Segmentation-Assisted Dirt Detection for the Restoration of Archived Films

Jinchang Ren, Theodore Vlachos
Centre for Vision, Speech and Signal Processing,
University of Surrey, Guildford, GU2 7XH, U. K.
{j.ren,t.vlachos}@surrey.ac.uk

Abstract

A novel segmentation-assisted method for film dirt detection is proposed. Since dirt manifests as a cluster of pixels whose intensity differs from that of its neighbourhood, we employ segmentation and assume that each small region as a dirt candidate. The assumption is validated by considering raw (non-motion compensated) differences between the current frame and each of the previous and next frames which provides a measure of a confidence. Our experiments show that our method compares favourably with standard spatial, temporal and multistage median filtering approaches and provides efficient and robust detection even for fast moving sequences.

1 Introduction

Automatic restoration is a key enabling technology towards facilitating access to historic 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 and MPEG-4.

The emergence of new multimedia and broadcasting outlets has the potential of dramatically improving public access to cultural assets of unique educational and entertainment value. Consequently, film restoration has recently attracted a lot of interest and several high-profile collaborative projects have received EU funding 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 damage due to environmental hazards such as humidity and dust, chemical instabilities, improper storage and handling practices and even poorly maintained projectors [1], [2]. In this paper, we focus on the detection of impairments occasionally referred to 'dirt'. These are among the most commonly encountered impairments and consequently their successful detection is a priority issue in any archive restoration system [1].

In general, dirt is a temporally impulsive (single-frame) event, appearing mostly as dark or bright opaque spots of random size, shape and location (Fig 1). It is due to particles which attached to the film or abrasions which occurred during storage or when the film passed through various transport mechanisms [1], [7]. Due to their temporal characteristics inter-frame processing has proved a useful tool towards detection and concealment [2], [5].

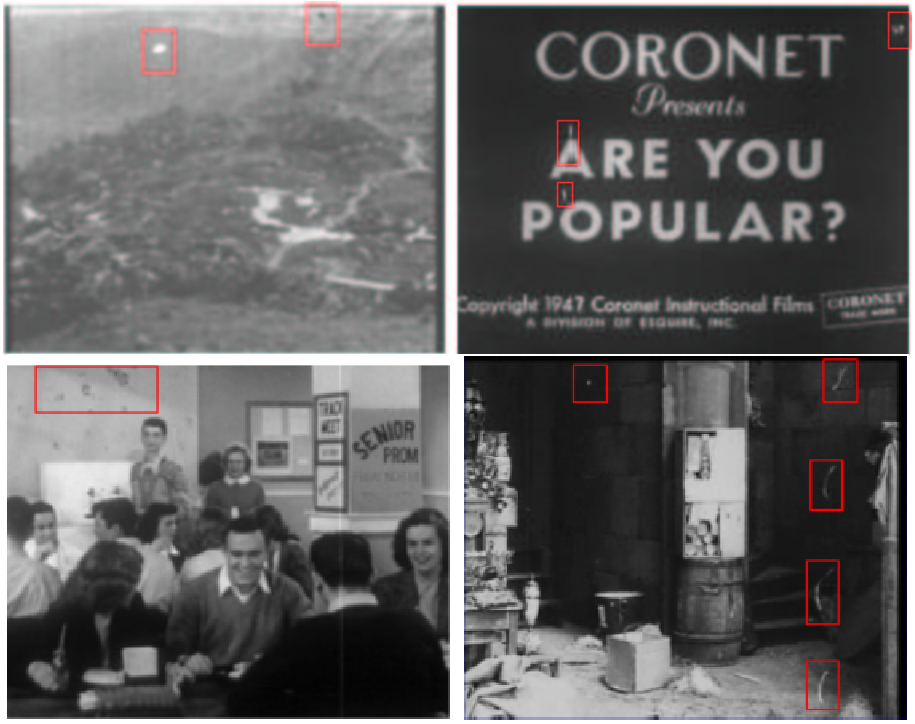


Figure 1: Examples of dirt

A common approach is to use bi-directional motion compensation and subsequently apply temporal median filtering using the current frame and its two motion compensated neighbours. 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. In Kokaram [4], the so-called "Spike Detection Index" (SDI) is proposed. This is also based on the identification of high absolute differences between the current frame and two compensated images. The extended SDI method, SDIp, additionally requires sign consensus of the two differences above. Nadenau, and Mitra [5], 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* extended ROD to five frames to improve accuracy in heavily corrupted images or occluded blotches [3].

Since motion-compensated prediction requires a high degree of complexity and can be unreliable when motion estimation fails, many spatial filtering techniques for dirt detection have also been proposed as alternatives [9-12]. Most of these methods employ

median filtering, as it can preserve edges to a greater extent than linear filters [8]. Subsequently, dirt detection is based on the identification of high difference values between the current frame and the filter output. Nieminem *et al* [9] presented a multi-stage median filter (MMF), which uses hierarchical median operations to reject sparkle distortions. In Arce [10], MMF filters are further evolved as multi-stage order statistic filters (MOS). Senel *et al* [11] proposed a topological median filter to extract edges in noise; however, the filtered images are of unacceptable visual quality in most cases. Hardie, and Boncelet [12], proposed LUM (lower-upper-middle) filters, which utilised two parameters for adjustable smoothing and sharpening of images. Compared with soft morphological filters (SMF), Hamid *et al* pointed out that LUM failed to restore fast-moving objects in image sequences [7]. However, SMF seems impractical for most applications because it needs sufficient representative dirt samples for training to optimise the size and shape of the filters.

In this paper, we propose an efficient and robust method for film dirt detection. Relative to the other methods in the literature, a unique feature of our algorithm is that, thanks to segmentation, dirt is detected at a higher semantic level as a region consisting of connected pixels rather than as isolated pixels. This is consistent with the actual manifestation of dirt in real film samples. A second attractive feature of our scheme is that a confidence measure is derived and attached to detected dirt regions. This is an invaluable feature towards both automatic and operator-assisted dirt concealment as it allows a variable degree of treatment according to preference. Finally our method does not employ motion estimation and motion compensated prediction which reduces its complexity considerably and makes it a good candidate for fast implementations.

2 Dirt Detection using Segmentation

We have employed region growing for image segmentation as follows. For a given image, we use a raster scan order to identify any previously unmarked pixel. Such a pixel is used as a seed for region growing. All other pixels previously merged to form a region are marked.

When growing, a normal distribution $N(\mu,\sigma^2)$ is utilised in which μ and σ are the mean and variance of pixel values in a given region. An unmarked pixel p which is adjacent to a pixel q of that region will be merged in that region if its intensity i_0 satisfies both $i_0 \in [f(q)-0.8\sigma, f(q)+0.8\sigma]$ and $i_0 \in [\mu-3\sigma,\mu+3\sigma]$, where $f(q)$ is the intensity value of the pixel q .

Whenever a new pixel is merged, the mean and variance will be updated. For efficiency, we use $\sigma=\sqrt{\mu/\lambda}$ to approximate the statistical variance, and normally we have $2 \le \lambda \le 6$. Finally, we assign the average intensity to all the pixels within the segmented region. When λ increases, the variance σ will decrease, thus more detail will be preserved in the segmented image. Fig 2 shows two different segmentation results from which we can see that dirt pixels have resisted merging and have remained visible.

It should be noted that detection performance is fairly insensitive to the choice of segmentation method and associated parameters.

Figure 2: Two segmented images with $\lambda = 2$ (left) and $\lambda = 4$ (right) respectively.

Next we assume that each small region in the segmented image is a dirt candidate and use a threshold s operating on region size to determine such candidates. Fig 3 shows the detection masks obtained using two different thresholds underlining the fact that good threshold selection is essential. We can also see that there are still many false alarms in the detected mask, mainly due to moving edges or static small objects. In the next section, inter-frame information will be employed to validate the contents of the detection mask and improve its accuracy.

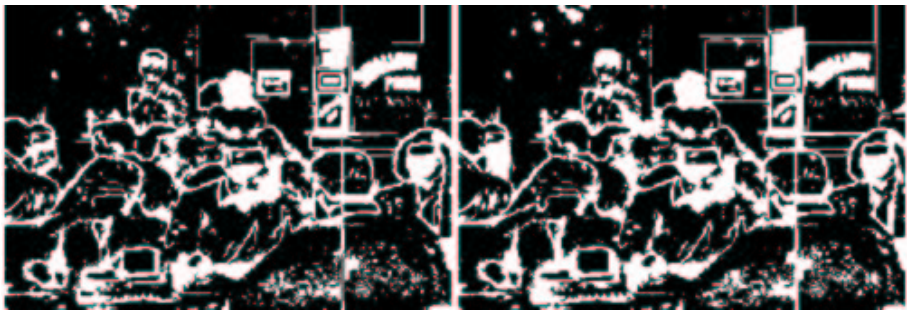


Figure 3: Detection dirt masks (for the right hand side image of Fig 2) with size threshold $s = 30$ pixels (left) and $s = 50$ pixels (right), respectively.

3 Confidence-Based Validation of Dirt

3.1 Definition of Dirt Confidence

The fundamental assumption that dirt is a single-frame event leads naturally to the idea of using inter-frame information for validation. Let f_{n-1}, f_n and f_{n+1} be three consecutive frames. We define $d_{n-} = f_n - f_{n-1}$ and $d_{n+} = f_n - f_{n+1}$ as the forward and backward frame differences, respectively. We also consider the absolute values of d_{n-} and d_{n+} as elementary confidence indicators. We define d_n as:

$$d_n = \begin{cases} \dfrac{2 d_{n-} d_{n+}}{|d_{n-}| + |d_{n+}|} & \text{if} \quad d_{n-} d_{n+} > 0 \\ 0 & \text{otherwise} \end{cases} \tag{1}$$

This attains its maximum value when an idealised dirt impulse occurs against a constant background i.e. when $d_{n-} = d_{n+}$ .If both d_{n-} and d_{n+} are negative or positive, this relates respectively to dark or bright dirt pixels (particles adhered on negative or positive film stock). Figure 4 shows the shape of d_n when d_{n-} and d_{n+} vary in the range between 1 and 25.

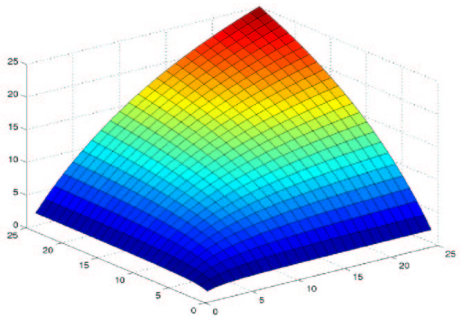


Figure 4. Partial surface of d_n.

Figure 5. One confidence image obtained.

Furthermore, for each value m in d_n, dirt probability is defined as:

$$p_n(m) = \lambda_0^{-1} \int_{m_0+1}^{m} p_d(x)dx \tag{2}$$

$$\lambda_0 + \int_0^{m_0} p_d(x)dx = 1 \tag{3}$$

In (2) and (3), p_d is the intensity distribution PDF (probability density function) of d_n, which can be derived from the histogram of d_n. Parameter λ_0 is used to normalise $p_n(m)$ within [0,1], and m_0 to control the removal of static background.

It is worth noting that $m_0 = 0$ in Eq. (2) amounts to histogram equalisation of d_n. Nevertheless, a static background in the three consecutive frames, may force most pixel values in d_n near zero therefore straightforward histogram equalisation is not useful in this context.

Let μ, γ and σ be the mean, median and variance of the distribution of values in d_n, and let m_0 be determined by

$$m_0 = \frac{\mu + \gamma}{2} + \sigma \tag{4}$$

Using p_n, we define a confidence image as $Conf_n(i,j) = g(p_n(d_n(i,j)))$. Let $z = p_n(d_n(i,j))$, we have $g(z) = (L-1) * \ln(1+z)/\ln 2$. Since $p_n(m) \in [0,1]$, we have $Conf_n(i,j) \in [0, L-1]$.

Figure 5 gives an example of a confidence image obtained for the bottom-left image of Fig 1. From Fig 5 we can see that dirt pixels appear very bright which means that they are detected with high confidence. On the other hand, there remain some false alarms mainly due to motion and moving edges.

3.2 Object-Based Dirt Validation

Assuming B_m is the binary dirt detection mask obtained in Section 2, and $Conf$ is the confidence image extracted in Section 3.1, we can use $Conf$ to validate B_m.

Firstly, a new image, $Conf_d$, is obtained as confidence of detection and given by

$$Conf_d(i,j) = \begin{cases} Conf(i,j) & if \quad B_m(i,j) \neq 0 \\ 0 & otherwise \end{cases} \tag{5}$$

Obviously, $Conf_d$ is a greylevel image in which all the non-zero pixels are the detected dirt of B_m, and the intensity in $Conf_d$ represents confidence of dirt detection.

Secondly, we attempt to remove false alarms caused by the movement of small objects in $Conf_d$. Due to motion, these false alarms occupy a larger area in $Conf$ than in $Conf_d$, thus they can be eliminated by comparing these two areas.

Finally, for a given confidence level c_q a binary dirt detection mask D_q, can be determined by thresholding as follows

$$D_q(i,j,c_q) = \begin{cases} 1 & if \quad Conf_d(i,j) \geq (L-1)*c_q \\ 0 & otherwise \end{cases} \tag{6}$$

In general, c_q can used to control the sensitivity of detection. Figure 6 shows the $Conf_d$ image and binary detection mask with $c_q = 0.9$ (or 90%).

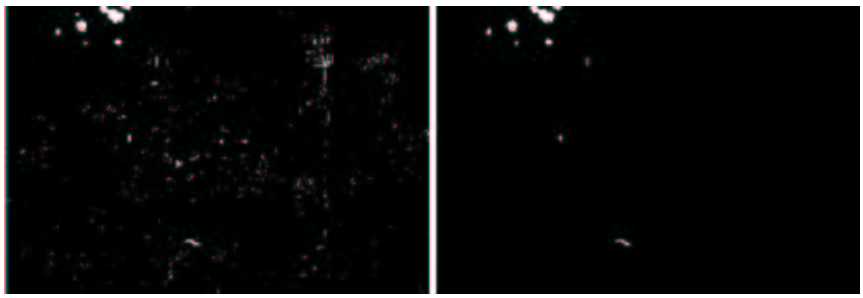


Figure 6. Confidence (left) and binary confidence mask at 90% (right).

4 Experimental Results

4.1 Visual Comparative Assessment

For comparison purposes we have used standard approaches based on spatial, temporal and so-called multistage median filtering. These are explained below.

In spatial median filtering, all the pixels are filtered before dirt detection. For each pixel (i,j) in the current frame f_n, a window W of radius r is defined as

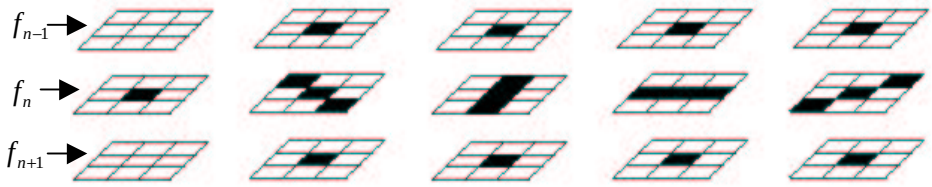
$$W(i,j,r) = \{(i_1,j_1) \mid abs(i_1 - i) \leq r, abs(j_1 - j) \leq r\} \tag{7}$$

The total number of pixels in W is $(2r+1)^2$. We sort the intensities of all the pixels in W and take the median as the filter output. If g_n is the filter output i.e. $g_n(i,j) = median(f_n(i',j')) \,|\, (i',j') \in W(i,j,r)$, the dirt detected by spatial median filtering (relative to a threshold t_s) is given by

$$D_s(i,j) = \begin{cases} 1 & if \quad |g_n(i,j) - f_n(i,j)| > t_s \\ 0 & otherwise \end{cases} \tag{8}$$

For temporal median filtering, at least three frames are needed: the current frame f_n and the two motion compensated frame neighbours, C_{n-} and C_{n+}. In [2] the detected dirt D_t is defined by (9) with thresholds, t_1 and t_2, satisfying $t_1 > t_2$.

$$D_t(i,j) = \begin{cases} 1 & if \quad \begin{array}{l} |C_{n-}(i,j) - f_n(i,j)| > t_1, |C_{n+}(i,j) - f_n(i,j)| > t_1, \\ |C_{n-}(i,j) - C_{n+}(i,j)| < t_2 \end{array} \\ 0 & otherwise \end{cases} \tag{9}$$



(a) W_0 (b) W_1 (c) W_2 (d) W_3 (e) W_4

Figure 7. Sub-windows defined in 3 frames for bi-directional MMF (radius = 1)

As for multistage median filtering (MMF), the bi-directional one is taken in our experiments [10]. The basic concept of MMF is based on the five sub-windows defined in three consecutive frames (see Fig 7). After filtering using (10)-(12), the dirt mask, D_m, is determined in the same way as spatial median filtering given above.

$$z_l = median[w_l] \qquad \forall l \in [0,4] \tag{10}$$

$$z_{\max} = \max[z_l], z_{\min} = \min[z_l] \quad l \in [1,4] \tag{11}$$

$$MMF(z) = median[z_{\max}, z_{\min}, z_0] \tag{12}$$

Fig 8 shows the outputs of the spatial, temporal and multistage median filtering approaches above. For spatial filtering, the parameters are $r = 3$ and $t_s = 15$ while for temporal filtering $t_1 = 25$ and $t_2 = 15$. In MMF, sub-windows in Fig 7 are used for filtering and then thresholded by $t_m = 5$. Motion compensation was implemented by using dense motion fields of sub-pixel accuracy using the well-known Black-Anandan optical flow algorithm [6]. From Fig 7 we can see that although small-area dirt can be detected by spatial and multistage median filtering to reasonable extent, a lot of false alarms also exist. On the other hand, our results at a confidence level of 90% (see Fig

6) are comparable to temporal median filtering, but expensive motion estimation is unnecessary with our method.

Figure 8. Detected dirt mask using spatial (left), temporal (middle) and multistage (right) median filtering.

4.2 Quantitative Comparative Assessment

We obtain a quantitative performance assessment using manually derived ground truth maps and three key criteria namely correct detection rate R_c, false alarm rate R_f and missed detection rate R_m. If D_g is a ground truth dirt mask and $D_x(x = m, q, s, t)$ is the detected mask obtained from any given method, these criteria are defined in as follows, where we have $y = g, m, q, s, t$ in D_y:

$$R_c = \frac{Count(D_g \otimes D_x)}{Count(D_g)} \tag{13}$$

$$R_f = \frac{Count(D_x \otimes \overline{D}_g)}{Count(D_g)} \tag{14}$$

$$R_m = \frac{Count(D_g \otimes \overline{D}_x)}{Count(D_g)} = 1 - R_c \tag{15}$$

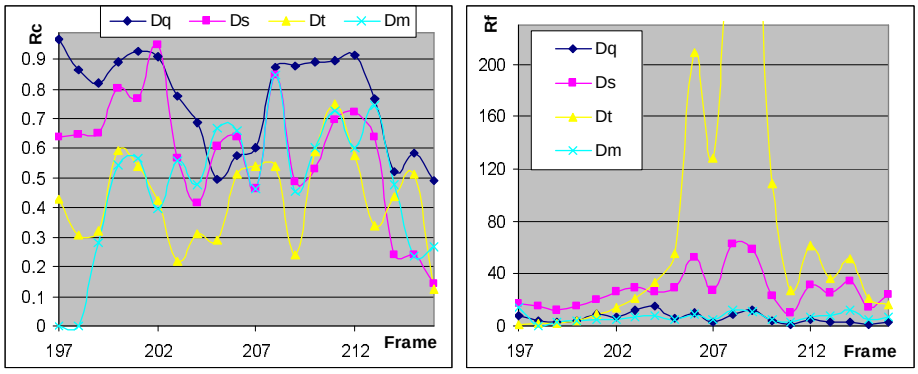
$$\overline{D}_y(i, j) = \begin{cases} 1 & if \quad D_y(i, j) = 0 \\ 0 & otherwise \end{cases} \tag{16}$$

where *Count* is a function counting the non-zero elements in a mask and operator $\otimes$ is the logical AND between two masks. This logical AND is defined as follows:

$$(D_x \otimes D_y)(i, j) = \begin{cases} 1 & if \quad (D_x(i, j) \neq 0, D_y(i, j) \neq 0) \\ 0 & otherwise \end{cases} \tag{17}$$

In this set of experiments we have used broadcast resolution (720x576) sequence "Pennine Way" which contains fast motion and textured background. Parameters are chosen as $r = 3$, $t_s = 15$, $t_1 = 25$, $t_2 = 15$, $\lambda = 4$, $s = 20$ and $c_q = 90$ for the three methods, respectively. Quantitative comparisons of detected results on consecutive 20 frames are given in Fig 9, from which we can see: (1) on the average our method yields the most accurate detection rate and the lowest missed rate; (2) due to failure of motion estimation, temporal median filtering may occasionally yield very poor detection rates

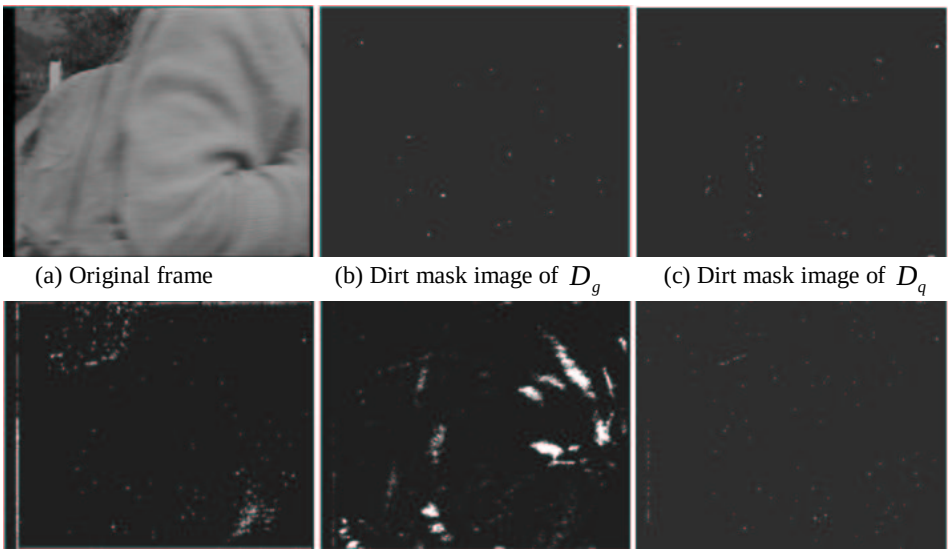
results (see Fig 10e); (3) spatial median filtering yields a higher detection rate for some frames, however, this comes at the expense of more false alarms; (4) MMF shows similar correct rate as spatial median filtering but much less false alarms.



(a) Correct rate (b) False alarm rate

Figure 9: Quantitative comparison of the three methods using R_c and R_f.

Fig 10 shows frame 210 of test sequence "Pennine Way" with the detected dirt from spatial, temporal and multistage median filtering. Comparing Fig 10 and Fig 7, it is interesting to note that false alarms in temporal median filtering are mainly due to inaccurate motion estimation while those in spatial median filtering and MMF are due to sparkle type of noise or sharp edges. Owing to spatial segmentation and confidence-based validation our method is less sensitive to both of those types of failure.



(a) Original frame (b) Dirt mask image of D_g (c) Dirt mask image of D_q

(d) Dirt mask image of D_s (e) Dirt mask image of D_t (f) Dirt mask image of D_m

Figure 10: One test image and dirt mask images of ground truth (D_g) and results from approaches of our method (D_q) and spatial (D_s), multistage (D_m) and temporal (D_t) median filtering.

5 Conclusions

We have presented a segmentation-assisted method for film dirt detection. We demonstrated that image segmentation can provide a useful framework for dirt detection in film restoration. The proposed method is effective, yields a useful measure of confidence and outperforms conventional spatial, temporal and multistage median filtering.

Acknowledgements

This work was carried out in the framework of project PrestoSpace, supported by the European Commission FP6-IST-507336. The authors would like to thank Mike Knee, Steve Simon, Kaaren May and Martin Weston from Snell and Wilcox for valuable discussions and for providing some of the test data.

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 Proc.*, 13(3): 397-415, 2004

[2] P. Schallauer, A. Pinz and W. Haas, "Automatic restoration algorithms for 35mm film", *J. Computer Vision Res.*, 1(3): 59-85, 1999.

[3] A. Gangal, T. Kayikcioglu and B. Dizdarogle, "An improved motion-compensated restoration method for damaged color motion picture films", *Signal Processing: Image Comm.*, 19: 353-368, 2004.

[4] A.C. Kokaram, "Motion picture restoration", Berlin, Germany: Springer-Verlag, 1998.

[5] M.J. Nadenau and S.K. Mitra, "Blotch and scratch detection in image sequences based on rank ordered differences", *Proc. of 5th Int. Workshop on Time-Varying Image Processing*, Florence, 1996.

[6] M.J. Black and P. Anandan, "The robust estimation of multiple motions: parametric and pecewise-smooth flow fields", *J. CVIU*, 63(1): 75-104, 1996.

[7] 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.*, 13(5): 406-416, 2003.

[8] J. Gallagher and G. Wise, "A theoretical analysis of the properties of median filters", *IEEE Trans. Acoust., Speech, Signal Proc.*, 29: 1136-1141, 1981.

[9] A. Nieminem, P. Heinonen and Y. Neuvo, "A new class of detail-preserving filters for image processing", *IEEE Trans. PAMI*, 9: 74-90, 1987

[10] G. Arce, "Multistage order statistic filters for image sequence processing", *IEEE Trans. Signal Proc.*, 39: 1146-1163, 1991.

[11] H. Senel, R. Peters and B. Dawant, "Topological median filters", *IEEE Trans. Image Proc.*, 11: 89-104, 2002

[12] R. Hardie and C. Boncelet, "LUM filters: a class of rank-order-based filters for smoothing and sharpening", *IEEE Trans. Signal Proc.*, 41: 1061-1076, 1993.

[13] R. Adams and L. Bischof, "Seeded region growing", *IEEE Trans. PAMI*, 16(6): 641-647, 1994