x_{(N-l+1)}\}. \tag{10}$$

In Arce [27], LUM was further applied to a $3 \times 3 \times 3$ spatial–temporal window, which can be denoted as LUM′(N,k), where we have $N = 27$ and $k \leqslant 14$. The output of the filter is given by

$$y' = \text{median}(x_{(k)}, x', x_{(N-k+1)}). \tag{11}$$

Regarding ML3D and ML3Dex filtering, six sub-windows in three consecutive frames are defined as shown in Fig. 2. Let z_l be the median value of all the pixels in W_l, then the output of the two filters can be given by

$$y'_{\mathrm{ML3D}} = \text{median}(z_0, z_5, z_6), \tag{12}$$

$$y'_{\mathrm{ML3Dex}} = \text{median}(z_5, z_6, z_7, z_8, z_9). \tag{13}$$

In both median and morphological filtering, an unsuitable size of window may cause massive false alarms with heavy degradation of edges. To solve this problem, in our earlier work [28] we introduced a segmentation-assisted method for dirt detection, in which the shape and size of the window are determined adaptively via segmentation. However, this method required full segmentation of a frame as well as thresholding according to the region size in order to determine dirt candidates. In this paper, we develop a new algorithm based only on local region-growing for better efficiency, robustness and lower computational complexity.

3. Confidence image extraction

Let f_{n-1}, f_n and f_{n+1} be three consecutive frames, and $h'(f_n)$ denotes the maximum value of the three color components in f_n, i.e., $h'(f_n)(i,j) = \max(r^{(n)}_{i,j}, g^{(n)}_{i,j}, b^{(n)}_{i,j})$. We then define $h(\cdot)$ as

$$h(f_n, f_{n-1}) = h'(f_n) - h'(f_{n-1}). \tag{14}$$

If $\delta_{n-} = h(f_n, f_{n-1})$ and $\delta_{n+} = h(f_n, f_{n+1})$, we define δ_n as

$$\delta_n = \begin{cases} \frac{2\delta_{n-}\delta_{n+}}{|\delta_{n-}|+|\delta_{n+}|} & \text{if } \delta_{n-}\delta_{n+} > 0, \\ 0 & \text{otherwise.} \end{cases} \tag{15}$$

The absolute values of δ_{n-} and δ_{n+} are considered as two elementary confidence indicators of dirt impairment. In (15), δ_n attains its maximum

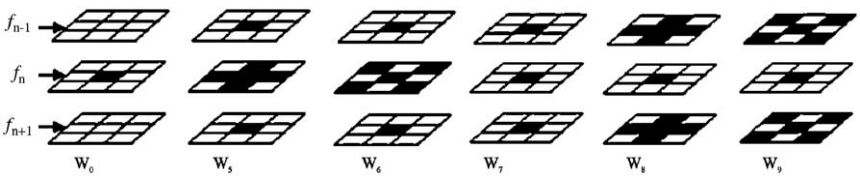

Fig. 2. Sub-windows defined in three consecutive frames for ML3D and ML3Dex filtering (radius = 1).

value when an idealized dirt impulse occurs against a constant background i.e. when $\delta_{n-} = \delta_{n+}$. If both δ_{n-} and δ_{n+} are negative or positive, this relates respectively to dark or bright dirt pixels. Eq. (15) reflects the fact that dirt is strictly a temporally impulsive event thus dirt pixels are expected to generate consistent differences between the current frame and its two neighboring frames. There are instances where the luminance difference between a dirt cluster on the one hand and its neighborhood on the other is low. However, individual color component differences for the same cluster may be significantly higher, such as non-gray dirt [29]. Typically one of the color components will exhibit such a characteristic. The purpose of the max operator in (14) and the combined effect of (14) and (15) are aiming at the exploitation of this characteristic.

For each value m in δ_n, its associated dirt probability is defined in (16), where p_d is the intensity probability density function (pdf) of δ_n, parameter L is the number of intensities, and $p_n(m) = 0$ when $m < m_0$.

$$p_n(m) = \frac{\sum_{x=m_0}^{m} p_d(x)}{\sum_{x=m_0}^{L-1} p_d(x)}. \tag{16}$$

In the presence of a static background in three consecutive frames, there will potentially be a substantial number of locations in δ_n with near-zero values. Parameter m_0 controls the influence of this static background and avoids assignment of high confidence to these near-zero values. Let μ, γ and σ be the mean, median and standard deviation of the distribution of values in δ_n, and m_0 is determined by $m_0 = 0.5(\mu + \gamma) + \sigma$.

Then, we can define a confidence image Conf as

$$\mathrm{Conf}(i,j) = (L-1)p_n(m),\ m = \delta_n(i,j). \tag{17}$$

If $m_0 = 0$, Conf is equivalent to the histogram equalization of δ_n. However, the straightforward histogram equalization is not useful in this context due to the static background. Fig. 3 shows examples of four confidence images extracted for the original images in Fig. 1. In Fig. 3, we can see that dirt pixels appear very bright in the confidence images, which means that they are correctly assigned a high confidence value. On the other hand, there remain some false alarms mainly due to motion and moving edges.

For gray-level images, $h(f_n, f_{n-1}) = f_n - f_{n-1}$, and δ_{n-} and δ_{n+} are forward and backward frame differences. In principle we can convert color images to gray ones for the extraction of confidence images. However, this fails to detect dirt locations of higher difference between color components but lower difference between intensity levels. In addition, δ_n extracted from $h(\cdot)$ has a larger dynamic range than the equivalent signal obtained by considering

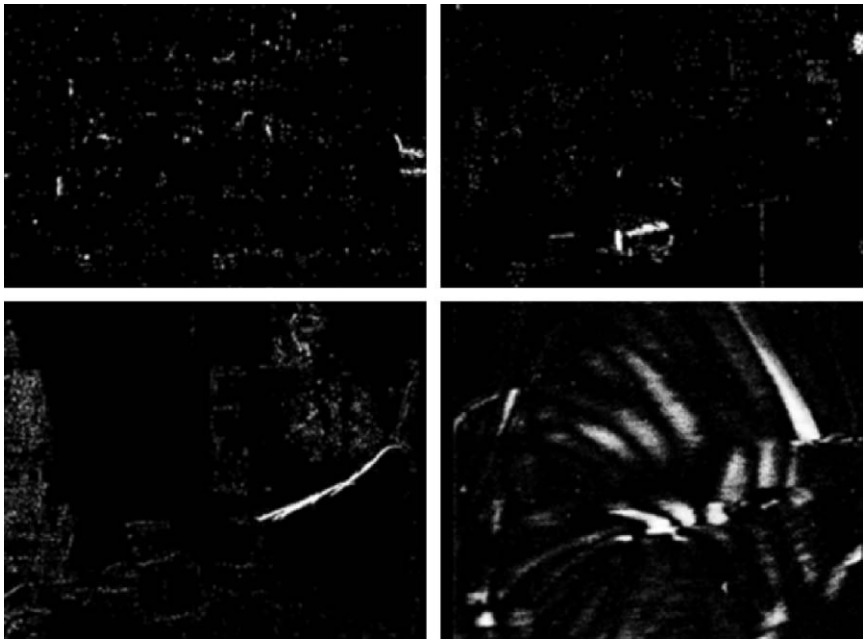

Fig. 3. Extracted confidence images for Fig. 1.

exclusively monochrome information. This has the effect of further emphasizing dirt areas and suppressing false alarms when δ_n is normalized as in (16). Fig. 4 shows one color image and its corresponding luminance component as well as the two corresponding confidence images. Comparing Fig. 4(c) and (d) we can see that the confidence image obtained using $h(\cdot)$ is (1) more accurate, especially in places where dirt has similar intensity to the background and (2) more robust owing to the suppression of false alarms.

4. Confidence-based dirt detection using local region-growing

We assume that the higher the intensity of the confidence image the higher the likelihood of a dirt particle being present. This allows us to subsequently localize regions of dirt via thresholding. Assuming Conf is a confidence image extracted as described in Section 3, we can obtain a binary mask B_0 as

$$B_0(i,j) = \begin{cases} 1 & \text{if } \mathrm{Conf}(i,j) \geqslant T_m, \\ 0 & \text{otherwise,} \end{cases} \tag{18}$$

where T_m is a threshold. This is proportional to a given confidence level C and determined as $(L-1)C$. In general, a larger threshold is used when there are fewer false alarms in the confidence image, and vice versa. However, to obtain an entire dirt cluster in B_0 this confidence level should not be too high, i.e., $<90\%$. After thresholding, dirt will appear as region islands in the binary mask B_0 which can subsequently be labeled and further processed with the intention to detect and remove false alarms.

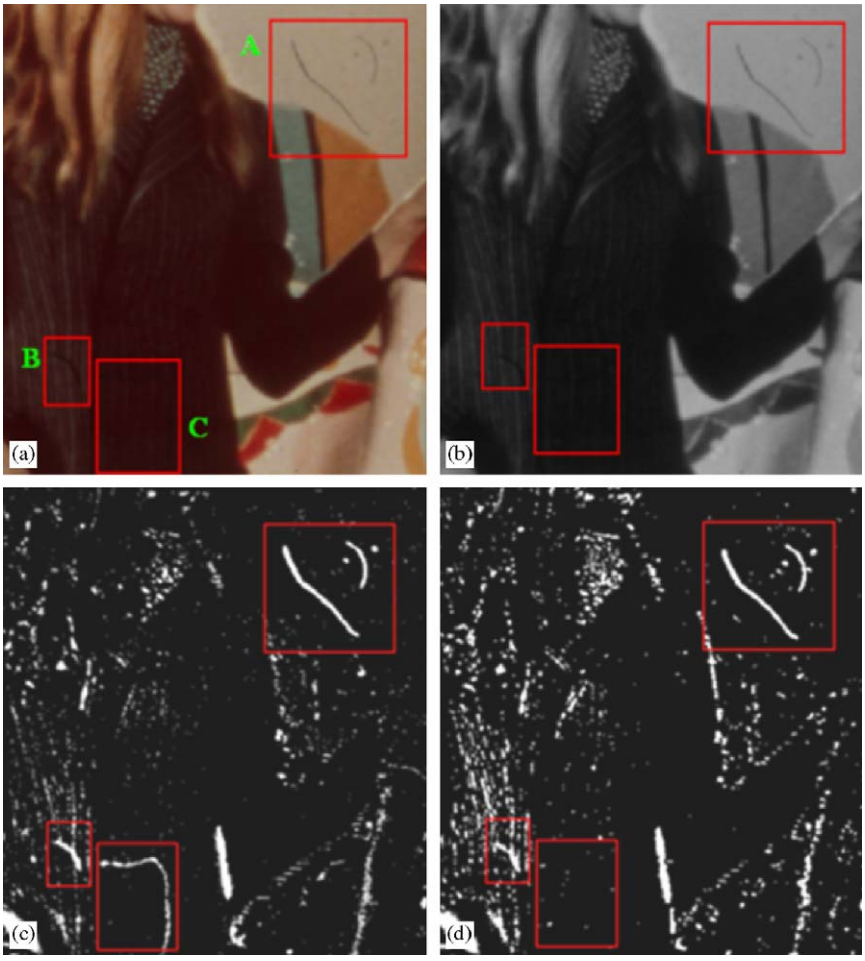

Fig. 4. Comparisons of confidence images extracted from luminance-only and color components, with sample dirt areas shown inside rectangles. (a), and (b) Color and gray image of frame #14. (c) and (d) Extracted confidence images.

Prior to this taking place we ignore all tiny regions of size less than a given threshold, i.e., 3 pixels.

In general, it is occlusions caused by motion which primarily lead to false alarms in the extracted confidence images. This can be easily verified by comparing the confidence image with original frames. The key principle used towards preventing false alarms (that would have otherwise resulted from using the confidence images) is local similarity. More specifically, it is expected that, on the average, false-alarm pixel locations and their neighbors should be strongly correlated i.e. due to intensity continuity in natural scenes. The opposite would be true for pixel locations occupied by genuine dirt for obvious reasons. At locations where abrupt intensity variations occur this assumption will no longer be valid but on the average this is less likely to occur. Local region-growing is therefore employed to exploit this similarity assumption and shape our false alarm detection strategy. It is worth reiterating that local region-growing is more efficient than global region-growing and has a lower computational complexity.

Region-growing is a well-established image processing concept which has been used for various tasks in image analysis (see for example [30]) and typically requires a similarity criterion which controls the merging of adjacent pixels to form a connected set or region. For our purposes, having identified labeled regions r_k as above, on the original frame f_n we compute the corresponding mean and standard deviation vectors per color component as $\mu_k = (\mu_k^{(r)}\mu_k^{(g)}\mu_k^{(b)})$ and $\sigma_k = (\sigma_k^{(r)}\sigma_k^{(g)}\sigma_k^{(b)})$, respectively. We then define a similarity criterion as

$$\left|c_{i,j}^{(n)} - \mu_k^{(c)}\right| \leqslant T_\sigma \sigma_k^{(c)} \tag{19}$$

where $c_{i,j}^{(n)}$ is the intensity value of adjacent pixels (i,j) in image f_n, $c = \mathrm{r,g,b}$. Parameter T_σ is to control the merging of pixels and in our experiments we have used $T_\sigma = 1.5$. It is interesting to note that this local region growing will generate similar results to the work in [23], in which morphological opening and closing are utilized. However, our method is more efficient and avoids substantial false alarms caused by the global nature of application of these operators.

Let n_k be the number of pixels in f_n which are adjacent to r_k. In (19) we assume that n'_k pixels will be merged into r_k, where $n'_k \leqslant n_k$. Next, we identify r_k as a false alarm if it holds true that $n'_k > n_k/T_s$, where T_s is another threshold satisfying $1 < T_s < 3$. If there are false alarms caused by less abrupt edges in an image, a larger T_s will be required. Otherwise, a smaller T_s is more suitable for accurate detection.

After the elimination of all false-alarm regions identified as above we obtain a new mask image, B, which is a subset of B_0. For visualization purposes, we overlay confidence values to each pixel of B, to generate a gray-level output image, $\mathrm{Conf_d}$. Fig. 5 shows detected dirt with attached confidence based on Fig. 3, and Fig. 6 illustrates detected dirt for Fig. 4(a) using the confidence images in Fig. 4(c) and (d), respectively, with $T_m = 155$, $T_s = 2.3$.

5. Experimental results

5.1. Objective ground truth and visual assessment

It is usual practice to assess performance using artificially added dirt such as in the work reported in [1,6] and [17]. In contrast, we are presenting results obtained by using GT dirt maps made available from Institut National de L' Audiovisuel (INA), Paris. As already mentioned such maps are obtained by using special infrared-film scanners and typically show dirt as darker areas set against a lighter background. Scratches are also present in those maps as mid-gray features owing to their partial transparency. We have used broadcast resolution (760×560) sequence 'Lady and doll' which, among other features, contains a fair amount of camera shake, local motion and partially textured background. In order to focus on dirt evaluation, the intensity of original GT is thresholded to obtain a binary GT dirt mask. For the source image in Fig. 4(a), the corresponding GT data in gray-level (original) and binary (thresholded) format are illustrated in Fig. 7(a) and (b), respectively.

Comparing Fig. 7(b) with Fig. 4(c) we can see that the extracted confidence is very consistent with the GT information. If we compare the detection results using our method (see Fig. 6) with those from SDIp, ROD, LUM and ML3Dex (see Fig. 7(c)–(f)) we can clearly see that our method is more accurate with fewer false alarms.

It should be noted that for the implementation of SDIp and ROD we used the well-known Black–Anandan optical flow algorithm to obtain dense motion fields of sub-pixel accuracy [31] for motion-compensation purposes. Since motion estimation is normally implemented on the luminance channel only, these motion-compensated methods like SDIp and ROD will fail in detecting dirt areas of low

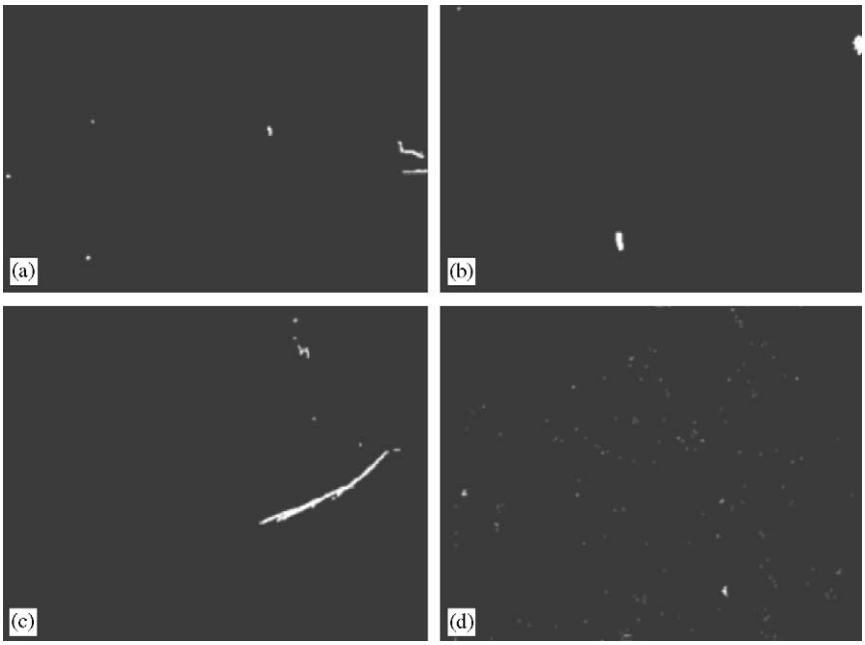

Fig. 5. Detected dirt with overlaid confidence for Fig. 1 using the confidence image in Fig. 4.

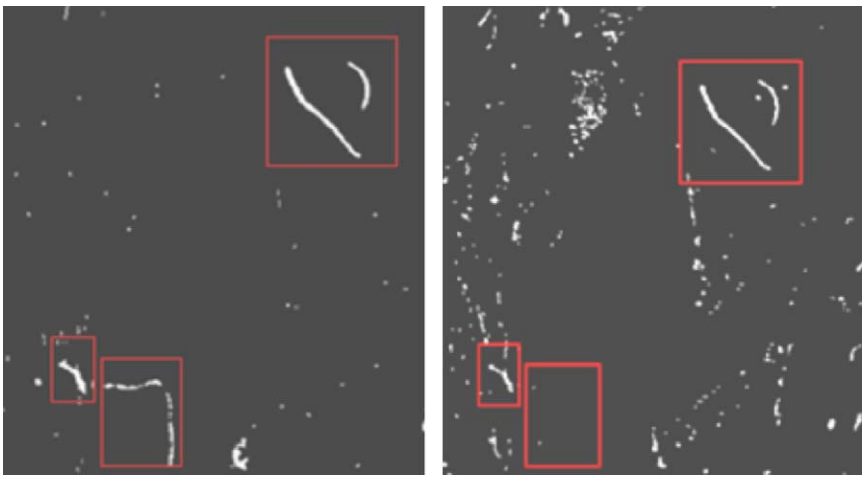

Fig. 6. Detected dirt with overlaid confidence for Fig. 5(a) using confidence from Fig. 5(c) (left) and Fig. 5(d) (right), respectively.

contrast in the corresponding luminance component, such as block C in Fig. 4(a), despite the fact that such dirt is fairly visible in the GT data in Fig. 7(a) and (b).

In addition, it is worth noting that our approach can successfully detect dirt impairments of large size, provided that the corresponding dirt area is represented by a high value in the extracted confidence image. In our method, each dirt region is identified via local region growing without any size restrictions being inherently imposed. Fig. 8 shows two examples of large impairments with source images, ground truth and detection results. The largest regions of dirt in the two images are 48×71 and 50×33 pixels, respectively. From the results we can see that they have been successfully detected by our method.

5.2. *Quantitative assessment*

We present a quantitative performance assessment using the available GT which has been utilized

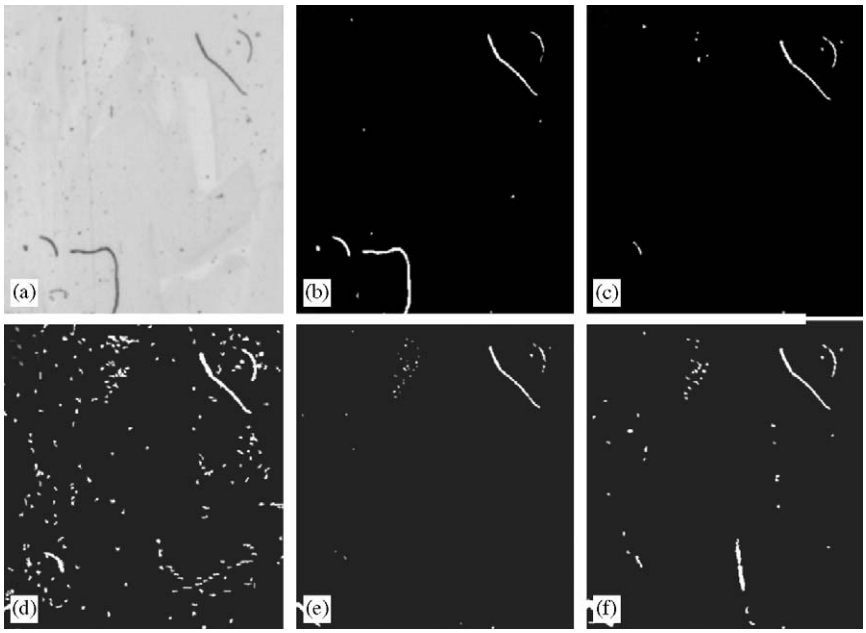

Fig. 7. Comparisons of GT and detected dirt for the image Fig. 5(a). (a) Gray-level GT map of dirt, (b) binary GT of dirt at threshold = 80, (c) detected dirt using SDIp with $t_2 = 10$, (d) detected dirt using ROD, $t_1 = 3$, $t_2 = 6$, $t_2 = 10$, (e) detected dirt using LUM (27,9) with $t_s = 10$ and (f) detected dirt using ML3Dex with $t_s = 10$.

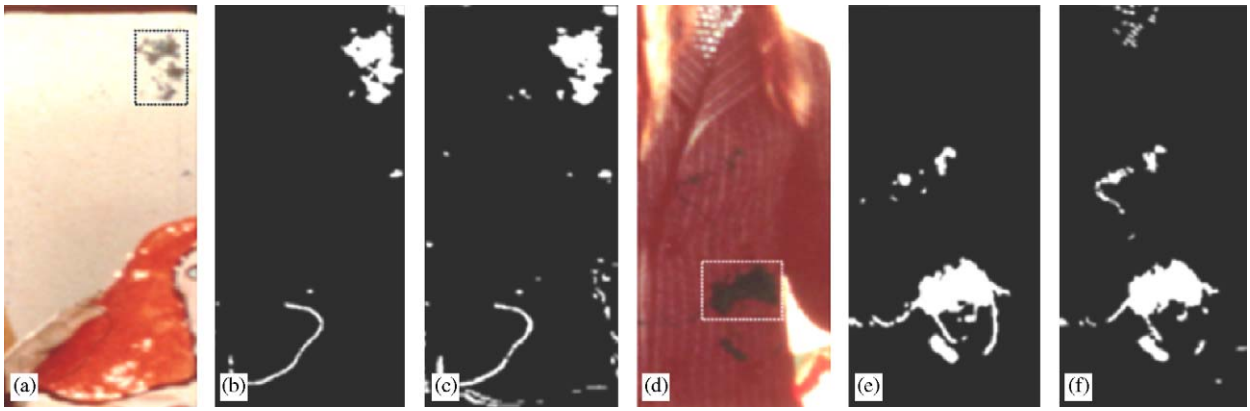

Fig. 8. Two examples of large impairments. (a) and (d) Two original images from frame ♯7 and frame ♯11. (b) and (e) Corresponding binary GT masks at threshold = 95. (c) and (f) Detected dirt using our method.

towards the computation of ROC type of curves [32]. It should be noted that this is the first time that a ROC-based evaluation has been used in such a context.

Let D_g be the binary GT mask of dirt and D_x be a dirt mask detected from any given method, we define the true positive rate R_{tp} and false positive rate R_{fp} by

$$R_{tp} = \frac{\text{Count}(D_x \otimes D_g)}{\text{Count}(D_g)}, \tag{20}$$

$$R_{fp} = \frac{\text{Count}(D_x \otimes \bar{D}_g)}{\text{Count}(\bar{D}_g)}, \tag{21}$$

where Count is a function counting the non-zero elements in a mask and operator $\otimes$ is the logical AND between the two masks. Besides, $\bar{D}_g$ is the complement of D_g.

ROC curves are plots of R_{tp} versus R_{fp}. The points required to determine ROC curves are obtained by varying threshold values i.e., t_1, t_2, t_3, or t_s. Typically, this would be the threshold that

determines whether the difference between the current frame on one hand and a motion-compensated or spatially filtered frame on the other are significant or not. If there are more than one threshold utilized in a particular method (i.e., such as ROD) we first adjust the smallest one and allow the others to change proportionally to it i.e., we take t_1 as the principle value and allow t_2 and t_3 to vary proportionally to t_1 yielding $t_1 < t_2 < t_3$. In our method ROC curves are similarly obtained by thresholding the gray-level output of the detector using progressively increasing confidence levels.

We applied our algorithm to the entire test sequence consisting of 286 frames. All the GT data were thresholded at level 95 to obtain binary masks of dirt. Fig. 9 shows global ROC performance comparing our method (denoted as 'Local Seg') with SDIp, ROD, LUM, and ML3Dex. These ROC curves were obtained by considering the sequence as a single data set in obtaining the corresponding measurements of true positive rate and false positive rate. From Fig. 9 we can see that our method offers the highest performance levels when the true positive rate is larger than 0.6. Spatial filtering methods like LUM and ML3Dex consistently generate the worst results. Less false positive rates may be obtained by SDIp and ROD owing to the suppression of false alarms by motion-compensated processing. However, these methods still produce fairly low true positive rates, especially for dirt locations of similar luminance intensity but different color relative to the background.

5.3. Computational complexity

We assessed all competing methods under consideration in terms of computational complexity.

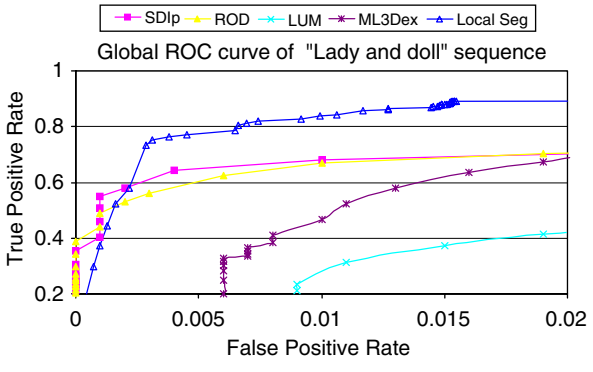

Fig. 9. Global ROC curves from SDIp, ROD, LUM, ML3Dex, and our algorithm (Local Seg).

On our processing engine (Pentium 4, CPU 3.2 GHz, RAM 1 GB), average times to process one frame of the 'Lady and doll' sequence were 53.8 s and 6.2 s, respectively, for motion-compensated filtering (including SDIp and ROD) and non-motion-compensated filtering (like LUM and ML3Dex) compared to 0.06 s for our method. It should be noted that for temporal median filtering, more than 99% of the computing time is spent towards bi-directional motion estimation. The above suggests that our method is nearly 1000 times faster than competing methods using motion estimation and compensation.

6. Conclusions

We have presented a dirt detection method for archived color film sequences using a combination of confidence extraction and local region-growing. One of the most attractive features of our scheme is that a confidence measure is attached to a detected region which supports fine tuning in both automatic and semi-automatic concealment of dirt. Another key feature of our method is its low computational complexity mainly owing to the fact that it does not rely on the use of motion estimation and motion-compensated prediction. We have shown that local region-growing is effective towards reducing false alarms caused by moving edges and we have demonstrated that the proposed method outperforms well-established methods including SDIp, ROD, LUM, and ML3Dex in terms of ROC performance as well as visually.

Acknowledgments

The authors would like to thank the research staff at Snell & Wilcox PLC, Hampshire, UK for valuable discussions and for providing some of the test data. The authors would also like to thank The Institut National de l'Audiovisuel (INA), Paris, France for providing the objective ground truth data. Finally, thanks are due to the anonymous reviewers for their constructive comments to further improve this paper.

References

[1] A.C. Kokaram, On missing data treatment for degraded video and film Archives: a survey and a new Bayesian approach, IEEE Trans. Image Process. 13 (3) (2004) 397–415.

[2] P. Schallauer, A. Pinz, W. Haas, Automatic restoration algorithms for 35 mm film, J. Comput. Vision Res. 1 (3) (1999) 59–85.
[3] A.C. Kokaram, Motion Picture Restoration, Springer-Verlag, Berlin, Germany, 1998.
[4] M.S. Hamid, N.R. Harvey, S. Marshall, Genetic algorithm optimization of multidimensional grayscale soft morphological filters with applications in film archive restoration, IEEE Trans. Circuits Syst. Video Technol. 13 (5) (2003) 406–416.
[5] D. Suter, P. Richardson, Historical film restoration and video coding, in: Proceedings of the Picture Coding Symposium, Melbourne, Australia, 13–15 March 1996, pp. 389–394.
[6] P. Read, M.-P. Meyer, Restoration of Motion Picture Film, Butterworth Heinemann, New York, 2000.
[7] O. Buisson, B. Besserer, S. Boukir, F. Helt, Deterioration detection for digital film restoration, in: Proceedings of CVPR, New York, 17–19 June 1997, pp. 78–84.
[8] R. Storey, Electronic detection and concealment of film dirt, J. SMPTE 94 (June 1985) 642–647.
[9] M.J. Nadenau, S.K. Mitra, Blotch and scratch detection in image sequences based on rank ordered differences, in: Proceedings of the Fifth International Workshop on Time-Varying Image Processing, Florence, Italy, 5–6 September 1996, pp. 27–35.
[10] A. Gangal, T. Kayikçioglu, B. Dizdaroglu, An improved motion-compensated restoration method for damaged color motion picture films, Signal Process.: Image Commun. 19 (4) (2004) 353–368.
[11] L. Tenze, G. Ramponi, S. Carrato, Robust detection and correction of blotches in old films using spatio-temporal information, in: Proceedings of SPIE, vol. 4667, International Symposium on Electronic Imaging, San Jose, California, 20–25 January 2002, pp. 348–357.
[12] M.K. Özkan, A.T. Erdem, M.I. Sezan, A.M. Tekalp, Efficient multiframe Wiener restoration of blurred and noisy image sequences, IEEE Trans. PAMI 1 (4) (1992) 453–476.
[13] D. Geman, G. Reynolds, Constrained restoration and the recovery of discontinuities, IEEE Trans. PAMI 14 (3) (1992) 367–383.
[14] A.C. Kokaram, R.D. Morris, W.J. Fitzgerald, P.J.W. Rayner, Detection of missing data in image sequences, IEEE Trans. Image Process. 40 (11) (1995) 1496–1508.
[15] M.N. Chong, D. Krishnan, An edge-preserving MRF model for the detection of missing data in image sequences, IEEE Signal Process. Lett. 5 (4) (1998) 81–83.
[16] A.C. Kokaram, S.J. Godsill, MCMC for joint noise reduction and missing data treatment in degraded video, IEEE Trans. Signal Process. 50 (2) (2002) 189–205.
[17] E. Abreu, M. Lightstone, S.K. Mitra, K. Arakawa, A new efficient approach for the removal of impulse noise from highly corrupted images, IEEE Trans. Image Process. 5 (6) (1996) 1012–1025.
[18] S. Kalra, M.N. Chong, D. Krishnan, A new auto-regressive (AR) model-based algorithm for motion picture restoration, in: Proceedings of ICASSP, vol. IV, Munich, Germany, 21–24 April 1997, pp. 2557–2560.
[19] R.D. Morris, Image sequence restoration using Gibbs distribution, Ph.D. Dissertation, Cambridge University, Cambridge, UK, 1995.
[20] D. Krishnan, M.N. Chong, S. Kalra, On the computational aspects of Gibbs–Markov random field modelling of missing data in image sequences, IEEE Trans. Image Process. 8 (8) (1999) 1139–1142.
[21] B. Alp, P. Haavisto, T. Jarske, K. Oistamo, Y.A. Neuvo, Median-based algorithms for image sequence processing, in: Proceedings of SPIE, vol. 1360, Lausanne, Switzerland, 2–4 October 1990, pp. 122–134.
[22] S.K. Mitra, T.-H. Yu, A new nonlinear algorithm for the removal of impulse noise from highly corrupted images, in: Proceedings of the IEEE International Symposium on Circuits and Systems, London, 30 May–2 June 1994, pp. 17–20.
[23] E. Decencière, Mathematical morphology and motion picture restoration, in: C. Kotropoulos, I. Pitas (Eds.), Nonlinear Model-Based Image/Video Processing and Analysis, Wiley, New York, 2001, pp. 93–120.
[24] R.C. Hardie, C.G. Boncelet, LUM filters: a class of rank-order-based filters for smoothing and sharpening, IEEE Trans. Signal Process. 41 (3) (1993) 1061–1076.
[25] A. Nieminen, P. Heinonen, Y. Neuvo, A new class of detail-preserving filters for image processing, IEEE Trans. PAMI 9 (1) (1987) 74–90.
[26] H.G. Senel, R.A. Peters II, B. Dawant, Topological median filters, IEEE Trans. Image Process. 11 (2) (2002) 89–104.
[27] G.R. Arce, Multistage order statistic filters for image sequence processing, IEEE Trans. Signal Process. 39 (5) (1991) 1146–1163.
[28] J. Ren, T. Vlachos, Segmentation-assisted dirt detection for the restoration of archived films, in: Proceedings of BMVC, vol. I, Oxford, England, 5–8 September 2005, pp. 359–368.
[29] A. Rares, M.J.T. Reinders, J. Biemond, Complex event classification in degraded image sequences, in: Proceedings of ICIP, vol. I, Thessaloniki, Greece, 7–10 October 2001, pp. 253–256.
[30] R. Adams, L. Bischof, Seeded region growing, IEEE Trans. PAMI 16 (6) (1994) 641–647.
[31] M.J. Black, P. Anandan, The robust estimation of multiple motions: parametric and piecewise-smooth flow fields, J. Comput. Vision Image Understanding 63 (1) (1996) 75–104.
[32] J.P. Egan, Signal Detection Theory and ROC Analysis, Academic Press, New York, 1975.