

Detection of Dirt Impairments from Archived Film Sequences: A Survey and Evaluations

Jinchang Ren and Theodore Vlachos

Abstract ----Film dirt is the most commonly encountered artefacts in archive restoration applications. Since dirt usually appears as temporal impulsive event, motion-compensated inter-frame processing has been widely applied for its detection. However, motion-compensated prediction requires a high degree of complexity and can be unreliable when motion estimation fails. Consequently, many techniques using spatial or spatio-temporal filtering without motion have also been proposed as alternatives. In this paper, a comprehensive survey and evaluation of existing methods is presented, in which both qualitative and quantitative performances are compared in terms of accuracy, robustness and complexity. After analysing these algorithms and identifying their limitations, we conclude with guidance in choosing from these algorithms and also promising directions for future research.

Keywords: Film dirt detection, archive restoration, spatio-temporal filtering, film restoration

Corresponding author:

Dr Jinchang Ren

Digital Imaging Research group
School of Informatics (EIMC)
University of Bradford
BD7 1DP
United Kingdom

Tel. +44-1274-235462

Email. j.ren@bradford.ac.uk

I. Introduction

Film archives of huge stocks of moving pictures represent a unique record of the historical, artistic and cultural development of essentially every aspect of life including sport, politics and entertainment. The emergence of new multimedia and broadcasting outlets has the potential of dramatically improving public access to cultural assets of such unique educational and entertainment value. Against this favourable socio-economic background, the level of public access today is limited by a number of technological factors. Many of the historically significant items are either unavailable in their original format or too fragile to survive any attempt at copy or playback.

In their lifetime, films may suffer various impairments due to environmental hazards such as humidity and dust, chemical instabilities, improper storage and handling practices and even poorly maintained projectors [3-6]. Automatic restoration is a key enabling technology towards facilitating access to historical film archives. By improving baseline picture quality and by reducing the perceptual impact of archive-related impairments 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 [1] [2]. Finally, since restoration processes almost invariably result in the enhancement of semantic content, they are also likely to contribute to more efficient management of pictorial databases and archives. Consequently, lots of interest has been attracted from many broadcasters and media holders along with several high-profile collaborative projects have been funded by EU (European Union), including 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 this paper, our aim is mainly focused on detection of impairments occasionally referred to as 'dirt'. These are among the most commonly encountered artefacts and hence their successful detection is a priority issue in any archive restoration system [3-4, 7]. Dirt may be airborne during film processing, may

be due to fibres from clothing or human tissue, may be due to wear and tear of moving mechanical parts or even due to fragments of the film itself owing to excessive friction as the film moves at speed in contact with mechanical components. Dirt particles can adhere to film at any processing stage like exposure, development, printing, telecine transfer or any of the intervening winding and handling operations. Dirt manifests itself in one of two ways; a dirt particle adhering to negative film will eventually appear as bright when printed or reproduced as a positive image while a dirt particle adhering to positive film will obviously impede the transmission of light and will appear dark.

In general, dirt is a temporally impulsive (single-frame) event, appearing mostly as dark or bright opaque spots of random size, shape and location. It is due to particles which are attached to the film or localised abrasions which occurred during storage or when the film passed through various transport mechanisms [3, 5]. Figure 1 gives some examples of dirt which are shown bounded by bright rectangle or polygon boxes.

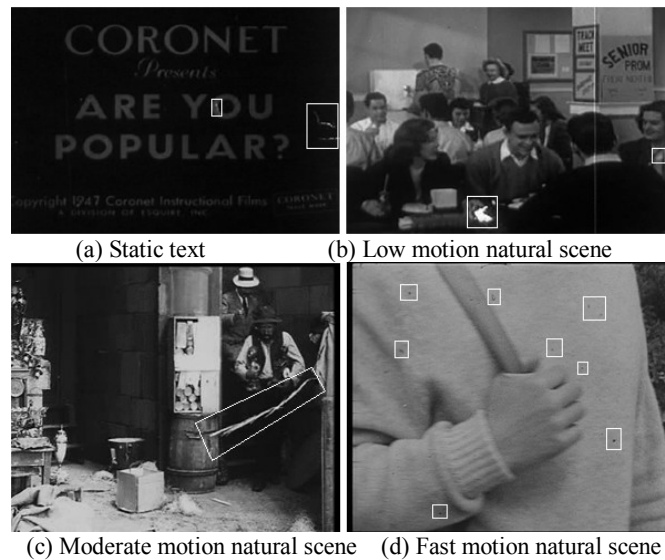


Figure 1: Examples of dirt

Although dirt pixels may be concealed without a detection procedure, i.e. via global filtering of frame images, however, global filtering has the potential to degrade, usually to over-smooth, non-dirt pixels as well. A detection process is used to identify candidate dirt pixels in order that the reconstruction algorithm will concentrate on these areas and reduce errors induced during recovery [5, 8].

Optical processing is employed in conventional (non-computer-assisted) detection of dirt, which requires a specially designed telecine and relies on the transparency of film emulsions to infrared light allowing dirt particles to be detectable by an infrared sensor. This has a number of practical limitations including the fact that detection is not ineffective for particles of small size. The latter is due to limitations of lens performance at infrared wavelengths. It should also be noted that this technique is applicable exclusively to colour film because the silver image in monochrome film is opaque to infrared radiation.

Typically, there are two generic steps involved in any dirt detection algorithm, i.e. inconsistency identification of a pixel in relation to its spatial and/or temporal neighbourhood followed by thresholding. The first step is conventionally implemented by a suitable combination of intra/inter type of filtering. In the second step a well-established principle is that the choice of threshold will influence the balance between detection accuracy and false alarms.

According to how inconsistency is determined, we can categorize current dirt detection methods into three classes, i.e. spatial filtering, temporal filtering and spatial-temporal filtering. In spatial filtering, only intra-frame information is used in filtering of dirt. Although inter-frame (at least three frames) information is utilized in both temporal and spatial-temporal filtering, the former considers only one pixel in each frame, and the latter takes a spatial neighbourhood. Besides, motion-compensated prediction may be contained in these two categories of algorithms.

This remainder of this paper is organized as follows. In Sections 2, spatial filtering approaches using median and morphological filtering are discussed. Section 3 describes temporal filtering from multiple frames. Spatio-temporal filtering with or without motion compensation is given in Section 4. In Section 5, combination of multiple techniques, including a confidence measurement on dirt detection and confidence weighting, are presented. Quantitative evaluations and results are provided in Section 6 and conclusions are drawn in Section 7.

II. Spatial Filtering for Dirt Detection

In spatial filtering, dirt is viewed as a spatial impulse and is typically detected using neighbourhood information in the current frame. Prominent among spatial detection methods are those featuring median and morphological filtering.

A. Median filtering

Standard median filtering is defined on a spatial template, which can preserve edges while remove impulsive noise [9]. Hardie, and Boncelet [10], proposed LUM (lower-upper-middle) filters, in which two parameters are utilised for adjustable smoothing and sharpening of images. Nieminen *et al* [11] presented a multi-level median filter (MLF) to reduce the influence of outlier values while preserving edges. Their filter firstly calculates separate median values for horizontal, vertical, and two diagonal transects in a given window, and then the minimum and maximum of these four values are found. The median of the minimum, maximum, and original raster value in the central of the window is taken as the output of the filter. Abreu *et al* [12] introduced rank-ordered mean (ROM) filter to remove impulsive noise from highly corrupted images. However, only noise of small size can be recovered due to 3×3 neighbouring window used. Senel *et al* [13] proposed a topological median filter to extract edges in noise; however, the filtered images are of unacceptable visual quality in most cases.

In both standard spatial median filtering (SSMF) and LUM, for each pixel (i, j) in the current frame f_n , a window W of radius r is defined as $W(i, j, r) = \{f(i_1, j_1)\}$, $|i_1 - i| \leq r, |j_1 - j| \leq r$, hence the total number of pixels in W is $N = (2r + 1)^2$, and we also define $N_0 = (N + 1) / 2$. We denote $W = \{x_1, x_2, \dots, x_N\}$, and the rank-ordered set is given by $x_{(1)} \leq x_{(2)} \leq \dots \leq x_{(N)}$. The central pixel in the original current frame and filtered image are denoted as x' and y' , respectively. In SSMF, we simply have $y' = x_{(N_0)}$. In LUM, two parameters, k and l , are introduced for smoothing and sharpening, respectively, where we have $1 \leq k \leq l \leq N_0$. Then, the filtered output is defined as

$$y' = \begin{cases} x_L, & \text{if } x' \leq (x_L + x_U) / 2 \\ x_U, & \text{otherwise} \end{cases} \quad (1)$$

where x_L and x_U are the corresponding outputs of the smoothing and sharpening processes given by

$$x_L = \text{median}\{x_{(k)}, x', x_{(l)}\} \quad (2)$$

$$x_U = \text{median}\{x_{(N-k+1)}, x', x_{(N-l+1)}\} \quad (3)$$

Let g_n be the output image after filtering, i.e. $g_n(i, j) = y'(i, j)$. Dirt mask D is then determined as

$$D(i, j) = \begin{cases} 1 & \text{if } |g_n(i, j) - f_n(i, j)| > t_s \\ 0 & \text{otherwise} \end{cases} \quad (4)$$

where t_s is a predefined threshold.

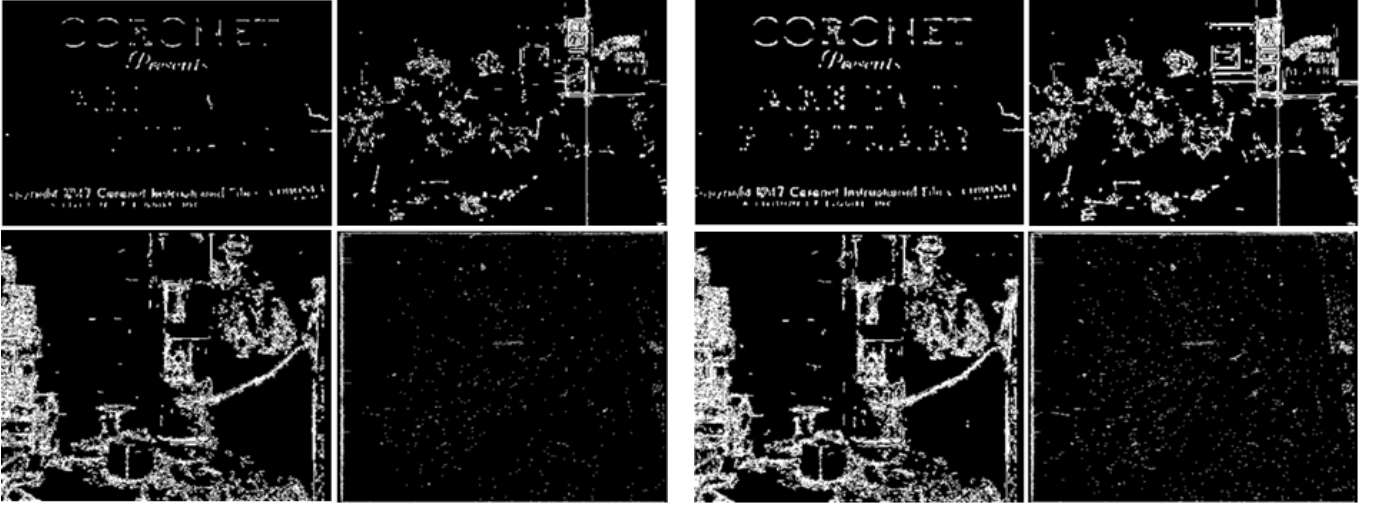


Figure 2: Dirt detected from Figure 1 by SSMF with $t_s = 10$ and window of 5×5 (left) and 7×7 (right), respectively.

With SSMF and LUM, detection performance is very sensitive to the size and shape of window W . A square window is usually adopted when we have no a priori information suggesting otherwise. With reference to the original images in Fig 1, Figure 2 and Figure 3 show detected dirt by SSMF and LUM using different window sizes, where $t_s = 10$, $k = 5$, $l = 10$ and $r = 2, 3$, respectively.

Comparing Fig 1, Fig 2 and Fig 3, we can conclude that: 1) SSMF performs better than LUM, though it generates more false alarms especially near sharp edges. However, SSMF is only suitable for a pre-detection, and additional processing is necessary to cope with false alarms; 2) larger windows are capable of detecting larger areas of dirt, however, more false alarms are also introduced.

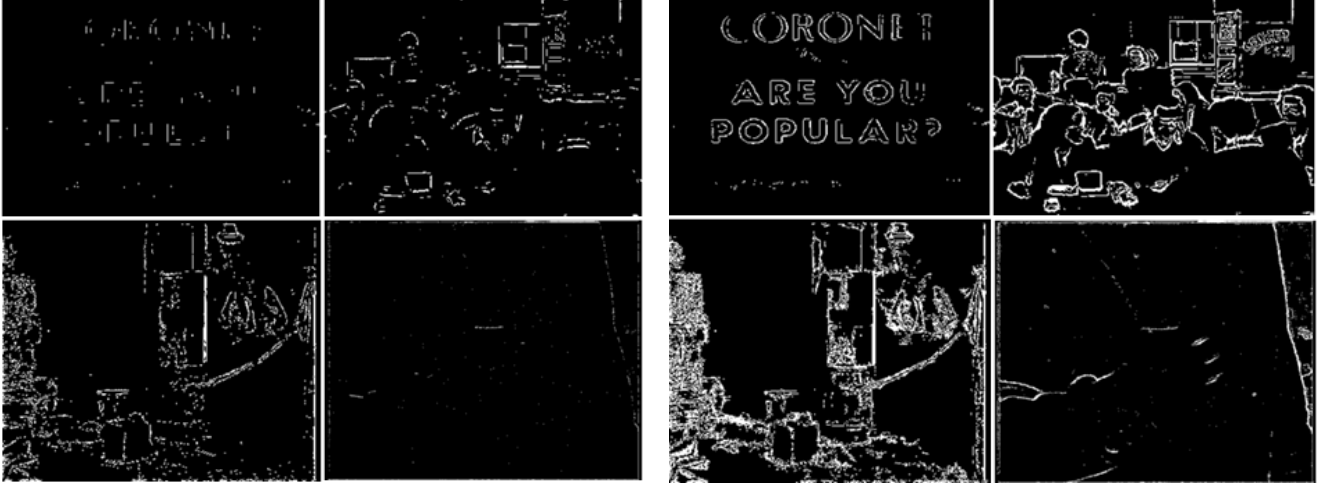


Figure 3: Dirt detected from Fig 1 by $LUM(5,10)$ with $t_s = 10$ and window size of 5×5 (left) and 7×7 (right), respectively.

B. Morphological filtering

As for morphological filtering, the top-hat and black-hat operators on image f are defined below [14]:

$$Top_hat(f) = f - (f \circ B) \quad (5)$$

$$Black_hat(f) = (f \bullet B) - f \quad (6)$$

where \circ and \bullet denote morphological opening and closing operators, and B is a structure element. The binary mask of dirt is then obtained by thresholding the filtered images. In Buisson *et al* [15], top-hat and black-hat are used to detect bright and dark pixels of dirt, respectively.

It is obvious that the size and shape of B is the very important parameter in both top-hat and black-hat operators. Again a rectangle template of B is usually used when there is no prior information. For the four images in Figure 1, we choose top-hat for the first three and black-hat for the last one, i.e. Fig 1(d), and show the detected dirt in Figure 4. A rectangle structure element of B with size of 3×3 and 5×5 is utilized, respectively, and the threshold to obtain a binary mask of dirt is set as 20.

From Figure 4 we can see large size of dirt have been successfully detected, although many false alarms are also flagged. Nevertheless, the overall accuracy shown in Figure 4 seems no superior to SSMF in Figure 2. Moreover, larger size of B usually cause to massive false alarms due to merging of objects in morphological closing operator.

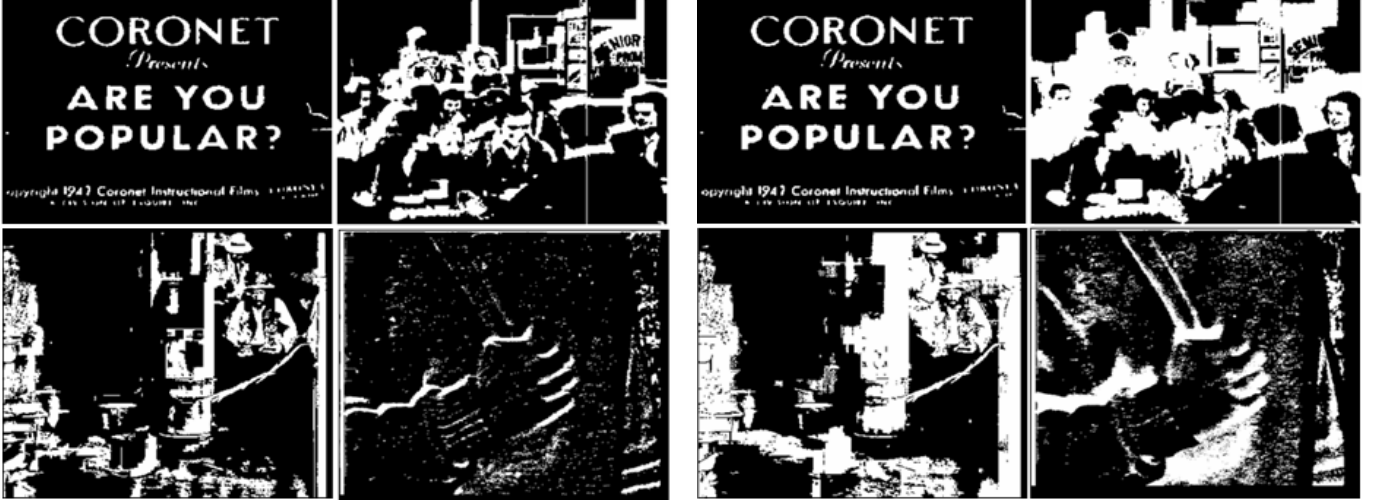


Figure 4: Two groups of detected dirt from Fig 1 with threshold $t_s = 20$ and structure element in 3×3 (left) and 5×5 (right), respectively, where top-hat is applied for Fig. 1(a-c) and black-hat for Fig. 1(d).

However, the difficulty is how to automatically choose either of these two operators, as they perform differently in bright or dark pixels of dirt as shown in Figure 5. This becomes more difficult when both bright and dark dirt exists; hence additional processing such as combining with other detectors is required [15-18]. In other words, morphological filtering is not an independent detector and hence it will not be evaluated separately.



Figure 5: Results from Figure 1(b) using black-hat (left) and top-hat (right) with 3×3 rectangle B and threshold $t_s = 35$.

III. Temporal Filtering for Dirt Detection

In temporal filtering approaches, dirt is viewed as a temporal impulse (single-frame incident) and hence treated by inter-frame processing by taking into account at least three consecutive frames, with motion-compensation [4, 8]. In fact, the concept of motion-compensated temporal filtering for noise removal in image sequences can be found in the later 1970's or early 1980's [19-22].

Storey's work was perhaps the earliest contribution to the electronic detection and concealment of dirt [23]. In his hardware-based system, 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. Kokaram extended this idea using motion-compensated differences [3, 8, 24], in which the so-called "Spike Detection Index" (SDI) was proposed. The basic SDI detector, SDIa, was based on the identification of high absolute differences between the current frame and two compensated images, and an expanded SDI detector, SDIp, additionally required sign consensus of the two differences above. Because of this additional constraint, SDIp is considered better than SDIa in almost all situations [3]. Schallauer *et al* [4] proposed a double-threshold temporal median filtering (DTMF), in which 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.

Let C_{n-} and C_{n+} denote two motion compensated frame neighbours of the current frame f_n , and we also define $D_{n-} = f_n - C_{n-}$ and $D_{n+} = f_n - C_{n+}$ as the differences between f_n and each of these two images. With a given threshold t_2 , possible dirt pixels are defined by

$$D_t = \begin{cases} 1 & \text{if } |D_{n-}| > t_2, |D_{n+}| > t_2 \\ 0 & \text{otherwise} \end{cases} \quad (7)$$

Then, further processing to filter false alarms is introduced in (8) and (9) before detection of dirt pixels in DTMF and SDIp, where t_1 is another threshold satisfying $t_1 < t_2$.

$$D_{DTMF} = \begin{cases} 1 & \text{if } D_t = 1, |C_{n-} - C_{n+}| < t_1 \\ 0 & \text{otherwise} \end{cases} \quad (8)$$

$$D_{SDIp} = \begin{cases} 1 & \text{if } D_t = 1, D_{n-} D_{n+} > 0 \\ 0 & \text{otherwise} \end{cases} \quad (9)$$

Figure 6 and Figure 7 show detected dirt using SDIp and DTMF with $t_1 = 10$ and $t_2 = 20$. Motion compensation was implemented by using dense motion fields of sub-pixel accuracy from the well-known

Black-Anandan optical flow algorithm [25]. These two figures confirm the fact that DTMF and SDIp generate comparable results, but their performances are sensitive to the availability of dirt-free background.

False alarms produced by these two methods are chiefly due to failure of motion estimation, see for example the results from Figure 1(c) and 1(d) in both Figure 6 and Figure 7. Generally, the performance of these methods is dependent on not only the accuracy in motion-compensated prediction but also the selected thresholds. Large or small threshold(s) will certainly lead to more or less accuracy and false alarms. With a given threshold t_2 , DTMF and SDIp will produce the same results, D_t , despite of the fact that different strategies are used to filter false alarms, i.e. $|C_{n-} - C_{n+}| < t_1$ in DTMF and $D_{n-}D_{n+} > 0$ in SDIp, respectively. Containing with simply one threshold, SDIp seems a better solution as it is more controllable.

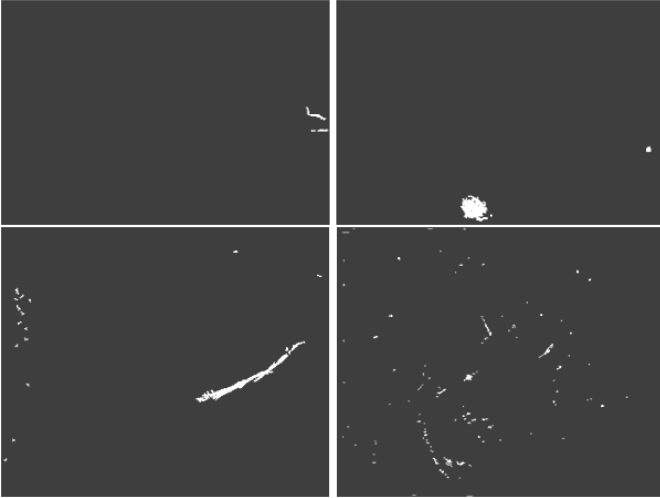


Figure 6: Dirt detected by DTMF ($t_1 = 10, t_2 = 20$).

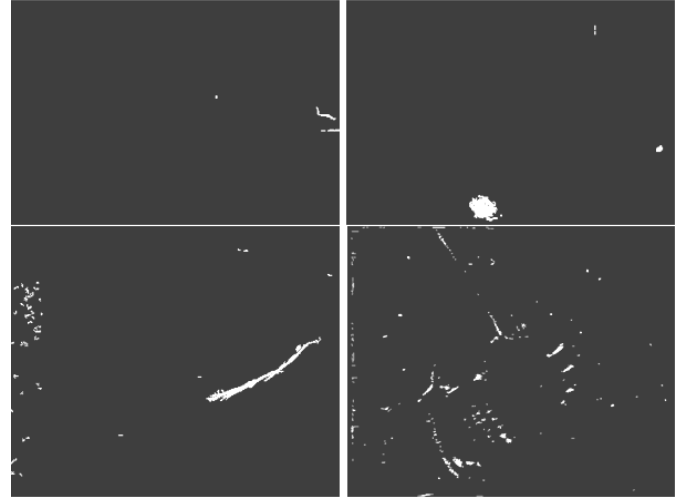


Figure 7: Dirt detected by SDIp with thresholds $t_2 = 20$.

IV. Spatial-temporal Filtering for Dirt Detection

In spatio-temporal filtering, pixel inconsistency is determined by the examination of both spatial and temporal neighbourhoods. In principle, any spatial-temporal method can be extended to motion-compensated filtering. However, those without motion-compensation prove to be good alternatives for more efficiency.

A. Non-motion-compensated techniques

Alp *et al* [26] introduced the so-called ML3D algorithm, in which two groups of windows are defined in three frames and their median values are determined. Then, the median of the two median values and the value of current pixel is taken as the output. Arce [27] applied a multi-stage order statistic filter approach (MOS) extending MLF to three consecutive frames, to noise suppression of image sequences. At the same time, Arce also proposed a three-frame LUM variant for image smoothing purposes, which is denoted as LUM' . In [5], Hamid *et al* employed soft morphological filtering (SMF) in three consecutive frames. However, SMF seems impractical for most applications because it needs sufficient representative samples of dirt for training to optimise the size and shape of the filters.

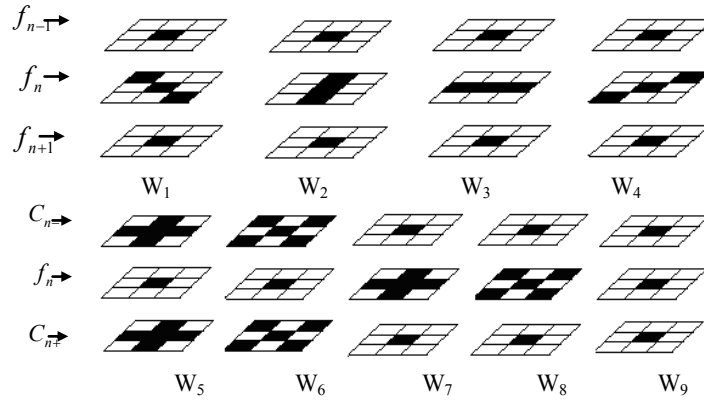


Figure 8: Sub-windows defined in three frames for bi-directional MOS filtering (W_1 - W_4) and ML3Dex filtering (W_5 - W_9).

Regarding MOS, the bi-directional variant used in our experiments is defined on the basis of four sub-windows in three consecutive frames [27], W_1 to W_4 , as shown in Figure 8. Firstly, the median of each sub-window is obtained as z_l , where $l \in \{1, 2, 3, 4\}$, and then z_{\max} and z_{\min} are defined as the maximum and minimum values among these four z_l values. Finally, the output of the filter is determined by

$$y' = \text{median}[z_{\max}, z_{\min}, x'] \quad (10)$$

As for $LUM'(N, k)$, it is defined as an extension of LUM applying to a $3 \times 3 \times 3$ spatial-temporal window [27], where $N = 27$ and $k \leq 13$. Firstly, a rank-ordered set of these pixels is obtained as $x_{(1)} \leq x_{(2)} \leq \dots \leq x_{(N)}$, then the output of the filter is given by

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

In Figure 9 and Figure 10, we show the dirt detection results using MOS and $LUM'(27,9)$, respectively. Owing to temporal information, most of the false alarms that occurred from processing Fig 1(a) and Fig 1(b) are avoided. However, a number of false alarms remain for Fig 1(c), while additional ones occurred for Fig 1(d). Moreover, LUM' seems superior to MOS due to the fact that it takes more pixels from neighbouring frames in filtering. However, this may also cause more false alarms. This demonstrates that the performance is also very sensitive to the designed shape of sub-windows.

In general, MOS fails in such a context, and LUM' performs better only when there is no moving edge in the images. Furthermore, if we compare the results here to those from spatial filtering like SSMF, we may find that SSMF provides better accuracy but also more false alarms.

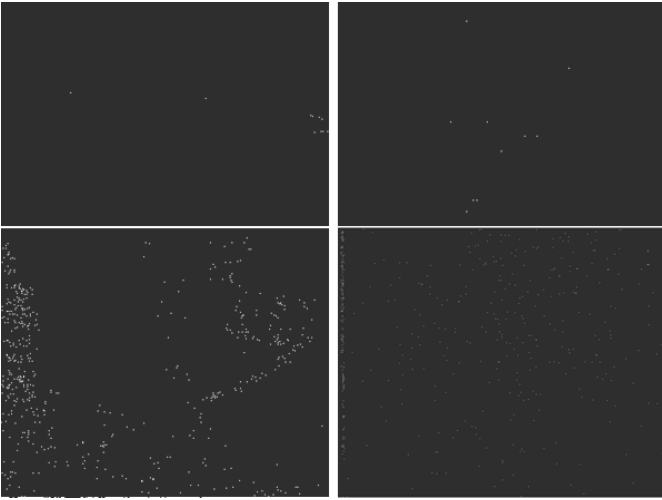


Figure 9: Dirt detected by MOS with $t_s = 10$.

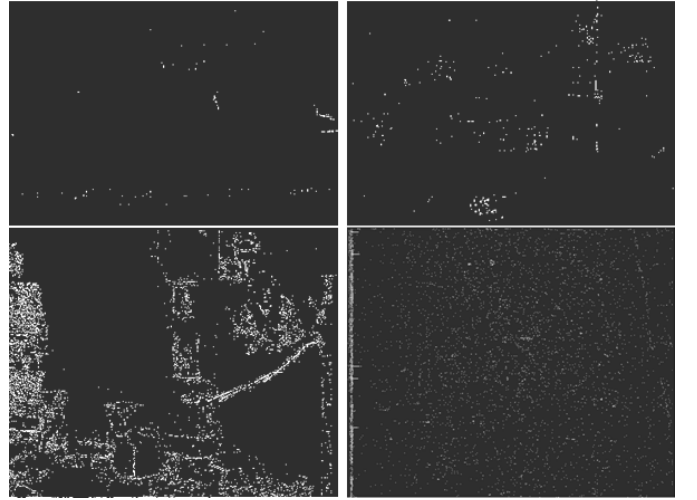


Figure 10: Dirt detected by $LUM'(27,9)$ with $t_s = 10$.

B. Motion-compensated algorithms

Nadenau, and Mitra [28], have used a ranked-order detector (ROD), which examines 7 pixels in three frames against three thresholds. In fact, it is an extension of the work in [12]. Gangal *et al* extended ROD to five frames to improve accuracy in heavily corrupted images or occluded blotches [7]. In Biemond *et al* [29], a simplified version of ROD with only one threshold is proposed, but complex post-processing is required to reduce false alarms. In [8], Kokaram presented an extended version of ML3D, ML3Dex, which applied ML3D filtering to five groups of windows with motion compensation (see Figure 8 from

w_5 to w_9). In [30], a more complex model using adaptive weighted average along the motion trajectory is presented.

In the framework of motion compensation, dirt can be also detected using model-based approaches, such as Wiener filtering, AR (auto-regressive), MRF (Markov random field), Gibbs distribution, and Gibbs-Markov random fields [24], [31-37]. The determination of a MRF prior allows the detection of dirt in a Bayesian framework [3], [38-39]. Since definitive statistical models are difficult to obtain, all the above methods have occasional constraints and will fail if the underlying statistical modelling assumptions cannot be satisfied or if accurate and robust motion compensation cannot be achieved [24], [40]. Most importantly, even with complex implementation, model-based approaches like MRF yield minor improvements compared with SDIa (using temporal filtering) as reported in [24].

To determine dirt in ROD, three pair of pixels at $(i-1, j)$, (i, j) and $(i+1, j)$ are extracted from C_{n-} and C_{n+} . These six pixels are sorted in increasing order in a list $[r_1, r_2, \dots, r_6]$ where r_6 is maximum. Then the median of the list is extracted as $med = (r_3 + r_4) / 2$. If $f_n(i, j) > med$, it refers to a potential bright dirt and define $e_k = f_n(i, j) - r_{7-k}$, otherwise $e_k = r_k - f_n(i, j)$ for dark dirt, where $k \in \{1, 2, 3\}$. Dirt is then detected if we have any e_k greater than t_k .

$$D_{ROD} = \begin{cases} 1 & \text{if } \vee (e_k > t_k) = 1, k \in \{1, 2, 3\} \\ 0 & \text{otherwise} \end{cases} \quad (12)$$

Figure 11 and Figure 12 shows detected dirt using ROD and ML3Dex, and the thresholds in these two methods are $t_1 = 10$, $t_2 = 20$, $t_3 = 30$ and $t_s = 10$. The Black-Anandan algorithm is also utilized here for motion compensation. From the two figures we can see that the two methods have overall comparable results, although slightly fewer false alarms are found from ROD in Figure 11. If we compare the results here to those from SDIp and DTMF (see Section III), we can find two differences: 1) All the four methods have similar detection accuracy, however, spatial templates in ROD and ML3Dex intend to extract extended boundary of dirt regions, especially in the results from Figure 1(b); 2) Although failure

of motion compensation may cause false alarms in all these methods, owing to spatial filtering more false alarms are produced by ROD and ML3Dex along moving edges.

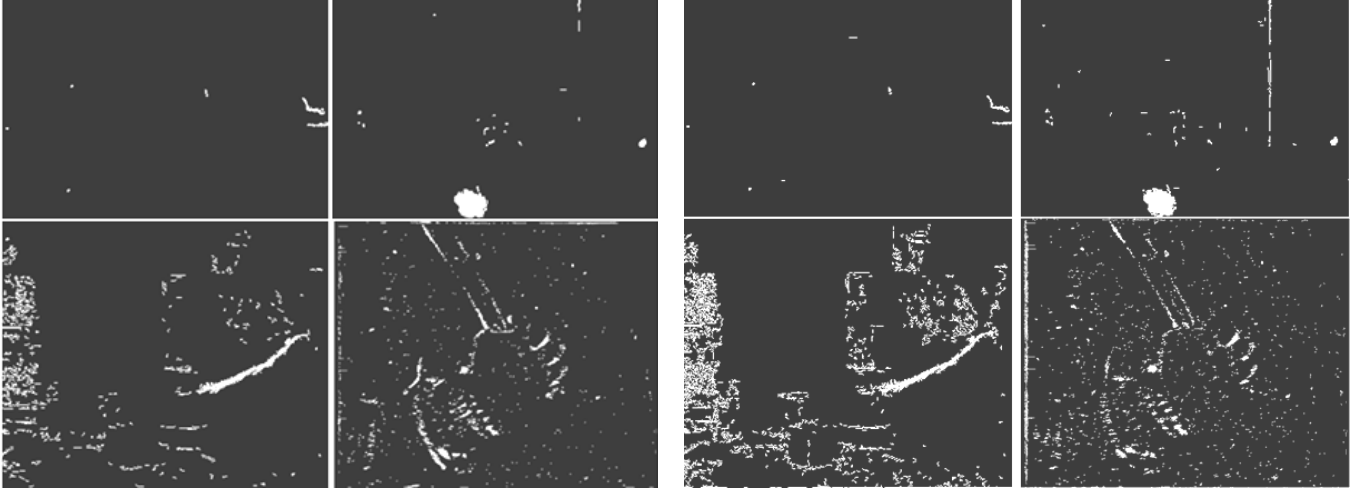


Figure 11: Dirt detected by ROD ($t_1 = 10, t_2 = 20, t_3 = 30$).

Figure 12: Dirt detected by ML3Dex with thresholds $t_s = 10$.

V. Combining Multiple Techniques

Approaches without motion-compensation cannot easily distinguish between genuine dirt clusters and moving objects of a similar spatial structure and also fail when such clusters exceed the filter size. Motion-compensated approaches on the other hand perform poorly when motion cannot be accurately estimated. Consequently, many people also proposed methods to combine different approaches together towards to more accuracy and fewer false alarms. In this section, existing schemes combining different techniques are discussed. In particular a confidence measurement is introduced, for it has great potential to be combined with many other algorithms.

A. Combined schemes

As we mentioned in Section II(B), morphological filtering is not independent, hence additional processing is always required to remove false alarms. In Buisson *et al* [15], a combination of morphological spatial filtering and motion-based dynamic detector is employed. The final result is obtained as logical AND of the results from the two standalone algorithms. In Ferrandiere [17],

morphological closing and opening are employed in spatial-temporal domains, where some post-processing and global motion compensation is required to filter false alarms. Fusion of morphological filtering and a simplified ROD detector can even be found in [41].

Morris and Fitzgerald [42] presented MRF-based motion field segmentation for motion-compensated interpolation of dirt areas. In [58], Markov modelling using HMM and MRF has been applied in detection of all film defects including dirt and sparkle, scratches, and salt and pepper like noise, where spatial and temporal constraints are employed in additional two stages to remove false alarms. In Rares *et al* [43], complex event areas (where motion estimation fails) is firstly detected, from which dirt and other blotch artefacts are then classified through matching of segmented regions in consecutive frames. In [56] and [59], similar concept is applied and namely as moving edge detector and pathological motion for blotch detection and missing data recovery, respectively.

In addition, recent progress using combined schemes in archive restoration for scratch detection and concealment can be found in [52, 53, 54], and reported work on detection of blotches including both dirt and sparkles can be found in [55-57, 59]. In general, scratch and blotch are significantly different from film dirt. Scratch intends as lines appearing at the same location in subsequent frames of the image sequence, caused by spurious particles present during the sequence acquisition phase or in the transport mechanism of the equipment used for the development of the film [52, 53]. Sparkle and dirt, on the contrary, normally appear in a single frame, i.e. a temporal impulsive defect. Dirt can be seen as opaque or semi-transparent clusters with random size, shape and position, caused by dust and dirt stuck on the film, while sparkle are white clusters, caused by the local abrasion of film gelatine [57]. Due to these differences, relevant models and methods for scratch and sparkle detection cannot be directly applied for dirt detection.

In Ren and Vlachos [44], dirt is detected on the basis of segmented frame images, and then a confidence measurement is proposed for validation and removal of false alarms. This confidence-based validation is also utilized in their proposed adaptive spatial-temporal filtering [45], in which either spatial

(such as SSMF or LUM) or motion-compensated filtering like SDIp is applied to each separate image block depending on error residuals after their filtering. As this confidence measurement provides a direct clue as dirt, it can be used to remove false alarms in many existing methods [44-47]. The way how this confidence measurement defines is summarised below.

B. Confidence measurement

On the basis of raw forward and backward frame differences, a confidence measurement of dirt, $Conf_n(i, j)$, is derived by removing static background from a combined measurement of these raw differences [44-47], in which $Conf_n(i, j) \in [0, 1]$ represents the likelihood that pixel (i, j) is dirt in frame f_n . Figure 13 provides a visual impression of the estimated confidence for the material shown in Figure 1, from which we can see that the results are, on the average, intuitively correct with actual dirt pixels being detected with high confidence and hence appearing bright and vice versa. This is an invaluable feature towards both automatic and operator-assisted semi-automatic concealment of dirt as it allows performance fine tuning according to preference especially towards achieving a desirable balance between false alarm and correct detection. As real dirt samples are considered of a hard core but soft boundary [5], this greylevel confidence measurement is useful in representing such a structure.

For a given confidence value $q \in [0, 1]$, a binary mask of dirt can be obtained by thresholding $Conf_n$ using q , and a suitable q is usually decided as a trade-off between desired accuracy and false alarms. According to the confidence images in Figure 13, Figure 14 shows binary masks of dirt obtained with $q = 0.80$ and $q = 0.90$, respectively. It is easily to find out that the overall accuracy and false alarms are comparable to those from motion-compensated spatio-temporal filtering, such as ML3Dex and ROD. Most of the false alarms, especially those from Figure 1(d), are due to fast motion. In addition, higher value of q will lead to fewer false alarms, and vice-versa.

To remove false alarms in the confidence images, local similarity is usually considered based on the assumption that false alarms should be within homogenous regions due to intensity continuity of natural

scenes. This assumption even holds true for occluded areas [49], hence image segmentation like region-growing seems a good solution for this purpose [44, 47]. Besides, false alarms can be also filtered in a combined scheme as discussed below, in which confidence weighting is utilized to the results from other detectors.

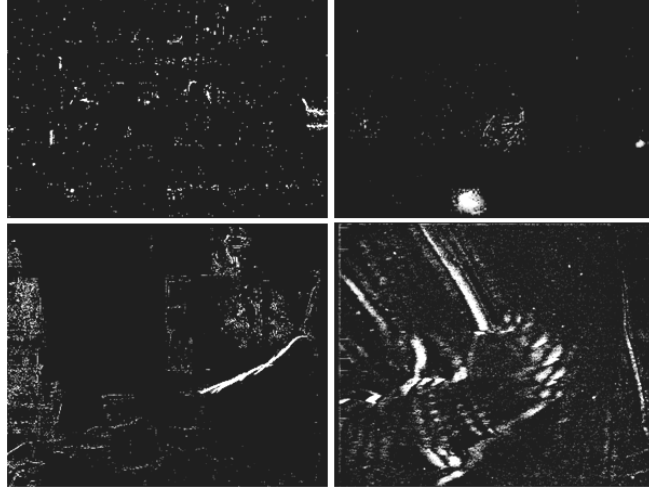


Figure 13: Confidence of dirt extracted for the four images in Figure 1.

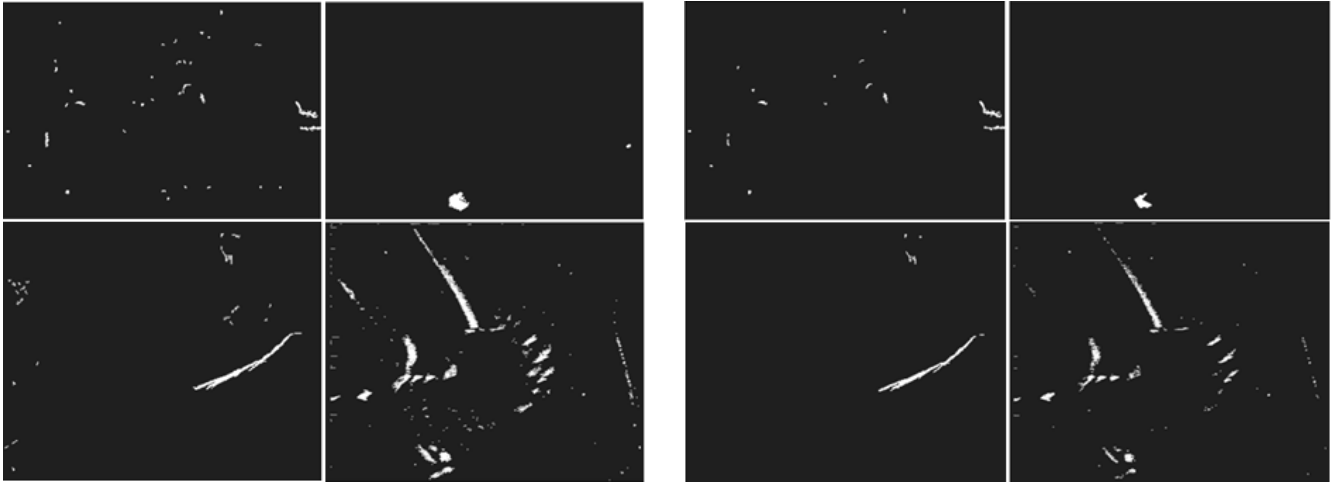


Figure 14: Binary masks of dirt by thresholding confidence images of Figure 13 using $q = 0.80$ (left) and $q = 0.90$ (right), respectively.

C. Confidence weighting in a combined solution

For an estimated binary mask of dirt, D , a confidence weighted result D' can be further defined as

$$D'(i, j) = \begin{cases} (L-1)Conf_n(i, j) & \text{if } D(i, j) \neq 0 \\ 0 & \text{otherwise} \end{cases} \quad (13)$$

Moreover, if D is a greylevel image, the validated results is simply defined as

$$D'(i, j) = Conf_n(i, j) \cdot D(i, j) \quad (14)$$

Essentially D' is a greylevel image whose non-zero values correspond to dirt particles with an associated confidence value. Again, a final binary mask of dirt D_c can be obtained by thresholding D' under a given confidence value $q \in [0,1]$.

$$D_c(i, j) = \begin{cases} 1 & \text{if } D'(i, j) \geq (L-1)q \\ 0 & \text{otherwise} \end{cases} \quad (15)$$

It should be obvious that confidence weighting does not improve detection accuracy but can be used as a mechanism for controlling false alarms. Figure 15 is an example to show the effect of confidence weighting on the results from morphological filtering, and original results before weighting can be found as the left column in Figure 4. Obviously the results have been significant improved by removing most of the false alarms.

Comparing Figure 15 with Figure 14 and Figure 7, we can find that confidence weighting in the combined scheme with spatial morphological filtering yields high accuracy and fewer false alarms. In terms of both accuracy and false alarms, the combined scheme has better or comparable performance than ROD and ML3Dex or to SDIp and DTMF, respectively. However, it does not require expensive motion-compensation.

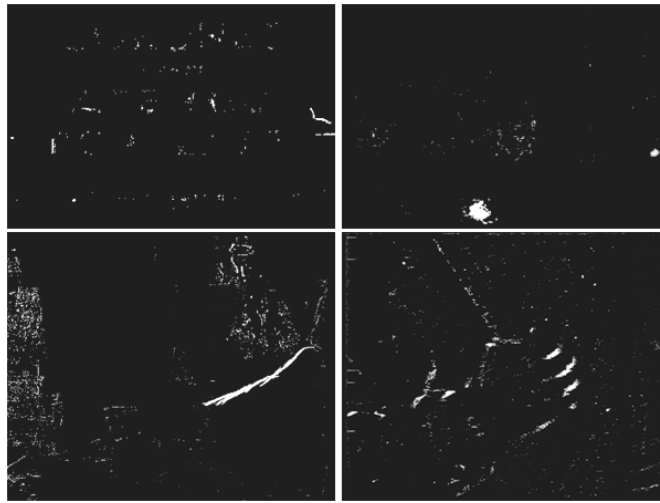


Figure 15: An example of confidence weighting on the results from morphological filtering, and the results before weighting can be found in Figure 4 (3*3 window).

VI. Evaluation and Discussion

Although qualitative evaluations can be achieved by visual assessment of detected dirt, this is insufficient as the results are dependent on certain pre-defined thresholds/parameters. Consequently, quantitative evaluation is always desired in such a context, and relevant criteria and results will be discussed in detail in this section.

A. Ground truth and evaluation criteria

In some approaches, such as Hamid *et al* [5], a corresponding ideal (dirt-free) reference sequence is required in order to carry out an evaluation by comparing recovered images with reference images using the mean absolute error (MAE) criterion. However, it fails to accurately show the performance of detection as the recovered quality also relies on the relevant reconstruction algorithms. Yet, in practice such a reference is unavailable. Therefore, ground truth maps of dirt are necessary for such purposes.

To obtain these ground truth maps, traditional approaches usually involve artificially added dirt in either grey or B/W (Black/White) format such as the work reported in [3], [6] [12]. Although good performances may be achieved with these manual degradations, especially in those model-based approaches, they almost certainly will not reflect the real situations hence may lead to unreliable results and evaluations. In Ren and Vlachos [44-46], binary masks of ground truth maps are manually obtained from real degraded sequences. However, it is difficult to accurately define the boundary of a dirt cluster due to the fact that dirt contains a hard core but soft boundary [5]. Consequently, these masks are inaccurate in representing such a structure.

In contrast, we are presenting evaluation results using objective ground truth maps of dirt made available from INA (Institut National de L' Audiovisuel, Paris). These ground truth maps are greylevel images obtained by infrared scanning of archived films, in which several physical defects are located, such as blotches and scratches. Due to non-transparency or semi-transparency in the films, these defects usually appear dark or grey in the scanned images. Darker a defect pixels is, more likely it refers to a

pixel of dirt. Consequently, the grey levels in an inverted ground-truth image show (probabilistic or likelihood) intensity of dirt.

Three sequences of broadcast resolution (720x576 pixels) are used in our evaluation, namely “cigarette”, “door” and “lady and doll”, and the number of frames tested from these sequences is 261, 255 and 283, respectively. An example of objective ground-truth from the sequence “lady and doll” is shown in Figure 16, compared with corresponding block of source image (enhanced for better contrast). At the same time, two binary masks of dirt obtained at different thresholds are also given. To ignore the effect of scratch and other semi-transparent artifacts while keeping dirt particles, we threshold the ground truth images using 85.

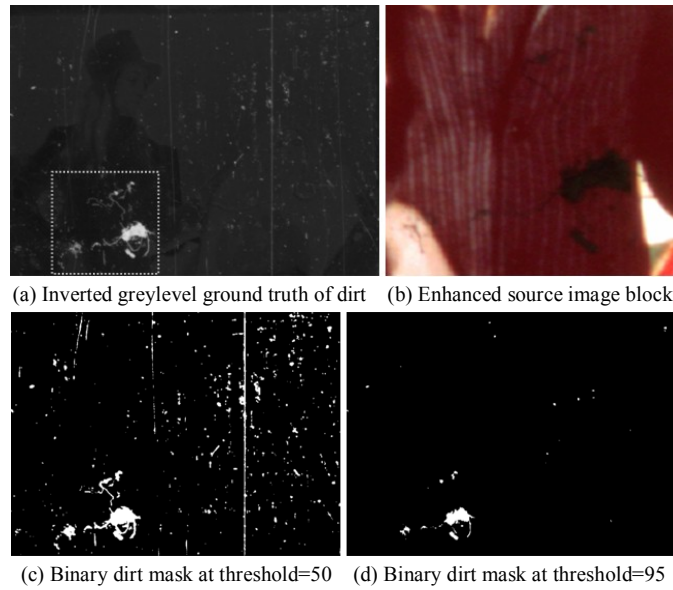


Figure 16: One objective ground-truth image (a) compared with the source image (b) and two binary masks of dirt, (c-d).

Based on these ground truth maps, a quantitative performance assessment was carried out using the well-known Receiver Operating Characteristics (ROC) type of curves [50]. For each ROC curve, it is defined as a function between true positive rate R_{tp} and false positive rate R_{fp} . Let D_g and D_x be binary masks of ground truth maps and detection results, respectively, R_{tp} and R_{fp} are then defined below.

$$R_{tp} = \frac{\text{Count}(D_x \& D_g)}{\text{Count}(D_g)}, R_{fp} = \frac{\text{Count}(D_x \& \bar{D}_g)}{\text{Count}(\bar{D}_g)} \quad (16)$$

where *Count* is a function counting the non-zero elements in a mask and operator $\&$ is the logical AND between two binary masks. Besides, \bar{D}_g is the complement of D_g . In addition, R_{tp} and R_{fp} are also denoted as accuracy and false alarm rate, respectively.

B. Results and discussions

For each test sequence, the results using spatial and/or temporal filtering with or without motion compensation are compared. The approaches used in our evaluation include SDIp, ROD, SSMF, *LUM'*, ML3Dex, and Confidence (confidence measurement). At the same time, results from ML3Dex but without motion compensation, denoted as ML3Dex-m, and results from confidence measurement with global motion compensation, denoted as GM Confidence, are introduced as two additional comparisons. Finally, results using Markov modelling is also compared and labelled as “MRF” in the evaluations, similar approaches can be found in [24, 33, 34, 36, 37, 40, 56, 58]. Consequently, in total nine ROC curves are calculated for each sequence in our evaluation.

In GM Confidence, firstly shifts between f_n and each of its two neighbours (f_{n-1} or f_{n+1}) are estimated by using phase correlation [51], and then these shifts are taken to compensate f_{n-1} or f_{n+1} , respectively. Finally, confidence measurement of dirt is extracted from f_n and its two global motion-compensated neighbouring frames. This is useful in overcoming the effect of camera shaking for more robustness. Moreover, it is much more efficient than that using local motion compensation like optical flow. More details on GM Confidence as well as how to recover detected dirt pixels can be found in [48].

In plotting each of ROC curves, a group of points, R_{tp} versus R_{fp} , are obtained by varying threshold values i.e. t_1 , t_2 , t_3 , or t_s . These thresholds are typically used to determine whether the difference between the current frame on the one hand and a (spatially and/or temporally) filtered frame or a motion-compensated frame on the other is significant or not. If there are more than one thresholds contained 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 confidence and GM confidence approaches, these points are decided by thresholding the grey-level output of the detector using progressively increasing confidence levels.

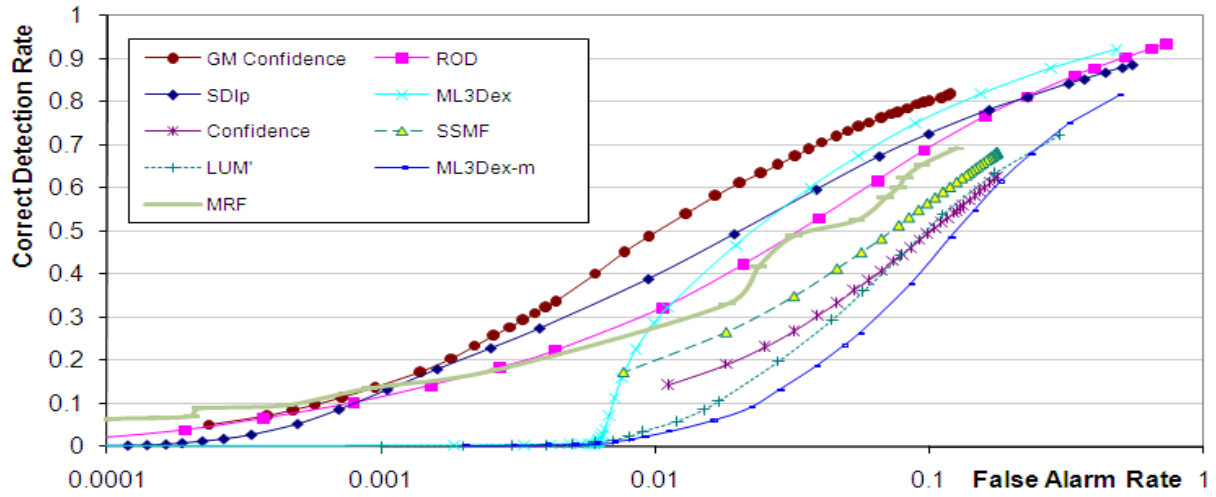
Using the nine approaches mentioned above, Figure 17 illustrates three groups of ROC curves extracted respectively from “Cigarette”, “Door” and “Lady and doll” sequences. Sequence “Cigarette” contains poor contrast, apparent camera shaking, and non-rigid human motion. In Sequence “Door”, big area of dirt exists with fast motion and camera shaking. In “Lady and doll” sequence, there are large dirt areas with camera zooming and shaking when a lady moved her head and hands near a doll of color appearance. From Fig 17 we can easily find several facts as follows.

Firstly, in all the three sequences ML3Dex-m produces the worst performance, i.e. low accuracy but high false alarm rate, followed by *LUM'*. This demonstrates that generally spatio-temporal filtering without motion compensation will fail in such a context.

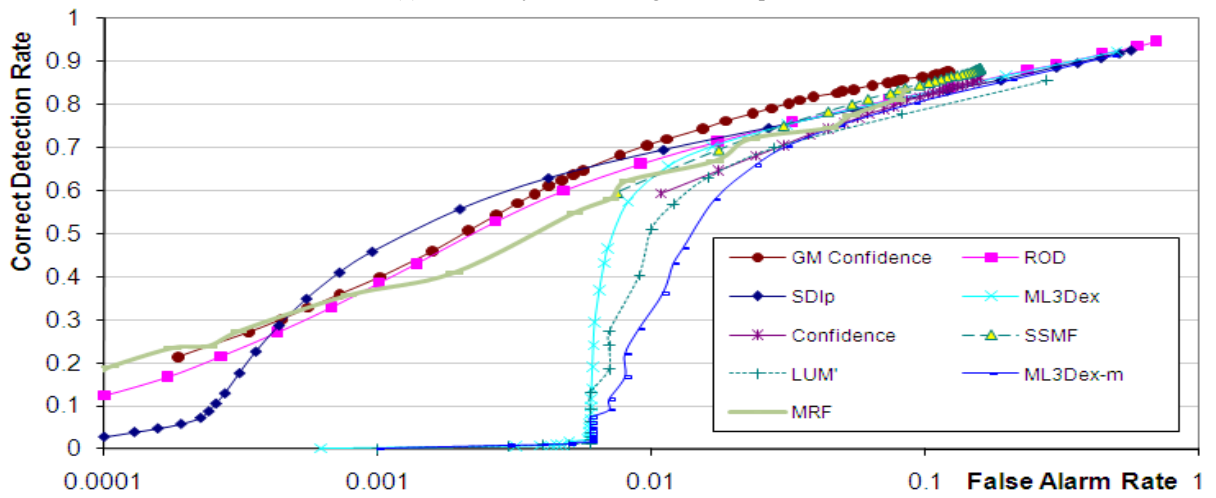
Secondly, SSMF and Confidence, the other two methods without motion-compensation, yield similar performance which seems much better than those from ML3Dex-m and *LUM'*. Occasionally they may have better performance than those motion-compensated ones like SDIp and ROD, with a pretty high false positive rate, say 5%, as shown in Fig. 17(c). However, the high accuracy with much more false alarms is useless in a practical solution as a desired false alarm rate should be less than 1% [24].

Thirdly, even with motion-compensation, ML3Dex still yields surprisingly worse results than other motion-compensated methods such as SDIp and ROD. On the other hand, its performance may appear better than non-motion-compensated approaches provided that the false alarm rate is higher than 0.5%.

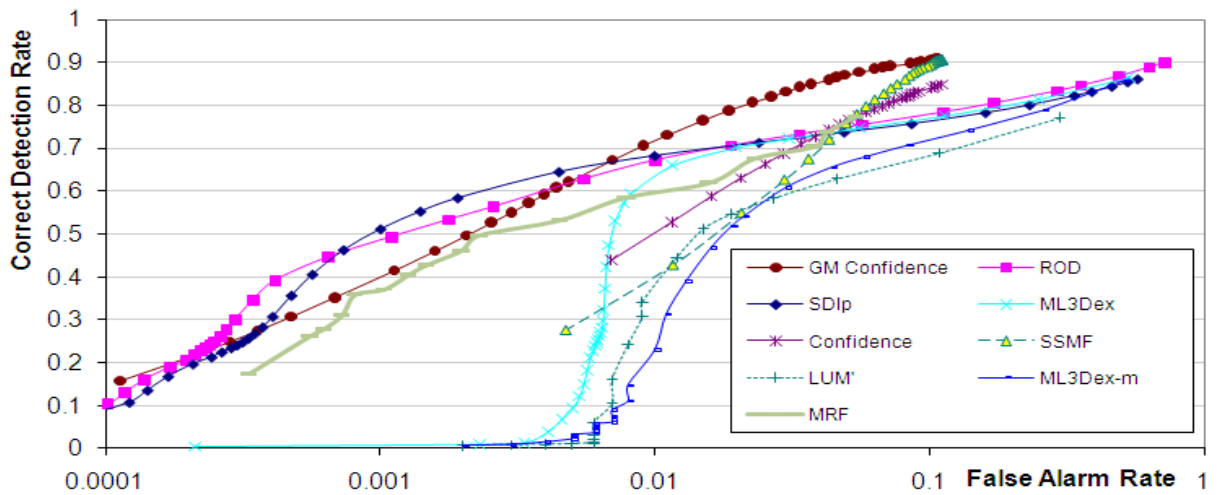
Fourthly, SDIp and ROD generate similar performance, though better performance is obtained from SDIp when the false alarm rate is between 0.1% and 1%. Otherwise, ROD produces better performance when the accuracy is less than 45% or the false alarm rate is higher than 1%.



(a) ROC analysis from "Cigarette" sequence



(b) ROC analysis from "Door" sequence



(c) ROC analysis from "Lady and doll" sequence

Figure 17: ROC analysis of eight different approaches on film dirt detection from three test sequences.

Finally, GM Confidence generates best results in particular when the accuracy is more than 65%. At the same time, the false alarm rate is only 0.4% and up. When asking for fewer false alarms, GM

confidence respectively produces superior, similar and slightly worse performance in the three sequences compared with SDIp and ROD. In addition, GM Confidence also shows best performance when the false alarm rate is less than 0.04% while the accuracy less than 25%. To further remove the false alarms, analysis of complex events [43] or pathological motions [59] may be employed.

Furthermore, the results above confirm the conclusion that more spatial support seems to yield a worse performance [24]. Since SSMF outperforms ML3Dex-m and *LUM'*, however, it seems that this worse performance is mainly due to spatial support from temporal neighbouring frames. In [24], Kokaram *et al* also found that model-based approaches using MRF and AR produced comparable or slightly better results to SDIa, while this SDIa yields worse results than SDIp [3, 8]. As a result, the performance of MRF should be between those from ROD and ML3Dex due to a spatial support commonly considered in several neighbouring frames. This has been confirmed by the results in Fig. 17 as well as those reported in [57]. This finding is different from the results of [58], where significant improvements are achieved against SDIp and confidence-based approach using Markov modelling. Unfortunately, the evaluation in [58] is inappropriate as the results for evaluation are inconsistent. Since both SDIp and confidence-based approaches in nature only suite for detecting impulsive film defects like dirt and sparkles, they will inevitably miss massive of other defects such as scratch. Therefore, different defects need to be separated in order to obtain more convincing and persuasive evaluations since the approach in [58] intends to detect impulsive film defects as well as scratch and even pepper and salt like noise. Furthermore, it is worth noting that MRF approach may occasionally outperforms SDIp and others when the false alarm rate is less than 0.1%, as seen in Fig. 17(a) and (b). However, the overall detection rate remains less than 40%. This makes MRF approach less appropriate in such a context where a relative high detection rate is desired to enable successful recovery of missing data from detected dirt masks, though it is helpful in achieving a low false alarm rate. The reason behind is that the false alarms introduced will not necessarily degrade the quality of restored images, which is further discussed in details in the next subsection.

Therefore, the overall performance of these methods can be summarized below. The worst ones including ML3Dex-m and *LUM'*, followed by Confidence and SSMF. Then, the others should be ML3Dex, MRF, ROD, SDIp, and GM Confidence. Besides, soft morphological filtering introduced in [5] generates comparable results to ML3Dex.

C. Restoration of detected regions of dirt

After detection, recovery of missing data is usually achieved via filtering in spatial and/or temporal domain(s), with or without interpolation. In general, interpolation-free approaches don't need require motion compensation and utilize spatial filtering for restoration, such as the soft morphological filtering in [5] and different variations of median filtering [9-13]. On the other hand, interpolation-based approaches employ spatial-temporal filtering for missing data recovery, thus motion compensation is required to accurately locate corresponding pixels in neighboring frames for interpolation [1, 9]. With the additional support of temporal information for improved accuracy, interpolation-based approaches generally outperforms interpolation-free ones in recovery of missing data.

Let $f_{n-1}(i_-, j_-)$ and $f_{n+1}(i_+, j_+)$ be two motion-compensated pixels corresponding to the missing data of $f_n(i, j)$, a simple method is to take the average of $f_{n-1}(i_-, j_-)$ and $f_{n+1}(i_+, j_+)$ as an estimate of $f_n(i, j)$ [4]. In Kokaram [8], a two-stage motion-compensated median filtering, ML3Dex, is introduced to recover $f_n(i, j)$ as follows, where the definitions of the five windows W_l are given in Fig. 8.

$$\begin{cases} z_l = \text{median}[W_l] & l \in \{5, 6, 7, 8, 9\} \\ \tilde{f}_n(i, j) = \text{median}[z_1, z_2, z_3, z_4, z_5] \end{cases} \quad (17)$$

Further extension to this ML3Dex filtering is to apply motion estimation and ML3Dex filtering only to detected dirt areas, rather than the whole image [48]. It is found such a scheme can not only yield much improved efficiency but also avoid over-smoothing of image details. Please note that over-smoothing problem persists in all approaches where global image filtering is employed [2, 3, 5, 8]. Another extension is the weighted spatial-temporal filtering introduced for missing data recovery, with the so-

called *alpha-modulated bilateral filter* to determine the weights in spatial-temporal neighbors [60]. Again, such method will fail due to the assumptions being unrealistic or if accurate motion compensation cannot be achieved.

In the following, results of recovered dirt from different approaches are compared, including ML3Dex for global filtering, local ML3Dex, local SDIp to intensity component, and local SDIp to three color components. In Fig. 18, one color image and the recovered results are given, where the regions of interests are highlighted in rectangle boxes for easy comparisons. In Fig. 18(a), there is one dirt region on the hand of an actor, and the differences among these results are summarized as follows. Firstly, local SDIp fails to reconstruct all the missing data owing to incomplete masks of dirt detected. Secondly, our method generates comparable results to global ML3Dex filtering in fully recovery of missing data, though global filtering shows slightly better result in recovering the boundary pixels. Thirdly, our approach can avoid over-smoothing caused by global filtering, which can be found in two regions between fingers as highlighted in rectangle boxes. Further details of our local ML3Dex filtering can be found in [48].

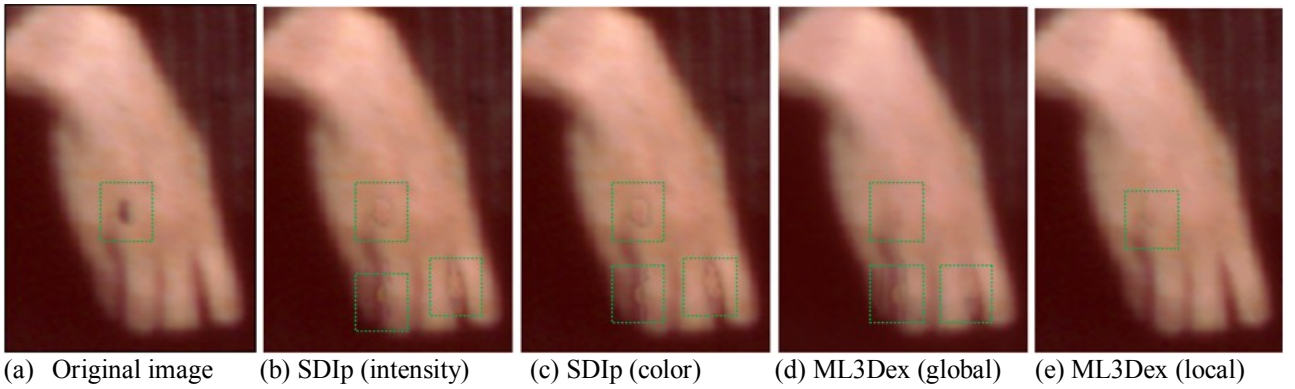


Figure 18: Results of recovered dirt regions for the image in Fig. 16 (b) using different approaches [48].

D. Computational complexity

Computational complexity is one of the most important issues for practical implementation, and this becomes more serious when dealing with massive videos of broadcasting solution. As a result, such expenses cannot be ignored in film dirt detection. Moreover, some approaches contain a training process

and may thus introduce additional cost, such as model-based MRF and soft morphological filtering [5]. However, this is not considered in the following comparisons.

It is obvious that motion-compensated approaches are more expensive than those non-motion-compensated ones. In addition, considering a neighbouring block is more complex than dealing with single pixel. Therefore, the most complex algorithms are model-based approaches, such as MRF, followed by ML3Dex, ROD, DTMF and SDIp. As for the cheapest approaches, these include Confidence, GM Confidence, SSMF, *LUM'*, LUM, and ML3Dex-m. It is interesting to note that GM confidence is the 2nd cheapest approach, only more complex than Confidence.

Using our processing engine (Pentium 4, CPU 3.2GHz, RAM 1GB), Confidence and GM confidence takes about 0.1s and 1.3s in dealing with one frame, compared with 6s to 8s in SSMF, *LUM'*, LUM, and ML3Dex-m. As for ML3Dex, ROD, DTMF and SDIp, each of them takes more than 60s in which about 59s are used for bi-directional motion estimation. The remaining time excluded motion estimation is about 1s in SDIp and DTMF, 2s and 7s in ROD and ML3Dex, respectively.

In practice, a successful dirt detector is expected to achieve a higher accuracy (more than 60%) and lower false alarm rate (less than 1%). Also it should have less computational complexity. Consequently, it seems that GM confidence provides a promising solution to match all these requirements. In addition, GM confidence seems more robust as global motion compensation is more reliable than local ones. More importantly, GM confidence avoids thresholding of filtered or motion-compensated results in determining dirt candidates; hence it is more effective in detecting real dirt samples of a hard core and soft boundary.

VII. Conclusions and Further Directions

Although motion compensation is generally taken as an essential component on film dirt detection, either in heuristics or model-based approaches, it does not degrade gracefully on failure of motion estimation or the model assumptions and may thus generate unpredictable results. On the other hand, spatial filtering is regarded as a useful tool either as a complement or even as an alternative to motion-

compensated approaches. However, it seems spatial and spatio-temporal filtering without motion compensation will normally fail in such a context. One possible reason is that such filtering only yields pixel-level accuracy, where motion-compensated approaches can bring more accuracy through sub-pixel interpolation.

As for motion-compensated processing, temporal filtering such as SDIp demonstrates better or much better performance than spatio-temporal filtering like ROD, ML3Dex and those model-based approaches. This shows again that more spatial support will lead to worse performance, especially to those using spatial support from neighbouring pixels. In particular, approaches with large spatial support, such as ML3Dex and those model-based ones like MRF are not encouraged.

The reason that GM confidence is recommended is not only it excludes spatial support for efficiency but also it avoids information lose from thresholding the discontinuity measurement in filtering the dirt. Although there is local motion in the frames, GM Confidence works on the assumption of intensity continuity in natural scenes thus co-site pixel values (after global compensation) can be used for both robustness and efficiency. However, GM Confidence may fail to work when there are complex motions in big areas of textured background, consequently, combined solutions with local motion estimation is expected in such cases.

Furthermore, the evaluations taken so far are chiefly on standalone approaches, rather than combined solutions. However, these results are certainly useful in giving guidance on choosing different methods in a combined solution. In fact, the main purpose here is to remove false alarms, and these combined solutions will undoubtedly be more practical in achieving this target. Combining GM Confidence with other approaches like segmentation and morphological filtering, the work which is being undertaken seems more promising in terms of accuracy, robustness and efficiency.

Finally, the evaluations used at present are based on binary masks, as the detected results from many existing approaches such as SDIp are binary. However, a more reasonable way is to utilise greylevel evaluation between ground truth maps and detection. Consequently, thresholding is no longer used and a

normalization process is then required to measure the discontinuity, perhaps in a similar way how confidence measurement is extracted.

Acknowledgements

This work forms part of the PrestoSpace project, supported by the European Commission FP6-507336. The authors would like to thank the research staff at Snell & Wilcox 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. The authors would also like to thank anonymous reviewers and the editor for their constructive comments to further improve the quality of this paper.

REFERENCES

- [1] D. Suter and P. Richardson, "Historical film restoration and video coding," in *Proc. Picture Coding Symp.*, pp.389-394, 1996.
- [2] P. V. M. Roosmalen, A. Kokaram, and J. Biemond, "Noise reduction of image sequences as preprocessing for MPEG2 encoding," in *Proc. Eur. Conf. Signal Processing*, vol. 4, pp. 2253-2256, 1998
- [3] A.C. Kokaram, "On missing data treatment for degraded video and film Archives: a survey and a new Bayesian approach," *IEEE Trans. Image Proc.*, vol. 13-3, pp. 397-415, 2004
- [4] P. Schallauer, A. Pinz, and W. Haas, "Automatic restoration algorithms for 35mm film," *J. Computer Vision Res.*, vol. 1-3, pp. 59-85, 1999.
- [5] M.S. Hamid, N.R. Harvey, and S. Marshall, "Genetic algorithm optimization of multidimensional grayscale soft morphological filters with applications in film archive restoration," *IEEE Trans. Cir. Sys. for Video Tech.*, vol. 13-5, pp. 406-416, 2003.
- [6] P. Read and M. P. Meyer, *Restoration of Motion Picture Film*, New York: Butterworth Heinemann, 2000.
- [7] A. Gangal, T. Kayikcioglu, and B. Dizdaroglu, "An improved motion-compensated restoration method for damaged color motion picture films," *Signal Processing: Image Comm.*, vol. 19, no. 4, pp. 353-368, 2004.

- [8] A.C. Kokaram, *Motion Picture Restoration*, Berlin, Germany: Springer-Verlag, 1998.
- [9] N. C. Gallagher and G. L. Wise, "A theoretical analysis of the properties of median filters," *IEEE Trans. Acoust., Speech, Signal Proc.*, vol. 29, pp. 1136-1141, 1981.
- [10] R. Hardie and C. Boncelet, "LUM filters: a class of rank-order-based filters for smoothing and sharpening," *IEEE Trans. Signal Proc.*, vol. 41, no. 3, pp. 1061-1076, 1993.
- [11] A. Nieminen, P. Heinonen, and Y. Neuvo, "A new class of detail-preserving filters for image processing," *IEEE Trans. PAMI*, vol. 9, no. 1, pp. 74-90, 1987.
- [12] E. Abreu, M. Lightstone, S. K. Mitra, and K. Arakawa, "A new efficient approach for the removal of impulse noise from highly corrupted images," *IEEE Trans. Image Proc.*, vol. 5-6, pp. 1012-1025, 1996.
- [13] H. G. Senel, R. A. Peters, II, and B. Dawant, "Topological median filters," *IEEE Trans. Image Proc.*, vol. 11, no. 2, pp. 89-104, 2002.
- [14] P. Maragos, "Morphological filtering for image enhancement and feature detection," in A. C. Bovik (Ed.), *The Image and Video Processing Handbook*, 2nd edition, pp. 135-156, Elsevier Academic Press, 2005.
- [15] O. Buisson, B. Besserer, S. Boukir, and F. Helt, "Deterioration detection for digital film restoration," in *Proc. CVPR*, New York, vol. 1, pp. 78-84, 1997.
- [16] L. Joyeux, S. Boukir, B. Besserer and O. Buisson, "Reconstruction of degraded image sequences: application to film restoration," *Image and Vision Computing*, vol. 19, pp. 503-516, 2001.
- [17] E. D. Ferrandière, *Mathematical Morphology and Motion Picture Restoration*, New York: John Wiley and Sons, pp. 93-120, 2001.
- [18] L. Tenze, G. Ramponi, and S. Carrato, "Robust detection and correction of blotches in old films using spatio-temporal information," in *Proc. SPIE*, vol. 4667, pp. 348-357, 2002.
- [19] E. Dubois and S. Sabri, "Noise reduction in image sequences using motion-compensated temporal filtering," *IEEE Trans. Commu*, vol. 32-7, pp. 826-831, July 1984.

- [20] T. J. Dennis, "Nonlinear temporal filter for television picture noise reduction," *IEE Proc.*, part G, vol. 127, pp. 52-56, 1980.
- [21] R. H. Mcmann, S. Kreinik, J. K. Moore, A. Kaiser, and J. Rossi, "A digital noise reducer for encoded NTSC signals," *SMPTE J.*, vol. 87, pp. 129-133, 1979.
- [22] T. S. Huang and Y. P. Hsu, "Image sequence enhancement," in *Image Sequence Analysis*, T. S. Huang, Ed. Berlin, Germany: Springer-Verlag, pp. 289-309, 1981.
- [23] R. Storey, "Electronic detection and concealment of film dirt," *SMPTE J.*, vol. 94, pp. 642-647, June 1985.
- [24] A. C. Kokaram, R. Morris, W. J. Fitzgerald, and P. J. W. Rayner, "Detection of missing data in image sequences," *IEEE Trans. Image Proc.*, vol. 40-11, pp. 1496-1508, 1995.
- [25] M.J. Black and P. Anandan, "The robust estimation of multiple motions: parametric and pecewise-smooth flow fields," *J. CVIU*, vol. 63-1, pp. 75-104, 1996.
- [26] B. Alp, P. Haavisto, T. Jarske, K. Oistamo, and Y. Neuvo, "Median-based algorithms for image sequence processing," in *Proc. SPIE*, vol. 1360, pp. 122-134, 1990.
- [27] G. R. Arce, "Multistage order statistic filters for image sequence processing," *IEEE Trans. Signal Proc.*, vol. 39, no. 5, pp. 1146-1163, 1991.
- [28] M.J. Nadenau and S.K. Mitra, "Blotch and scratch detection in image sequences based on rank ordered differences," in *Proc. of 5th Int. Workshop on Time-Varying Image Processing*, Florence, 1996, pp. 27-35.
- [29] J. P. Biemond, B. M. V. Roosmalen and R. L. Lagendijk, "Improved blotch detection by postprocessing", *Proc. of IEEE Conf. Acoustics, Speech, Signal Proc.(ICASSP)*, 6: 3101-04, 1999.
- [30] M. K. Özkan, M. I. Sezan, and A. M. Tekalp, "Adaptive motion-compensated filtering of noisy image sequences," *IEEE Trans. Cir. Sys. For Video Techno.*, vol. 3-4, pp. 277-290, 1993
- [31] S. Kalra, M. N. Chong and D. Krishnan, "A new auto-regressive (AR) model-based algorithm for motion picture restoration," *Proc. of ICASSP*, Munich, Germany, pp. 2557-2560, 1997.

- [32] M. K. Özkan, A. T. Erdem, M. I. Sezan, and A. M. Tekalp, "Efficient multiframe Wiener restoration of blurred and noisy image sequences," *IEEE Trans. PAMI*, vol. 1, no. 4, pp. 453-476, 1992.
- [33] R. D. Morris and W. J. Fitzgerald, "Detection and correction of speckle degradation in image sequences using a 3D Markov random field," in *Proc. Int. Conf. Image Processing: Theory and Applications*, pp. , 1993.
- [34] A. C. Kokaram and S. J. Godsill, "MCMC for joint noise reduction and missing data treatment in degraded video," *IEEE Trans. Signal Proc.*, vol. 50-2, pp. 189-205, 2002.
- [35] S. Kalra, M. N. Chong, and D. Krishnan, "A new auto-regressive (AR) model-based algorithm for motion picture restoration," in *Proc. ICASSP*, vol. 4, pp. 2557-2560, 1997.
- [36] R. D. Morris, "Image sequence restoration using Gibbs distribution," PhD. Dissertation, Cambridge University, Cambridge, U. K., 1995
- [37] D. Krishnan, M. N. Chong, and S. Kalra, "On the computational aspects of Gibbs-Markov random field modelling of missing data in image sequences," *IEEE Trans. Image Proc.*, vol. 8-8, pp. 1139-1142, 1999
- [38] S. Geman and D. Geman, "Stochastic Relaxation, Gibbs Distributions, and the Bayesian Restoration of Images," *IEEE Trans. PAMI*, vol. 6, no. 6, pp. 721-741, 1984.
- [39] D. Geman and G. Reynolds, "Constrained restoration and the recovery of discontinuities," *IEEE Trans. PAMI*, vol. 14, no. 3, pp. 367-383, 1992.
- [40] M. N. Chong and D. Krishnan, "An edge-preserving MRF model for the detection of missing data in image sequences," *IEEE Signal Proc. Letters*, vol. 5-4, pp. 81-83, 1998.
- [41] S. Tilie, L. Laborelli, and I. Bloch, "Blotch detection for digital archives restoration based on the fusion of spatial and temporal detectors," in *Proc. 9th Int. Conf. Information Fusion*, Florence, Italy, 2006
- [42] R. D. Morris and W. J. Fitzgerald, "Replacement noise in image sequences, detection and interpolation by motion field segmentation," in *Proc. ICASSP*, vol. V, pp. v/245-v/248, 1994.
- [43] A. Rares, M. J. T. Reinders, and J. Biemond, "Complex event classification in degraded image sequences," in *Proc. ICIP*, vol. 1, pp. 253-256, 2001.

- [44] J. Ren and T. Vlachos, "Segmentation-assisted detection of dirt impairments in archived film sequences". *IEEE Trans. Sys. Man and Cyber. Part B.* vol. 37, no. 2, pp. 463-470, 2007.
- [45] J. Ren and T. Vlachos, "Dirt detection for archived film restoration using an adaptive spatio-temporal approach," *Proc. 2nd European Conf. on Visual Media Production*, pp. 219-228, London, Nov 2005.
- [46] J. Ren and T. Vlachos, "Non-motion-compensated region-based dirt detection for film archive restoration". *Optical Engineering*, vol. 45, no. 8, pp. 087004(1-6), Aug 2006.
- [47] J. Ren and T. Vlachos, "Efficient detection of temporally impulsive dirt impairments in archived films". *Signal Processing*, vol. 87, no. 3, pp. 541-551, 2007.
- [48] J. Ren and T. Vlachos, "Missing data recovery from dirt sparkles on degraded color films". *Optical Engineering*, vol. 46, no. 7, pp. 077001(1-6), July 2007.
- [49] K. P. Lim, A. Das, and M. N. Chong, "Estimation of occlusion and dense motion fields in a bidirectional Bayesian framework", *IEEE T-PAMI*, vol. 24, no. 5, pp. 712-718, 2002.
- [50] J.P. Egan, *Signal Detection Theory and ROC Analysis*, Academic Press, New York, 1975.
- [51] E. De Castro and C. Morandi, "Registration of translated and rotated images using finite Fourier transforms". *IEEE Trans. PAMI*, vol. 9, no. 5, pp. 700-703, Sept. 1987.
- [52] K.-T. Kim and E. Y. Kim, "Film line scratch detection using texture and shape information". *Pattern Recognition Letters*, vol. 31, no. 3, pp. 250-258, 2010.
- [53] L. Maddalena and A. Petrosino, "Restoration of blue scratches in digital image sequences". *Image and Vision Computing*, vol. 26, no. 10, pp. 1314-1326, 2008.
- [54] L. Maddalena, A. Petrosino, and G. Laccetti, "A fusion-based approach to digital movie restoration". *Pattern Recognition*, vol. 42, no. 7, pp. 1485-1496, 2009.
- [55] A. Licsar, T. Sziranyi, and L. Czuni, "Trainable blotch detection on high resolution archive films minimizing the human interaction". *Machine Vision and Applications*, DOI 10.1007/s00138-007-0106-y, 2009.

- [56] S.-C. Nam, M. Abe, and M. Kawamata, "Fast and efficient MRF-based detection algorithm of missing data in degraded image sequences". *IEICE Trans. Fundamentals of Electronics, Communications and Computer Sciences*, vol. E91-A, no. 8, pp. 1898-1906, 2008.
- [57] S. Tilie, I. Bloch, and L. Laborelli, "Fusion of complementary detectors for improving blotch detection in digitized films". *Pattern Recognition Letters*, vol. 28, no. 13, pp. 1735-1746, 2007.
- [58] X. Wang and M. Mirmehdi, "HMM based archive film defect detection with spatial and temporal constraints". *Proc. British Machine Vision Conf. (BMVC'09)*, 2009.
- [59] D. Corrigan, N. Harte, and A. Kokaram, "Pathological motion detection for robust missing data treatment". *EURASIP J. on Advances in Signal Processing*, vol. 2008, Article ID 542436.
- [60] J. Cai and S. B. Kang, "Matte-based restoration of vintage video". *IEEE Trans. Image Processing*, vol. 18, no. 10, pp. 2185-2197, Oct. 2009.