

Research Article

Pathological Motion Detection for Robust Missing Data Treatment

David Corrigan, Naomi Harte, and Anil Kokaram

Department of Electronic and Electrical Engineering, Trinity College Dublin, Dublin 2, Ireland

Correspondence should be addressed to David Corrigan, david.corrigan@gmail.com

Received 31 August 2007; Accepted 3 April 2008

Recommended by Peter Schallauer

This paper presents a new missing data detection algorithm that is robust to pathological motion (PM). PM causes clean image data to be misdiagnosed as missing data, resulting in damage to the image during restoration. The proposed algorithm uses a probabilistic framework to jointly detect PM and missing data. It builds on an existing technique of using five frames for detection instead of the standard three frames. This allows the temporally impulsive intensity profile of blotches to be distinguished from the quasiperiodic profile of PM. Another diagnostic for PM is defined on the motion fields of the five-frame window. This follows the observation that PM results in motion fields which are not smooth. A ground truth comparison with standard missing data detectors shows that the proposed algorithm dramatically reduces the number of falsely detected missing data regions. It is also shown to reduce image damage during missing data treatment.

Copyright © 2008 David Corrigan et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

1. INTRODUCTION

Missing data treatment is an important process in the digital restoration of archived media. It is concerned with the removal of visible impairments, known as blotches, commonly present in archived sequences. Also referred to as “dirt and sparkle,” blotches are caused by either adherence of dirt to the surface of the film emulsion, or by abrasion of the emulsion. Consequently, blotches manifest as typically small dark or bright regions in the digital sequence which obscures the true image data. Missing data treatment (MDT) algorithms attempt to detect these blotches and to somehow recover the true image data.

Most blotch detection algorithms are derived from the assumption that if a blotch is present at a particular location in a frame, then no blotches are present at the same point in neighbouring frames. Blotches are therefore considered to be a temporally impulsive event due to the change in intensity from the true intensity to the blotch intensity at a particular pixel. Detection of blotches can be performed by searching for sites where temporal discontinuities exist between the current frame and the previous and next frames. Consequently, missing data detectors employ a three-

frame window and detect discontinuities by measuring the displaced frame difference (DFD) between the appropriate frames. Examples of detectors using a three-frame window include the deterministic spike detection index (SDI) [1, 2] and rank order distance (ROD) [3, 4] detectors, which apply deterministic thresholds to the DFDs to detect discontinuities, and the probabilistic Morris [2, 5] and blotch MRF [6] algorithms, which allow for the inclusion of prior knowledge of the result into the decision. Other detectors [1, 7] propose methods for the estimation of a nonbinary blotch confidence index.

Motion estimation is a necessary part of the blotch detection process. It is used to compensate for motion over the three-frame window, preventing moving objects from being confused with blotches. Although the state-of-the-art motion estimators can robustly estimate the most common motion patterns, there are complex motion patterns which cannot be estimated accurately. This complex motion is referred to as *pathological motion* (PM) and can lead to true image data being detected as a blotch (see Figure 1). Pathological motion is usually caused by the occlusion and uncovering of objects in a sequence, or by fast moving objects (resulting in the phenomenon of motion blur). More

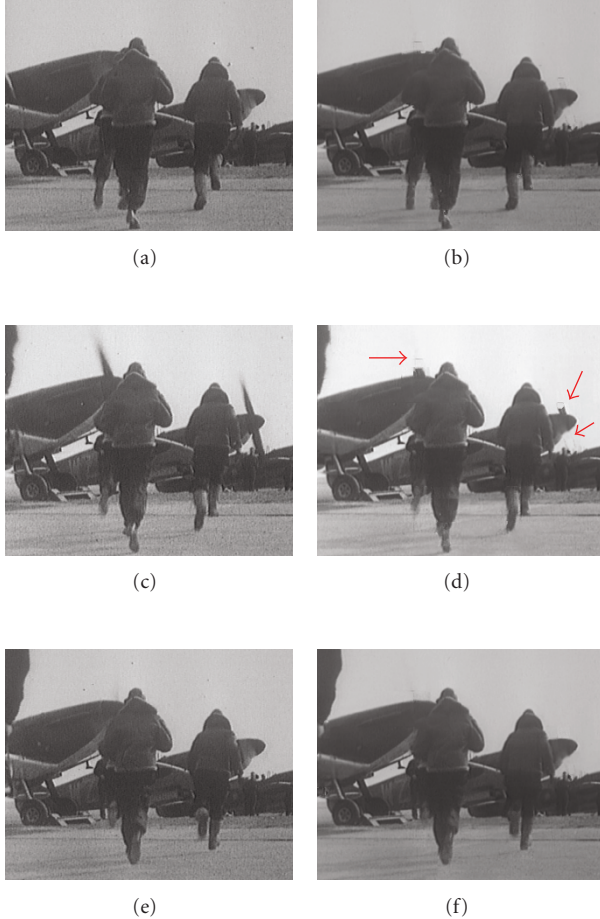


FIGURE 1: Left: three consecutive frames of a sequence where the motion of the propeller blades is PM (repetitive occlusion); right: the images after MDT using [10]. The arrows indicate the removal of the blades due to the blades being misdiagnosed a blotch.

extensive taxonomies of PM are presented in [8, 9]. PM affects the robustness of blotch detectors as it can also cause temporal discontinuities.

In this paper, a new missing data detection algorithm is presented. The algorithm is made more robust by preventing PM regions being detected as blotches, thereby preventing damage to image data during restoration of dirty sequences. A five-frame window is used instead of the standard three frames. This allows blotches to be distinguished from long term forms of PM, since blotches have a temporally impulsive intensity profile and PM often causes the intensity to vary in a periodic manner (see Figure 2). The next section introduces a review of the existing approaches to improve the robustness of blotch detection to PM and outlines the motivation for the new algorithm. The algorithm is then outlined in Sections 3 and 4. Section 5 presents an evaluation of the algorithm, including the results of an experimental comparison of the new detector with existing blotch detectors on two ground truth sequences. The paper concludes with a discussion of the algorithm, outlining areas for future development.

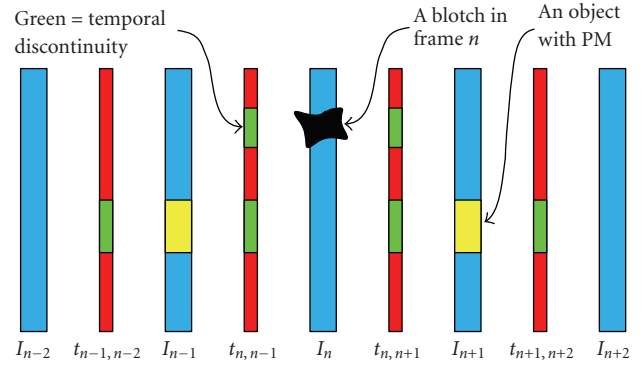


FIGURE 2: This figure illustrates the arrangement of the five frames window for blotch detection (in blue). The window is centred on frame $n(I_n)$ which is the frame on which blotch detection is to be performed. Between the five frames, four temporal discontinuity fields ($t_{*,*}$) (in red) can be estimated between pairs of consecutive frames. In this diagram a blotch is present on frame n and the yellow object is occluded in every second frame (an example of PM). Due to the pattern of occlusion it is impossible to distinguish between the blotch and the object when detecting blotches based on temporal discontinuities over a three-frame window. By extending the window it is possible to perform a correct diagnosis.

2. REVIEW OF PM DETECTION ALGORITHMS

Previous approaches for dealing with PM have largely fallen into three categories. The first strategy is to recognise that PM is associated with the motion of foreground objects. Therefore, by performing a foreground/background segmentation, likely PM regions can be associated with the estimated foreground and missing data detection can be made more conservative in these regions [6]. The second uses the colour statistics of blotches and PM regions (specifically motion blur) to distinguish between them [11]. This is possible as blotches typically manifest as regions of dark or bright intensity while motion blur is the apparent smearing of a foreground object over a background.

The final approach is to extend the temporal window for blotch detection from three to five frames. This approach was first adopted by van Roosmalen in [12], in which the ROD detector was modified to operate on a five-frame window. Another more elaborate algorithm adopting this approach is outlined by Bornard in [9]. In the Bornard algorithm a probabilistic framework is used to estimate the four binary temporal discontinuity fields over a five-frame window (see Figure 2). This is followed by a deterministic interpretation step to perform the classification. Missing data is classified if temporal discontinuities only exist at a site in the two central discontinuity fields. Sites at which discontinuities also exist in the two outer fields are classified as PM.

The algorithm proposed in this paper adopts a similar strategy to the Bornard algorithm with two key modifications. Instead of having separate temporal discontinuity detection and interpretation stages, these stages are integrated into a single probabilistic framework which estimates a blotch and PM mask from the five frames. This allows

for prior knowledge of the blotch and PM masks as well as the temporal discontinuity fields (TDFs) to be included in the framework. The second novelty to algorithm is the addition of a further decision criterion for PM based on the smoothness of the motion fields. This follows from the observation that PM causes the local smoothness assumption of a motion field to be violated. The vector field divergence is used as the smoothness measure in the proposed algorithm.

Section 3 outlines the blotch and PM detection model employed in the algorithm and how this relates to the temporal discontinuity and motion field smoothness information. Section 4 describes how the two detection criteria are integrated into the probabilistic framework and subsequently introduces the new algorithm.

3. ALGORITHM OVERVIEW

The goal of the algorithm is to segment the current frame into regions of PM, missing data, and uncorrupted sites. To achieve this, a label field, $l(\mathbf{x})$, is defined as follows:

$$l(\mathbf{x}) = \begin{cases} 0 & \text{no missing data or pathological motion,} \\ 1 & \text{missing data,} \\ 2 & \text{pathological motion.} \end{cases} \quad (1)$$

Every pixel in the image is assigned one of the three labels. Like the Bornard algorithm, a pixel cannot be both a missing data site and a PM site. Although there is no reason why a blotch could not occur in a PM region, it would not be possible to reliably detect a blotch using a purely temporal discontinuity-based detector. Furthermore, a PM detection implies that the data at any blotch cannot be interpolated temporally. Consequently, interpolation must be performed using spatial rather than temporal information.

3.1. Temporal discontinuity-based detection

Like the Bornard algorithm, the DFD between a pair of neighbouring frames is used as the measure of temporal discontinuity. A temporal discontinuity is said to exist at a given site if the absolute value of the DFD at the site is sufficiently large. In the five frame window, there are four pairs of neighbouring frames where the frames are denoted by $I_{n-2}(\mathbf{x})$, $I_{n-1}(\mathbf{x})$, $I_n(\mathbf{x})$, $I_{n+1}(\mathbf{x})$, and $I_{n+2}(\mathbf{x})$. A DFD can be measured between each pair with a binary temporal discontinuity field associated with each DFD. Consequently, four DFDs ($\Delta_{n-2}(\mathbf{x})$, $\Delta_{n-1}(\mathbf{x})$, $\Delta_{n+1}(\mathbf{x})$, and $\Delta_{n+2}(\mathbf{x})$) and TDFs ($\mathbf{t}(\mathbf{x}) = [t_{n-2}(\mathbf{x}), t_{n-1}(\mathbf{x}), t_{n+1}(\mathbf{x}), t_{n+2}(\mathbf{x})]$) are estimated over the five-frame window. In the proposed algorithm, all configurations of temporal discontinuity are considered. There are sixteen possible configurations of the four TDFs and a state field $s(\mathbf{x})$ is defined which describes the configuration of a site. Each configuration is directly mapped to a value of $l(\mathbf{x})$ (Table 1).

Missing data regions are characterised by an impulsive temporal intensity profile. Therefore, if a blotch exists in frame n , the absolute values of the DFDs between both frames n and $n-1$ ($\Delta_{n-1}(\mathbf{x})$) and frames n and $n+1$ ($\Delta_{n+1}(\mathbf{x})$) will be large but the other two DFDs will have

TABLE 1: The temporal discontinuity model for the framework.

State, $s(\mathbf{x})$	$\mathbf{t}(\mathbf{x})$	$l(\mathbf{x})$	State, $s(\mathbf{x})$	$\mathbf{t}(\mathbf{x})$	$l(\mathbf{x})$
0	0, 0, 0, 0	0	8	1, 0, 0, 0	0
1	0, 0, 0, 1	0	9	1, 0, 0, 1	2
2	0, 0, 1, 0	0	10	1, 0, 1, 0	2
3	0, 0, 1, 1	2	11	1, 0, 1, 1	2
4	0, 1, 0, 0	0	12	1, 1, 0, 0	2
5	0, 1, 0, 1	2	13	1, 1, 0, 1	2
6	0, 1, 1, 0	1	14	1, 1, 1, 0	2
7	0, 1, 1, 1	2	15	1, 1, 1, 1	2

TABLE 2: The divergence state model for the probabilistic framework.

Divergence state, $v(\mathbf{x})$	$l(\mathbf{x})$	$s(\mathbf{x})$
0	0, 1	0, 1, 2, 4, 6, 8
1	2	3, 5, 7, 9, 10, 11, 12, 13, 14, 15

small values (see Figure 2). This results in a predictable temporal discontinuity profile described by $s(\mathbf{x}) = 6$ in Table 1. On the other hand, the quasiperiodic profile of long-term PM will result in high absolute values in more than two DFDs. Consequently, any temporal discontinuity configuration with two or more temporal discontinuities present, apart from the missing data case, is considered to be associated with PM (Table 1). In the table, states 3 and 12 can correspond to missing data in the neighbouring frames. It is considered here as PM as it is not caused by missing data in the current frame.

In the Bornard algorithm, all candidate blotch sites are found (i.e., where $t_{n-1}(\mathbf{x}) = t_{n+1}(\mathbf{x}) = 1$) before each candidate is classified as either being a blotch site or PM site. This is performed based on the frequency of discontinuity in $t_{n-2}(\mathbf{x})$ and $t_{n+2}(\mathbf{x})$ in the neighbourhood of the candidate. Effectively, this means that not every configuration of temporal discontinuity is detected, distinguishing the Bornard algorithm from the proposed algorithm.

3.2. Motion field smoothness based detection

The second feature used to detect PM is the smoothness of the motion field. The vector field divergence is used as the basis of the smoothness measure. By estimating the divergences of all the motion fields in the five-frame window, a smoothness energy term, $E_{\text{div}}(\mathbf{x})$, can be defined which describes the smoothness of the motion fields. This measure of divergence will be large in the presence of long-term PM and low in other cases. The divergence measure is associated with a binary field $v(\mathbf{x})$, where a value of 0 corresponds to a low divergence measure (i.e., when motion is not pathological) and a value of 1 corresponds to a high divergence measure (i.e., for PM). This relationship is described in Table 2.

The fields s and v act as auxiliary variables which dictate the final value of $l(\mathbf{x})$ in the probabilistic framework. For a

given value of s or ν , the value of l is directly defined through Tables 1 and 2.

4. PROBABILISTIC FRAMEWORK

The framework derives an estimate of $s(\mathbf{x})$ from the posterior $P(s(\mathbf{x}) \mid \Delta_n(\mathbf{x}), E_{\text{div}}(\mathbf{x}))$. For convenience of notation, the four DFDs have been grouped into a vector-valued function Δ_n where $\Delta_n(\mathbf{x}) = [\Delta_{n-2}(\mathbf{x}), \Delta_{n-1}(\mathbf{x}), \Delta_{n+1}(\mathbf{x}), \Delta_{n+2}(\mathbf{x})]$.

The posterior is factorised in a Bayesian fashion as follows:

$$\begin{aligned} P(s \mid \Delta_n, E_{\text{div}}) &\propto P(s, \Delta_n, E_{\text{div}}) \\ &\propto P_t(\Delta_n \mid s) \times P_d(E_{\text{div}} \mid s) \times P_r(s), \end{aligned} \quad (2)$$

where the index \mathbf{x} has been dropped for clarity. Rather than considering the unknown variables $l(\mathbf{x})$, $\nu(\mathbf{x})$, and $s(\mathbf{x})$ as separate random variables, only one random variable $s(\mathbf{x})$ is considered. The values of $l(\mathbf{x})$ and $\nu(\mathbf{x})$ can then be determined from the estimate of $s(\mathbf{x})$ according to Tables 1 and 2.

There are two likelihoods associated with the framework: $P_t(\cdot)$ associated with the DFDs of the window and $P_d(\cdot)$ associated with the divergence measure. The final term is a prior on the state field $s(\mathbf{x})$. Mathematically, the correct form of this expansion is $P_t(\Delta_n \mid E_{\text{div}}, s) \times P_d(E_{\text{div}} \mid s) \times P_r(s)$. Although it seems likely that some relationship between the DFDs and divergence measure exists, for the purposes of this framework it is assumed that they are statistically independent.

4.1. Temporal discontinuity likelihood

Before the temporal discontinuity likelihood can be introduced, the method for calculating the DFDs must be outlined. As has been outlined in Section 3.1, there are four DFDs to be estimated described by the four dimensional DFD vector $\Delta_n(\mathbf{x})$. The DFDs are calculated by first compensating the motion of each frame of the window relative to the central frame. If the image sequence model is the local translation model given by

$$\begin{aligned} I_n(\mathbf{x}) &= I_{n-1}(\mathbf{x} + \mathbf{d}_{n,n-1}(\mathbf{x})) + e(\mathbf{x}) \\ &= I_{n+1}(\mathbf{x} + \mathbf{d}_{n,n+1}(\mathbf{x})) + e(\mathbf{x}), \end{aligned} \quad (3)$$

where $\mathbf{d}_{h,k}$ is the motion field representing the relative position of pixels in the k th frame relative to the h th frame, then the five motion compensated frames ($I'_{n-2}, I'_{n-1}, I'_n, I'_{n+1}, I'_{n+2}$) are given by

$$\begin{aligned} I'_{n-2}(\mathbf{x}) &= I_{n-2}(\mathbf{x} + \mathbf{d}_{n,n-1} + \mathbf{d}_{n-1,n-2}(\mathbf{x} + \mathbf{d}_{n,n-1})), \\ I'_{n-1}(\mathbf{x}) &= I_{n-1}(\mathbf{x} + \mathbf{d}_{n,n-1}), \\ I'_n(\mathbf{x}) &= I_n(\mathbf{x}), \\ I'_{n+1}(\mathbf{x}) &= I_{n+1}(\mathbf{x} + \mathbf{d}_{n,n+1}), \\ I'_{n+2}(\mathbf{x}) &= I_{n+2}(\mathbf{x} + \mathbf{d}_{n,n+1} + \mathbf{d}_{n+1,n+2}(\mathbf{x} + \mathbf{d}_{n,n+1})). \end{aligned} \quad (4)$$

Consequently, the four DFDs of the window are given by

$$\begin{aligned} \Delta_{n-2}(\mathbf{x}) &= I'_{n-1}(\mathbf{x}) - I'_{n-2}(\mathbf{x}), \\ \Delta_{n-1}(\mathbf{x}) &= I'_n(\mathbf{x}) - I'_{n-1}(\mathbf{x}), \\ \Delta_{n+1}(\mathbf{x}) &= I'_n(\mathbf{x}) - I'_{n+1}(\mathbf{x}), \\ \Delta_{n+2}(\mathbf{x}) &= I'_{n+1}(\mathbf{x}) - I'_{n+2}(\mathbf{x}). \end{aligned} \quad (5)$$

4.1.1. Motion estimation

The motion fields used to perform the motion compensation above are performed as a preprocess to the algorithm. For every frame of the sequence, a backward and forward motion field must be estimated (i.e., for motion in the current frame relative to the previous and next frames of the sequence, resp.). Any motion estimator can be used to generate the motion fields. The choice of motion estimator affects the correct detection and false alarm rate for missing data and also the PM detection rate. In the tests described in Section 5, the gradient-based algorithm outlined in [13] is used.

4.1.2. The likelihood expression

The data likelihood $P_t(\Delta_n \mid s)$ constrains each DFD to be low when a temporal discontinuity does not exist. An expression for the likelihood of a simple temporal discontinuity detector is given by

$$P(\Delta(\mathbf{x}) \mid t(\mathbf{x})) \propto \exp - \left\{ \frac{\Delta(\mathbf{x})^2}{2\sigma_e^2} (1 - t(\mathbf{x})) + \frac{\alpha^2}{2} t(\mathbf{x}) \right\}, \quad (6)$$

where σ_e^2 is the variance of the model error $e(\mathbf{x})$ and α acts as a threshold on temporal discontinuities. In the new algorithm there are four DFDs and the likelihood expression is formed by multiplying the expressions for four separate temporal discontinuity detection likelihoods. The likelihood expression is given in vector form by

$$P_t(\Delta_n \mid s) \propto \exp - \sum_{k=0}^3 \left\{ \frac{\Delta_n[k]^2}{2\sigma_e^2} (1 - \mathbf{t}[k]) + \frac{\alpha_k^2}{2} \mathbf{t}[k] \right\}, \quad (7)$$

where k refers to the component index of a vector. For each component the maximum likelihood condition for a temporal discontinuity is

$$(\Delta_n[k](x))^2 > \sigma_e^2 \alpha_k^2. \quad (8)$$

The value of the model error variance, σ_e^2 , is determined by estimating the variance of the DFDs when $s(\mathbf{x}) = 0$. The threshold α is allowed to vary for each component of the likelihood. The α_2 and α_3 thresholds for the central DFDs (Δ_{n-1} and Δ_{n+1}) are linked to a deterministic threshold on the DFD, δ_t . The equivalent α threshold to δ_t can be derived from (8) and is given by

$$\alpha_2 = \alpha_3 = \sqrt{\frac{\delta_t^2}{2\sigma_e^2}}. \quad (9)$$

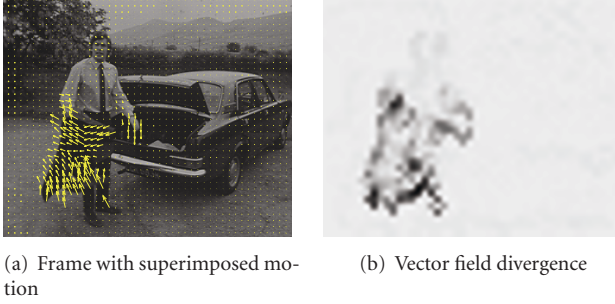


FIGURE 3: The motion of the jacket in the left image is an example of PM. The motion field in this region is not smooth and as a result the divergence of the motion field in this region has a high absolute value. The dark colours represent high divergence values.

Detection of missing data is made more conservative by using a lower α (i.e., for α_1 and α_4) on the outer DFDs (i.e., Δ_{n-2} and Δ_{n+2}) than on the central DFDs. The outer threshold ratio, α_r , is defined as the ratio of outer thresholds α_1 and α_4 to the inner thresholds α_2 and α_3 . A typical value of α_r is 0.5.

4.2. Divergence likelihood

The divergence of a vector or flow field measures the rate at which flow exits a given point of the flow field. The divergence of a 2D motion field, $\mathbf{d}(\mathbf{x})$, is defined as

$$\text{div}(\mathbf{d}(\mathbf{x})) = \frac{\partial \mathbf{d}_x(\mathbf{x})}{\partial x} + \frac{\partial \mathbf{d}_y(\mathbf{x})}{\partial y}, \quad (10)$$

where \mathbf{d}_x and \mathbf{d}_y are the horizontal and vertical components of the motion field, respectively. When PM occurs, the smoothness of the motion field is violated and the absolute value of the divergence is large (Figure 3). By finding the divergence of the motion fields involved in the motion compensation of the five-frame window, a divergence measure $E_{\text{div}}(\mathbf{x})$ can be defined as the sum of the absolute divergences of the four motion fields used for motion compensation. That is

$$E_{\text{div}}(\mathbf{x}) = \sum_{k=n-1}^n |\text{div}(\mathbf{d}_{k,k-1}(\mathbf{x}))| + \sum_{k=n}^{n+1} |\text{div}(\mathbf{d}_{k,k+1}(\mathbf{x}))|. \quad (11)$$

The divergence likelihood constrains E_{div} to be low for either missing data or unaffected states. The expression used for the likelihood energy in the framework is

$$\begin{aligned} & -\log(P_d(E_{\text{div}} | s)) \\ & \propto \begin{cases} \Lambda_d(1 - \phi(E_{\text{div}})) & \text{if } s = \{0, 1, 2, 4, 6, 8\}, \\ 0 & \text{if } s = \{3, 5, 7, 9, 10, 11, 12, 13, 14, 15\}, \end{cases} \end{aligned} \quad (12)$$

where Λ_d is the weight of the divergence likelihood in the framework and where $s = \{3, 5, 7, 9, 10, 11, 12, 13, 14, 15\}$ is the set of PM states (i.e., $l = 2, v = 1$) and $s = \{0, 1, 2, 4, 6, 8\}$

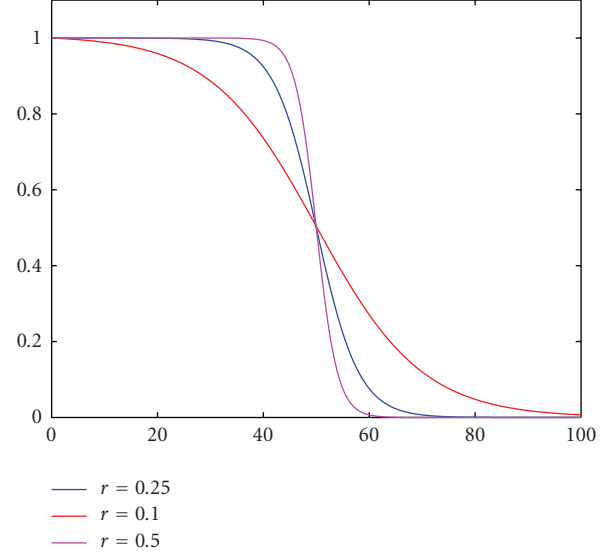


FIGURE 4: This figure shows three plots of $\phi(a)$ for $a \in [0, 100]$. In this example $e_t = 50$. As r increases the drop in energy about e_t becomes the more abrupt.

is the set of unaffected and missing data sites (i.e., $l = \{0, 1\}$, $v = 0$). $\phi(E_{\text{div}}(\mathbf{x}))$ is a sigmoidal function defined on the divergence measure, E_{div} , and is given by

$$\phi(a) = \frac{1 + e^{r \cdot e_t}}{e^{r \cdot e_t}} \times \frac{e^{-r(a-e_t)}}{1 + e^{-r(a-e_t)}}, \quad (13)$$

where r and e_t are positive real numbers (see Figure 4). Effectively, the divergence likelihood acts as a bias towards the detection of PM when the divergence measure is large. e_t acts as a soft threshold for detecting high divergences.

4.3. The smoothness prior

The prior expression used in the framework is

$$P_r(s(\mathbf{x})) = P_s(l(\mathbf{x}) | L) \quad (14)$$

which is a spatial smoothness prior on $l(\mathbf{x})$. The smoothness prior is enforced on $l(\mathbf{x})$ rather than $s(\mathbf{x})$. This reduces computational complexity since the complexity is proportional to the number of values a field can hold. As $l(\mathbf{x})$ is the field of interest the spatial smoothness of the result is not compromised.

The prior on l is given by

$$P(l(\mathbf{x}) | L) \propto \exp - \left\{ \Lambda_l \sum_{\mathbf{y} \in \mathcal{N}_s(\mathbf{x})} \lambda_{\mathbf{y}} (1 - u(\mathbf{x}, \mathbf{y})) \rho(l(\mathbf{x}), l(\mathbf{y})) \right\}, \quad (15)$$

where Λ_l is the weight of the prior in the framework (a value of 1 is commonly used), where $\mathcal{N}_s(\mathbf{x})$ is a spatial neighbourhood of \mathbf{x} and where $\lambda_{\mathbf{y}}$ is a weight inversely proportional to $\|\mathbf{x} - \mathbf{y}\|$. $\rho(l(\mathbf{x}), l(\mathbf{y}))$ is the energy penalty for neighbours \mathbf{x} and \mathbf{y} having the labels $l(\mathbf{x})$ and $l(\mathbf{y})$. $u(\mathbf{x}, \mathbf{y})$



(a) Result without the missing data bias (b) Result with the missing data bias

FIGURE 5: This figure shows the effect of including the missing data bias in the framework. In the images above a propellor blade undergoing PM is highlighted by the blue circle. Detected PM is highlighted in red and detected missing data is highlighted in green. When the standard smoothness prior is used (i.e., $K = 1$), the top part of the blade is incorrectly detected as missing data. Introducing the PM bias penalises against such a configuration by favouring the detection of PM over missing data. Consequently, the entire blade has been classified as PM when the bias is introduced.

is a binary edge value between neighbouring pixels which turns off smoothness across image edges, allowing sharp transitions in $l(\mathbf{x})$ across edges [14]. $u(\mathbf{x}, \mathbf{y})$ is defined on the central frame $I_n(\mathbf{x})$.

4.3.1. The PM bias

Usually the value of $\rho(l(\mathbf{x}), l(\mathbf{y}))$ is 0 for all $l(\mathbf{x}) = l(\mathbf{y})$ and 1 for all $l(\mathbf{x}) \neq l(\mathbf{y})$. However, in this framework a bias is introduced into ρ to prevent a part of an “object” being classified as PM and another as missing data (Figure 5). This is achieved by increasing the penalty for assigning a site which has a neighbouring PM site as a missing data site ($\rho(1, 2)$). The effect is to make blotch detection more conservative by favouring PM detection at such sites. The expression used for ρ is

$$\rho(l(\mathbf{x}), l(\mathbf{y})) = \begin{cases} 0 & \text{if } l(\mathbf{x}) = l(\mathbf{y}), \\ K & \text{if } l(\mathbf{x}) = 1, l(\mathbf{y}) = 2, \\ 1 & \text{otherwise,} \end{cases} \quad (16)$$

where $K > 1$. A value of 5 is used for K in the experiments performed in Section 5.

4.4. Solving for $l(\mathbf{x})$

An estimate for $l(\mathbf{x})$ is found by finding the MAP estimate of $s(\mathbf{x})$ (see (2)) using the iterated conditional modes (ICM) algorithm [15]. This process is iterated until the result converges with a maximum of 20 iterations. The pixels are updated using a “checkerboard” scan.

ICM gives a suboptimal estimate of s . The converged estimate represents a local maximum in the posterior PDF. Consequently, a good initialisation of s is necessary to ensure that the converged result is close to the global maximum.

A deterministic temporal discontinuity detector is used to estimate $\mathbf{t}(\mathbf{x})$ and is described by

$$t_k(\mathbf{x}) = \begin{cases} 1 & \text{if } \Delta_k(\mathbf{x}) > \delta_t, k \in \{n-2, n-1, n+1, n+2\}, \\ 0 & \text{otherwise,} \end{cases} \quad (17)$$

where

$$\mathbf{t}(\mathbf{x}) = [t_{n-2}(\mathbf{x}), t_{n-1}(\mathbf{x}), t_{n+1}(\mathbf{x}), t_{n+2}(\mathbf{x})]. \quad (18)$$

From the initial estimate of $\mathbf{t}(\mathbf{x})$, estimates of $s(\mathbf{x})$ and $l(\mathbf{x})$ can be found.

4.4.1. Multiresolution

A multiresolution scheme [16] is incorporated into the algorithm. Using multiresolution results in faster convergence and the state field, $s(\mathbf{x})$, is more likely to converge to the global maximum of the posterior PDF.

A hierarchical pyramid of state fields of differing resolution is constructed with the full resolution field at the bottom level of the pyramid (level 0) and with fields downsampled by a factor of two in each dimension at the next level up in the pyramid. A similar pyramid is constructed for each of the DFDs and for the divergence measure, with all being filtered at each level with a gaussian kernel of variance 2.5 before being downsampled.

The algorithm proceeds by initialising the state field, $s(\mathbf{x})$, at the coarsest level of the pyramid using (17). A new estimate of $s(\mathbf{x})$ at the coarsest level (four levels are used) is obtained from the probabilistic framework and the new estimate is then used to initialise the framework at the level below. This process continues until $s(\mathbf{x})$ has been estimated at full resolution.

An adjustment is made to the prior energy expression when estimating $s(\mathbf{x})$ at the coarser resolution. The weight λ_y is modified to reflect the increased spatial correlation between a pixel and horizontal and vertical neighbours relative to its diagonal neighbours. At a level i of the pyramid, the value of λ_y is given by

$$\lambda_y = \begin{cases} 2^i + \sqrt{2}(2^i - 1) & \text{for vertical and horizontal neighbours} \\ \frac{1}{\sqrt{2}} & \text{for diagonal neighbours.} \end{cases} \quad (19)$$

5. RESULTS

The main contribution of the proposed algorithm is the introduction of a joint Markov random field (MRF) model for blotches and PM (the label field $l(\mathbf{x})$). Pixels can either be clean, a blotch, or PM. Even though blotches can exist at PM sites, such sites cannot be detected as blotches. Effectively, the correct detection rate for blotches is compromised in order to prevent the false detection of blotches due to PM. Another feature of the algorithm is the application of a smoothness prior to the label field $l(\mathbf{x})$ instead of the full state field $s(\mathbf{x})$.

This allows smoothness to be maximised on the field of interest (i.e., $l(\mathbf{x})$) while at the same time minimising the computational cost of estimating the spatial smoothness energy. This facilitates the introduction of a PM bias, which penalises the detection of neighbouring blotch and PM sites by adding an energy penalty K to the assignment of the site as a blotch for every PM site in its neighbourhood (see Figure 5). The other significant features of the algorithm are the divergence likelihood and the outer threshold ratio α_r .

This section presents an evaluation of the proposed algorithm. It compares the performance of the algorithm to other missing data detectors, including conventional missing data detectors such as SDIp [1] and also the Bornard missing data/PM detection algorithm. This is achieved by comparing the results for $l(\mathbf{x})$ generated by each algorithm against a ground truth for blotches. Ground truths for dirt are generated using both infrared scans and by manual detection of blotches. The correct detection and false alarm rates for missing data in the sequence are measured and are arranged in the form of receiver operating characteristics (ROCs), with the control parameter being the value of the threshold δ_t .

However, ROCs cannot give a complete picture of the performance of the detector. Blotch detectors that model pathological motion will perform poorly in terms of correct detections as they are typically designed to prevent detection of blotches in regions of motion estimation failure. Ideally, it would be desirable to ignore ground truth blotches in PM regions when estimating the correct detection rates. However, defining an objective and unbiased ground truth for PM is not a straight forward task. In order to gain a complete understanding of the performance of the algorithm it is necessary to augment the ROC plots with a visual evaluation of the masks generated by the algorithm and also examine the detection rates of PM.

The remainder of this section outlines the experiments used to evaluate the performance of the algorithm. The algorithm is tested on two sequences with PM for which a ground truth for blotches was available and also on sequences for which no ground truth exists. The performance of the proposed algorithm is compared to other blotch detectors in Section 5.4, including conventional detection algorithms and the Bornard algorithm. Finally, a discussion of the results is presented. Before the evaluation begins, the two ground truth sequences are introduced and a description of the experimental procedure is given.

5.1. Ground truth acquisition

The first ground truth sequence is known as the “dance” sequence (see Figure 6). The motion of the dancer in the foreground is pathological. As the dancer turns, her arms become occluded behind the rest of her body. There is also occlusion of the background.

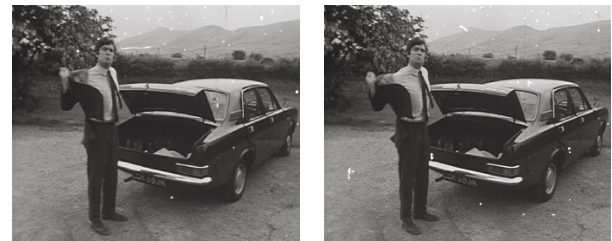
The blotch ground truth for this sequence is acquired using an infrared (IR) scanner. The IR scans were provided by INA [17] as part of the EU Prestospace project [18]. IR scanners are used to produce a transparency map of each frame, where the intensity of each pixel in the map is proportional to the degree of transparency of the pixel. A



FIGURE 6: The “dance” sequence is the first of the ground truth sequences. The movement of the dancer in the foreground is pathological. Its motion is self-occluding which also causes occlusion and uncovering of the background.



(a) 2 frames from the jacket sequence



(b) Frames with blotches highlighted in white

FIGURE 7: The second ground truth sequence is the “jacket” sequence. As the person in the sequence removes his jacket, the motion of the jacket is pathological and is an example of nonrigid object motion. This sequence has a higher blotch frequency than the “dance” sequence.

useful property of dirt is that it is opaque to IR and as a result is associated with low intensities in the transparency map. In the “dance” sequence a ground truth is estimated for 125 frames of 720×576 pixels. The ground truth is obtained by applying a threshold of 150 to each 8-bit transparency map.

Using IR scans is a good way to automatically generate a ground truth for dirt. However, there are a number of drawbacks to using IR scans. IR scans detect dirt even when dirt intensities are similar to the true image intensity. Such dirt cannot be detected by image-based missing data detectors. IR scans will also detect nonimpulsive dirt such as line scratches. As a result of these factors, the reported correct detection rate will be lower than the effective rate.

The second method used to generate a ground truth is to mark blotches in each frame by hand. Acquiring a ground truth in this manner ensures that only perceivable dirt is

TABLE 3: The experimental parameter values for the algorithm.

Parameter	Value
Number of multiresolution levels	4
Number of iterations at each level	20
α_2, α_3	See (9)
Outer threshold Ratio, α_r	0.5
Divergence weight, Λ_d	1
Divergence measure threshold, e_t	50
Roll-off rate, r	0.5
Spatial smoothness weight, Λ_d	1
PM bias, K	5

marked, allowing for a more realistic ground truth. A ground truth for 22 frames of the “jacket” sequence (see Figure 7) was generated in this manner. The resolution of this sequence is also 720×576 . The motion of the jacket is an example of the erratic motion of nonrigid bodies. This sequence also has a higher frequency of blotches than the “dance” sequence.

5.2. Experimental procedure

5.2.1. Motion estimation

The gradient-based motion estimator described by Kokaram in [13] is used to generate motion vectors for all the test sequences. The block size of vector field is 17×17 pixels with a 2 pixel horizontal and vertical overlap.

5.2.2. Configuration of the proposed algorithm

The key parameter of the algorithm is the deterministic threshold δ_t which is used as the control parameter for the ROC plots. The chosen values of δ_t are $\{5, 7, 9, 11, 13, 15, 17.5, 20, 25, 30, 35, 40\}$. These values are also used for the equivalent thresholds in the other missing data detectors used in the comparison outlined in Section 5.4, ensuring that each detector is detected over an equivalent range. The model error σ_e^2 is estimated before the first stage at each multiresolution level by finding the variance of the four DFDs when $s(\mathbf{x}) = 0$. The other relevant parameters are defined in Table 3.

5.2.3. ROC plots

The ROC curves are generated by finding the average correct detection and false alarm rates over all the frames of the test sequences for each value of δ_t . The missing data correct detection rate, r_c , and false alarm rate, r_f , are estimated for each frame (and for each δ_t), and are given by

$$r_c = \frac{\sum_{\mathbf{x}} b_{\text{gt}}(\mathbf{x}) b_{\text{det}}(\mathbf{x})}{\sum_{\mathbf{x}} b_{\text{gt}}(\mathbf{x})}, \quad (20)$$

$$r_f = \frac{\sum_{\mathbf{x}} (1 - b_{\text{gt}}(\mathbf{x})) b_{\text{det}}(\mathbf{x})}{\sum_{\mathbf{x}} (1 - b_{\text{gt}}(\mathbf{x}))},$$



(a) Validation image



(b) PM/blotch detection

FIGURE 8: This figure shows a validation image and a PM/blotch detection for the “jacket” sequence (correct detections—green, missed detections—blue, false alarms—red). The PM/blotch (red/green) images can be used to indicate whether or not a missed blotch detection occurs to the pixel being detected as PM (blue circle).

where $b_{\text{gt}}(\mathbf{x})$ is the binary ground truth blotch mask for the frame and $b_{\text{det}}(\mathbf{x})$ is the binary output of the blotch detector under test. The curve is a parametric plot of corresponding (r_c, r_f) pairs.

5.2.4. PM rate plots

Plots of PM rate, r_p , against the DFD threshold δ_t are evaluated by finding the average relative frequency of PM in a sequence at each threshold. The PM rate is given by

$$r_p = \frac{\sum_{\mathbf{x}} |l(\mathbf{x}) - 2|}{\sum_{\mathbf{x}} 1}, \quad (21)$$

where PM occurs when $l(\mathbf{x}) = 2$.

5.2.5. Visual evaluation

Two types of image are used for the visual evaluation. They are validation images and PM/blotch detections (see Figure 8). Validation images compare the estimated blotch masks to the ground truth and highlight correct detections



FIGURE 9: This figure shows intermediate values of the label field (PM—Red, Blotches—Green) at various iterations of the ICM optimisation process for a frame from the “dance” sequence. The standard parameter set is used (Section 5.2) and the value of δ_t is 7. Level 0 corresponds to the full resolution; the resolution at level 1 is 360×288 and so on. At the coarsest resolution (Level 3) the resolution is 90×72 pixels.

(green), missed detections (blue), and false alarms (red). PM/blotch detections display the masks generated by the proposed algorithm and the Bornard algorithm with PM regions highlighted red and blotches highlighted green.

5.2.6. Implementation of the algorithm

The tests outlined on the proposed algorithm in this paper utilised a C++ implementation of the algorithm, incorporating the IPP libraries [19] where appropriate. The algorithm was tested on a PC with a 1.4 Giga-Hertz Pentium M processor with 1 Giga-Byte of RAM.

5.3. Algorithm evaluation

Figure 9 gives an example of the evolution of the label field configuration for a frame of the “dance” sequence. Detection of missing data and PM at the coarser resolutions allows large PM and blotch features to be detected. As the resolution at which the optimisation occurs increases, the precision of the result increases, allowing small features to be detected. In

general, convergence at the coarser levels is fast as spatial information propagates quickly through the field, but as the resolution increases the speed of convergence decreases. In the example shown in Figure 9, the convergence occurs after six iterations at both Level 3 and Level 2. At the finer resolutions of Level 1 and Level 0, the maximum iteration limit of twenty iterations is reached. Computation time for this example is approximately ten seconds.

Applying the algorithm to all the frames of the test sequences gives the ROC and PM rate curves shown in Figure 10. Overall, the maximum correct detection rates, r_c , for both sequences are quite low (approximately 25% for the “dance” sequence and 40% for the “jacket” sequence). However, there is also a low false alarm rate, r_f , which does not increase above 0.1% (roughly 400 pixels in a 720×576 image) even at low threshold values. The PM rate plots for both test sequences (Figures 10(b) and 10(d)) show that the rate of PM detection increases exponentially as δ_t decreases.

An unusual feature of the ROC plots for the test sequences is the apparent ceiling on the correct detection rates at lower values of the DFD threshold δ_t . The expected

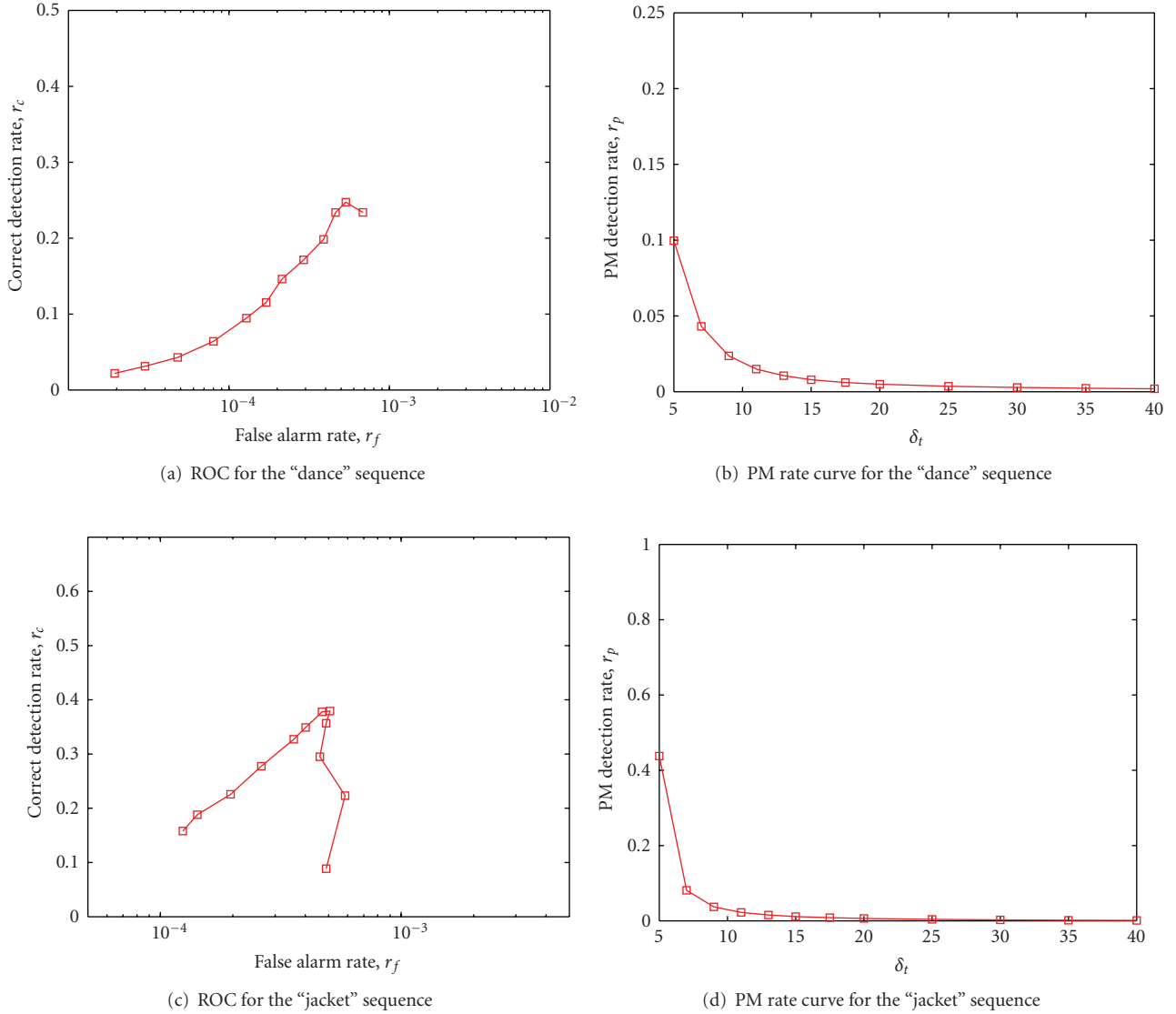


FIGURE 10: This figure shows ROC and PM rate plots (r_p versus δ_t) for both test sequences obtained from the proposed algorithm using the standard set of parameters.

trend in an ROC plot of a blotch detector is to have low correct detection and false alarm rates at higher threshold values and for the rates to increase as the value of the threshold drops, tending to a notional correct detection and false alarm rate of 100% at a zero threshold when every pixel in the image is detected as a blotch. The lowest value of the threshold used therefore gives the highest false alarm and correct detection rate. However, the ROC curves show that this pattern does not apply to the proposed algorithm. While the expected pattern is followed at high values of the threshold δ_t , at lower values of the threshold a value is reached, below which the correct detection rate no longer increases and continues to drop off. Although the rate at which temporal discontinuities are detected in the five-frame window continues to increase, an increasing number of pixels are classified as PM (see Figure 11) since the likelihood of temporal discontinuities at the same location between

multiple frames greatly increases. This is confirmed by the exponentially increasing PM rate at lower threshold values. Consequently, less pixels are detected as missing data and the correct detection rate decreases. For this detector, the notional false alarm rate and correct detection rate at the zero threshold level is 0% as all pixels will be detected as pathological motion.

5.4. Comparison with other missing data detectors

The proposed algorithm is compared with five existing missing data detectors. These include four conventional algorithms and the Bornard missing data detection algorithm [9]. The four other algorithms are the SDIp algorithm [2, 13], the sROD detector [4, 12], the Morris algorithm [2, 5] and the blotch MRF algorithm [6].



FIGURE 11: This figure illustrates how lowering the DFD threshold δ_t results in a lower correct detection rate. The blue circle highlights a large blotch which is not detected at the lower threshold ((a) and (b))—correct detections in green, missed detections in blue). At the lower threshold this blotch is classified as PM ((c) and (d))—red for PM, green for blotches).

5.4.1. Detector configuration

Both the SDIp and sROD algorithms have one associated parameter δ_t and are implemented directly from the description of the algorithms. The four probabilistic algorithms are implemented with as similar parameter values as possible. A multiresolution framework is employed for each algorithm, with a maximum of twenty iterations per level. An eight-pixel first-order spatial neighbourhood is used for either the blotch field (blotch MRF algorithm), the temporal discontinuity field (TDFs) (Morris/Bornard algorithms) or the blotch/PM field (the proposed algorithm). The Bornard algorithm also employs a two-pixel temporal neighbourhood linking the TDFs. Like the proposed algorithm, in the Blotch MRF algorithm a smoothness prior is only enforced on the blotch field and not on the state field.

The value of the spatial smoothness weight in the Morris and Bornard algorithms is 1, and the temporal prior weight of the Bornard algorithm, β_3 , is -0.1 . This reflects the setup recommended by Bornard in his thesis [9, Section 5.7.1.2], where the temporal prior weight is roughly one tenth the size of the spatial weight with a negative value. The search radius in the interpretation stage is 20 pixels with a temporal discontinuity threshold of 1. Although less conservative than the recommended setup (a radius of 38 pixels), PM will be flagged if a temporal discontinuity exists in one of $2 \times 20^2\pi$ pixels which is approximately two and a half thousand pixels. For the proposed algorithm, the standard set of parameters described in Section 5.2 is applied.

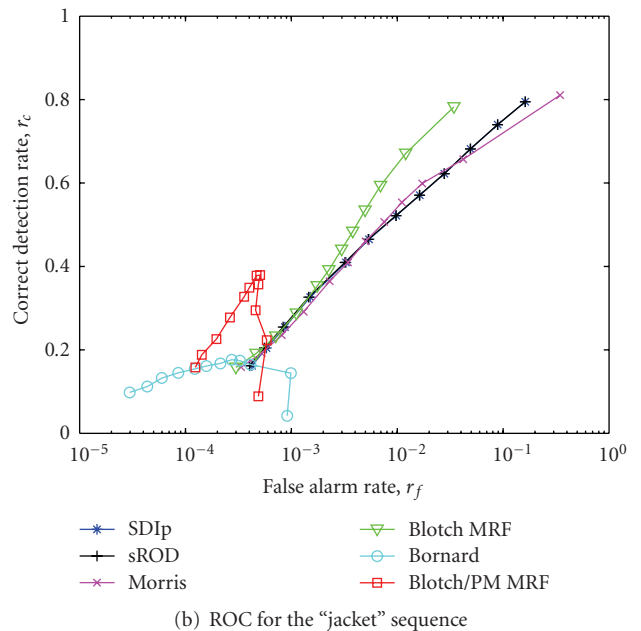
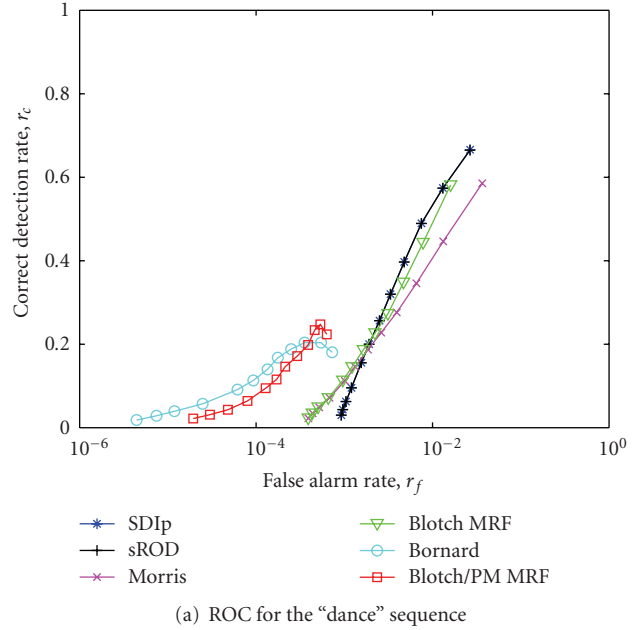


FIGURE 12: This figure shows the ROC curves for the 6 detectors under test for each test sequence. The standard parameter set is used for the proposed algorithm. The set up of each detector is described in Section 5.2.

5.4.2. Comparison with the conventional detectors

ROC curves are plotted for each detector, with the same set of values of δ_t as before ($\delta_t = \{5, 7, 9, 11, 13, 15, 17.5, 20, 25, 30, 35, 40\}$). Comparison (Figure 12) of the missing data detectors shows that both the proposed algorithm and the Bornard algorithm achieve a significantly reduced false alarm rate when compared to the four standard missing data

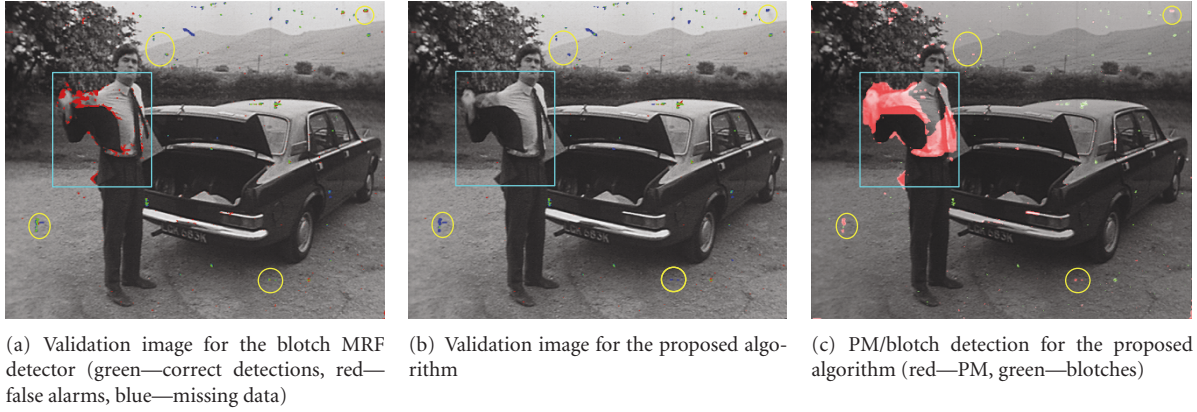


FIGURE 13: This figure consists of validation images for the blotch MRF algorithm and the proposed algorithm for a frame from the “jacket” sequence. The value of the DFD threshold, δ_t , is 15. Also shown is the PM/blotch detection for the proposed algorithm. The blue square highlights the region of the image where the proposed algorithm eliminates false alarms caused by pathological motion. The yellow circles highlight missed blotch detections in the proposed algorithm, which have been wrongly detected as PM.

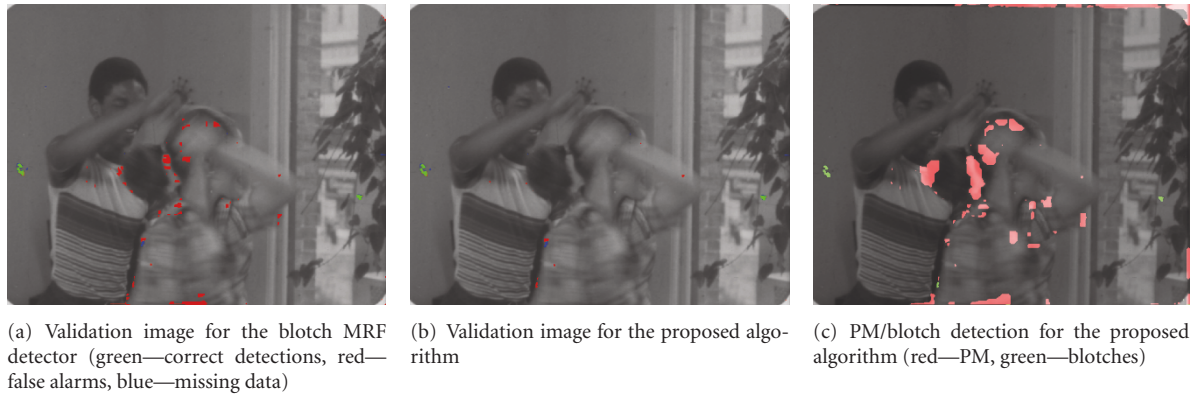


FIGURE 14: This figure consists of validation images for the blotch MRF algorithm and the proposed algorithm for a frame from the “dance” sequence, as well as the corresponding PM/blotch detection for the proposed algorithm. The value of the DFD threshold, δ_t , is 15. Using the proposed algorithm reduces the number of false alarms in the region of the head of the female dancer.

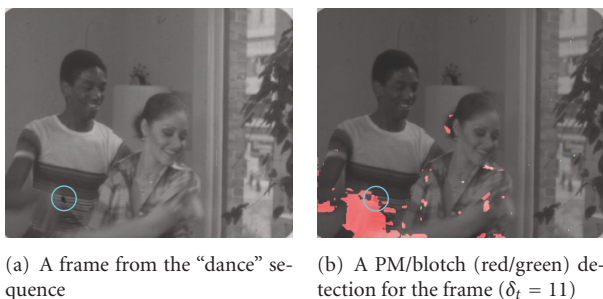


FIGURE 15: This blue circle in the two images highlights a blotch located near a region of motion blur. Due to its location the blotch is detected as PM by the proposed algorithm.

detectors, although the correct detection rates are also reduced. At higher values of the DFD threshold δ_t (i.e., before the maximum correct detection rate is reached), the

false alarm rate of the proposed algorithm is 75% to 90% less for the “dance” sequence and 50% to 80% less for the “jacket” sequence. On the other hand, both the Bornard and proposed algorithms have an upper limit on the rate of correct blotch detections. The average maximum correct detection rate of the proposed algorithm is two to three times less than the correct detection rates at the lowest threshold tested ($\delta_t = 5$), although the false alarm rate is 30 to 50 times lower.

Figures 13 and 14 visually highlight the difference in performance between the proposed algorithm and the blotch MRF algorithm, the standard blotch detector most closely related to the proposed algorithm. In the highlighted frames, the proposed algorithm reduces the number of false alarms in the pathological motion regions (i.e., the jacket and the female dancer), preventing damage to image data. The drop in blotch detections can mainly be attributed to two factors. Firstly blotches which are located near objects undergoing PM are likely to be missed (Figure 15). This is an intended

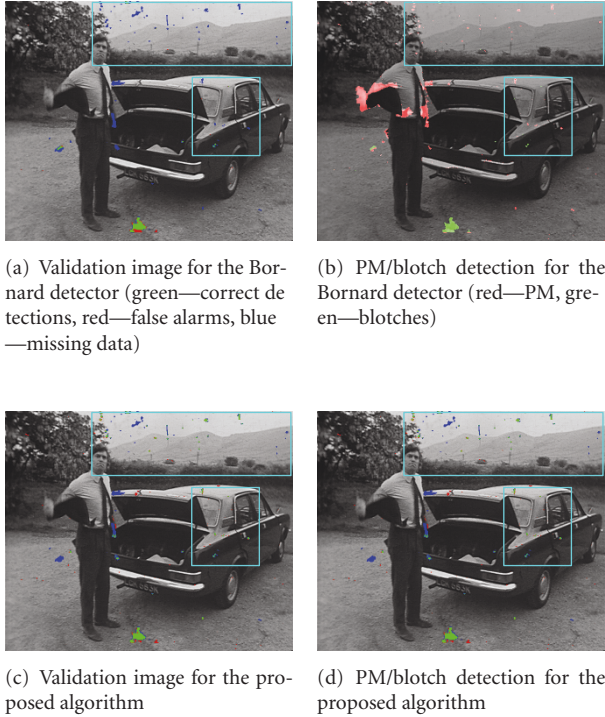


FIGURE 16: This figure illustrates the superior blotch detection rate of the proposed algorithm over the Bornard algorithm in the “jacket” sequence. The highlighted regions of the frame indicate blotches missed by the Bornard algorithm (image (a)) that are detected by the proposed algorithm (c). The blotches missed by the Bornard algorithm are detected as PM (b). The value of δ_t used to generate the images in this figure is 15.

consequence of the algorithm design, which is encouraged to maintain image quality rather than remove the maximum number of blotches. The second major cause of missed detections occurs because of a high frequency of blotches over a number of frames. A fundamental assumption made by missing data detectors is that blotches are rarely located at the same spatial location in neighbouring frames. However, if there is a high frequency of blotches in a sequence, then the probability of blotches occurring at the same location in neighbouring frames increases. In such situations, a blotch is likely to be incorrectly classified as pathological motion.

5.4.3. Comparison with the Bornard algorithm

The missing data ROC curves in Figure 12 also highlight the variable performance of the proposed algorithm when compared with the Bornard algorithm. The Bornard algorithm generally performs better on the “dance” sequence. At high thresholds, the false alarm rate, r_f , of the Bornard algorithm is lower for similar correct detection rates, r_c . However, at lower threshold values the difference in the false alarm rates decreases and the proposed algorithm achieves a higher maximum correct detection rate. On the other hand, the proposed algorithm is much more effective on the “jacket” sequence and achieves a much higher blotch detection rate.

The maximum correct detection rate is approximately 40%, roughly twice the maximum correct detection rate of the Bornard algorithm.

The poor correct detection rates of the Bornard algorithm on the “jacket” is clearly visible in the frame from the “jacket” sequence highlighted in Figure 16. Many of the blotches in the frame are detected as PM by the Bornard algorithm, caused by the high frequency of visible blotches in the sequence. The proposed algorithm is less prone to missed detections in such circumstances.

It is reasonable to conclude that the Bornard detector is more conservative than the proposed algorithm. The interpretation stage of the Bornard is more conservative than the cumulative effect of the divergence likelihood, the PM bias, and the outer threshold ratio. The interpretation stage acts as a “maximum caution” classifier, rejecting any candidate blotch site if a single temporal discontinuity exists within a radius (20 pixels) of the candidate in either of the outer temporal discontinuity fields. The ultra-cautious nature of the interpretation allows the Bornard detector to achieve lower false alarm rates. However, it also results in more missed detections at lower threshold values or when there is a high frequency of blotches in a sequence.

On the other hand, the proposed algorithm tries to make an informed decision and as such is less indiscriminate than the Bornard algorithm. For a site to be detected as PM rather than a blotch, the proposed algorithm requires either a temporal discontinuity to exist in at least one outer temporal discontinuity field at the same location or spatial propagation from neighbouring PM sites by means of the smoothness prior. The divergence likelihood, the PM bias, and the outer threshold are all intended to make the proposed algorithm more likely to detect PM. The divergence likelihood increases the probability of sites with high motion field divergences being detected as PM, the PM bias encourages the spatial propagation of PM states and the outer threshold ratio increases the likelihood of temporal discontinuities being detected in the outer fields.

5.5. Blotch restoration using the proposed algorithm

The critical design criterion of the proposed algorithm is that it eliminates image damage during missing data treatment. A primitive missing data treatment algorithm was implemented that takes a blotch mask from any detector and interpolates detected blotches with the mean motion compensated intensity of the previous and next frames.

Figure 17 compares images restored using both the blotch MRF algorithm and the proposed algorithm ($\delta_t = 15$). Frames from four sequences are shown. The first is the “jacket” sequence used in the ground truth experiments, in which the motion of the jacket is an example of self-occluding and nonrigid object motion. The second sequence shown is called the “vj” sequence and is an example of repetitive occlusions, as the aircraft propellers appear and disappear in alternative frames. The third sequence is the “birds” sequence. This sequence contains another example of repetitive occlusion (the bird’s wings), however, in this case, the region of PM is moving across the frame. The final

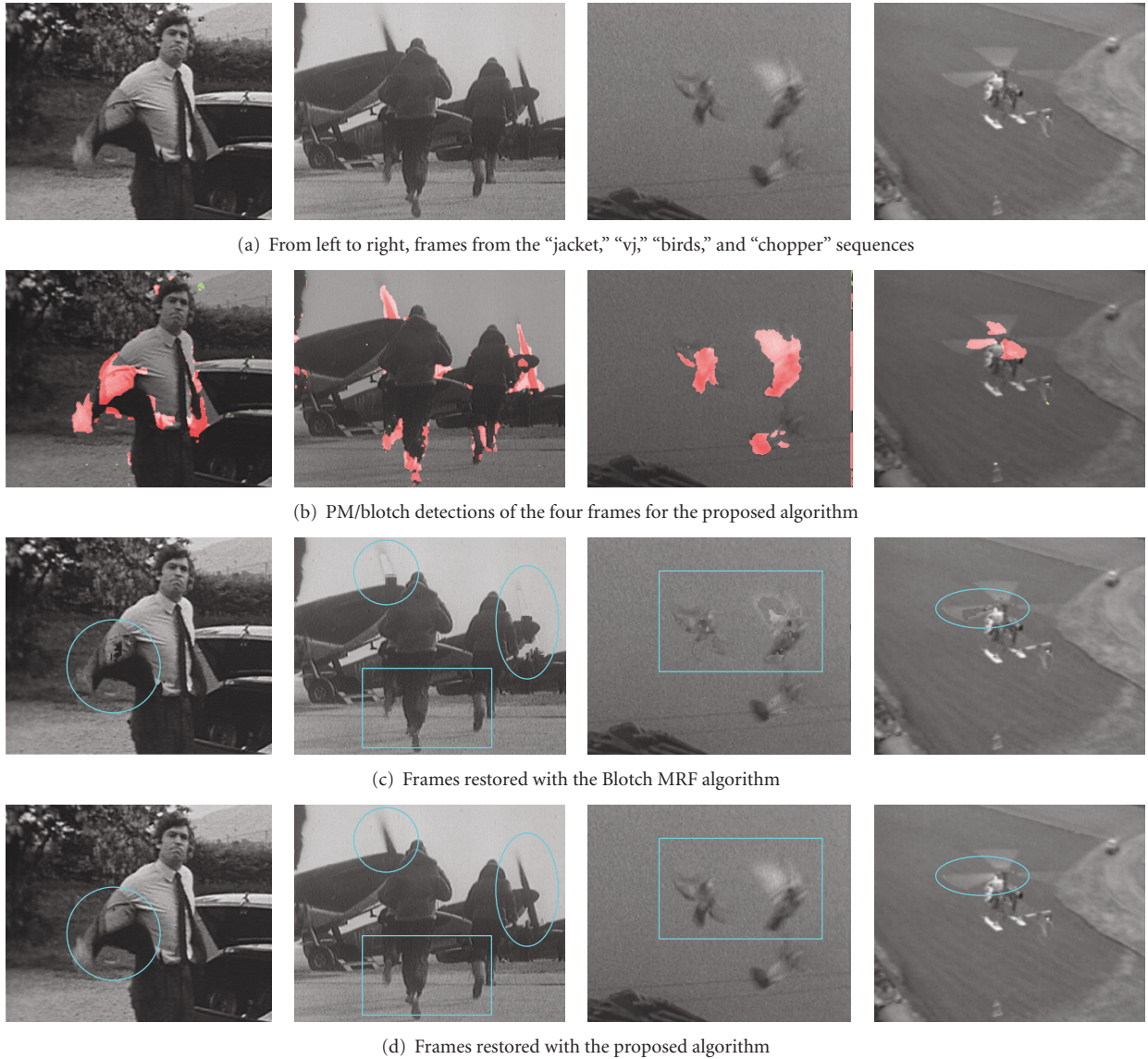


FIGURE 17: This figure shows frames from four sequences restored using the Blotch MRF and proposed algorithms as blotch detectors. The highlighted frames show image damage present if the Blotch MRF is used instead of the proposed algorithm.

sequence (the “chopper” sequence) contains an example of motion blur. The rotation of the rotors is also an issue, as the motion estimator employs a purely translational motion model.

The four examples shown in Figure 17 highlight the effectiveness of the proposed algorithm. Using the proposed algorithm prevents significant damage to each of the four images shown. However, the proposed algorithm is not always able to prevent all image damage, especially in cases where the region of PM is undergoing a translation (i.e., the motion of an object is the sum of PM and a nonpathological translation (see Figure 18)). The assumption is made that motion compensation can be performed accurately, even in regions of PM. Consequently, the algorithm is prone to failure as it involves the motion compensation of four frames with respect to the central frame. However, it should also be

noted that while some image damage may occur, any damage will be much less than the damage caused by performing missing data treatment without detection of PM.

5.6. Computational complexity

Estimation of the prior energy and estimation of the MAP state at each pixel is the most computationally intensive aspect of the algorithm, amounting to approximately 90% of the execution time for a single frame. The upper bound for the complexity of the ICM optimisation is $\mathcal{O}(\# \text{ iterations} \times \# \text{ pixels} \times \# \text{ states} \times \# \text{ neighbours})$ where the number of states is 16 and the number of pixels in a given neighbourhood is eight. The actual order of complexity is slightly reduced as the smoothness prior is enforced on the label field (three states) rather than the full state field (16 states). The number

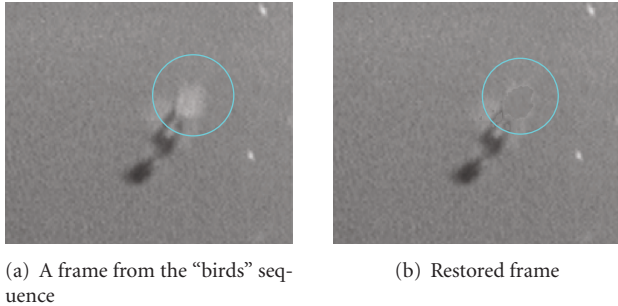


FIGURE 18: This figure shows an example from the “birds” sequence where the proposed algorithm fails to prevent damage to image data (the highlighted regions). In this example, the object (the bird) undergoing PM is also undergoing a translation. This motion causes the algorithm to fail on some frames but not others.

of iterations necessary for convergence also depends on the number of states and the size of the neighbourhood, with a higher number of states and greater neighbourhood size requiring more iterations for convergence. From observation of the algorithm, the size of the δ_i threshold also impacts on the number of iterations required. The average number of iterations performs decreases as δ_i increases. Computation time varies between 5 and 15 seconds per frame, depending on the number iterations performed.

Further reductions in computation would be possible by reducing the complexity of the probabilistic model. For example, the model could be approximated by reducing the number of labels in $l(\mathbf{x})$ to two (i.e., missing data or not missing data) and by using a mixture model to approximate the data likelihood distributions. This would also allow a graph cuts [20] based segmentation to be performed.

6. FINAL COMMENTS

This paper has presented a new robust missing data detection algorithm that aims to reduce false alarms due to pathological motion. It uses a probabilistic framework to detect regions of PM, preventing them from being detected as blotches. A quantitative comparison of the algorithm was presented, comparing the algorithm against four standard blotch detectors and the Bornard PM detection algorithm. The comparison showed that the proposed algorithm gives a significantly reduced false alarm rate when compared with the standard detectors. Although a lower false alarm rate can be obtained with the Bornard algorithm at high thresholds, higher correct detection rates are possible with the proposed algorithm. A further visual evaluation of the algorithm shows how the algorithm can prevent image damage when used as a detector in a missing data treatment framework.

However, the proposed algorithm does not address how the image is to be restored in PM regions. The missing data algorithms outlined in this paper will fail in these regions. Restoration in PM regions has been a topic touched on by Rares in [11] in which an alternative metric for detecting blotches in PM was proposed. As temporal reconstruction of missing data in PM regions is unreliable, spatial reconstruc-

tion techniques are necessary such as image inpainting [21] or texture synthesis [22].

The major area of further development in this area would be to integrate the algorithm into a missing data treatment framework. The algorithm outlined in this paper is presented as a stand-alone blotch detector which operates independently of a missing data interpolation stage. Previous work in missing data treatment [10, 23, 24] has shown that integrating the detection, interpolation, and motion estimation stages into a single Bayesian framework allows for a more visually pleasing restoration. A new missing data treatment algorithm along the lines of the JONDI algorithm [24] using the proposed algorithm for the detection stage could result in an integrated framework resistant to pathological motion.

ACKNOWLEDGMENTS

This work has been supported by the Irish Research Council for Science Engineering and Technology’s Embark initiative, the Adobe corporation, the European Union (EU) Prestospace project, and the European Marie Curie Transfer of Knowledge project Axiom.

REFERENCES

- [1] A. Kokaram, *Motion Picture Restoration: Digital Algorithms for Artefact Suppression in Degraded Motion Picture Film and Video*, chapter 6, Springer, New York, NY, USA, 1998.
- [2] A. Kokaram, R. D. Morris, W. J. Fitzgerald, and P. J. W. Rayner, “Detection of missing data in image sequences,” *IEEE Transactions on Image Processing*, vol. 4, no. 11, pp. 1496–1508, 1995.
- [3] M. J. Nadenau and S. K. Mitra, “Blotch and scratch detection in image sequences based on rank order differences,” in *Proceedings of the 5th International Workshop on Time-Varying Image Processing and Moving Object Recognition*, Florence, Italy, September 1996.
- [4] J. Biemond, P. M. B. van Roosmalen, and R. L. Lagendijk, “Improved blotch detection by postprocessing,” in *Proceedings of the International Conference on Acoustics, Speech, and Signal Processing (ICASSP ’99)*, vol. 6, pp. 3101–3104, Phoenix, Ariz, USA, March 1999.
- [5] R. D. Morris, “Image sequence distribution using Gibbs distributions,” Ph.D. thesis, Cambridge University, Cambridge, UK, 1995.
- [6] B. Kent, A. Kokaram, B. Collis, and S. Robinson, “Two layer segmentation for handling pathological motion in degraded post production media,” in *Proceedings of IEEE International Conference on Image Processing (ICIP ’04)*, vol. 1, pp. 299–302, Singapore, October 2004.
- [7] J. Ren and T. Vlachos, “Dirt detection for archive film restoration using an adaptive spatio-temporal approach,” in *Proceedings of the 2nd European IEE Conference on Visual Media Production (CVMP ’05)*, pp. 221–230, London, UK, November 2005.
- [8] A. Rares, M. J. T. Reinders, and J. Biemond, “Statistical analysis of pathological motion areas,” in *Proceedings of IEE Seminar on Digital Restoration of Film and Video Archives*, pp. 1–28, London, UK, January 2001.

- [9] R. Bornard, "Probabilistic approaches for the digital restoration of television archives," Ph.D. thesis, Ecole Centrale Paris, Paris, France, 2002.
- [10] A. Kokaram and S. J. Godsill, "MCMC for joint noise reduction and missing data treatment in degraded video," *IEEE Transactions on Signal Processing*, vol. 50, no. 2, pp. 189–205, 2002.
- [11] A. Rares, M. J. T. Reinders, and J. Biemond, "Complex event classification in degraded image sequences," in *Proceedings of IEEE International Conference on Image Processing (ICIP '01)*, vol. 1, pp. 253–256, Thessaloniki, Greece, October 2001.
- [12] P. M. B. van Roosmalen, "Restoration of archived film and video," Ph.D. thesis, Delft University of Technology, Delft, The Netherlands, 1999.
- [13] A. Kokaram, *Motion Picture Restoration: Digital Algorithms for Artefact Suppression in Degraded Motion Picture Film and Video*, chapter 2, Springer, New York, NY, USA, 1998.
- [14] J. Konrad and E. Dubois, "Bayesian estimation of motion vector fields," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 14, no. 9, pp. 910–927, 1992.
- [15] J. E. Besag, "On the statistical analysis of dirty pictures," *Journal of the Royal Statistical Society B*, vol. 48, no. 3, pp. 259–302, 1986.
- [16] F. Heitz, P. Perez, and P. Bouthemy, "Multiscale minimization of global energy functions in some visual recovery problems," *CVGIP: Image Understanding*, vol. 59, no. 1, pp. 125–134, 1994.
- [17] Institut national de l'audiovisuel, <http://www.ina.fr/index.en.html>.
- [18] "Prestospace: an integrated solution for audio-visual preservation and access," <http://www.prestospace.org/>.
- [19] Intel performance primitive (ipp) libraries, <http://www.intel.com/cd/software/products/asm-na/eng/302910.htm>.
- [20] Y. Y. Boykov and M.-P. Jolly, "Interactive graph cuts for optimal boundary & region segmentation of objects in N-D images," in *Proceedings of the 8th IEEE International Conference on Computer Vision (ICCV '01)*, vol. 1, pp. 105–112, Vancouver, Canada, July 2001.
- [21] A. Rares, M. J. T. Reinders, and J. Biemond, "Edge-based image restoration," *IEEE Transactions on Image Processing*, vol. 14, no. 10, pp. 1454–1468, 2005.
- [22] A. A. Efros and T. K. Leung, "Texture synthesis by non-parametric sampling," in *Proceedings of the 7th IEEE International Conference on Computer Vision (ICCV '99)*, vol. 2, pp. 1033–1038, Kerkyra, Greece, September 1999.
- [23] A. Kokaram, *Motion Picture Restoration: Digital Algorithms for Artefact Suppression in Degraded Motion Picture Film and Video*, chapter 7, Springer, New York, NY, USA, 1998.
- [24] A. Kokaram, "On missing data treatment for degraded video and film archives: a survey and a new Bayesian approach," *IEEE Transactions on Image Processing*, vol. 13, no. 3, pp. 397–415, 2004.