

# On Missing Data Treatment for Degraded Video and Film Archives: A Survey and a New Bayesian Approach

Anil C. Kokaram, *Member, IEEE*

**Abstract**—Image sequence restoration has been steadily gaining in importance with the increasing prevalence of visual digital media. The demand for content increases the pressure on archives to automate their restoration activities for preservation of the cultural heritage that they hold. There are many defects that affect archived visual material and one central issue is that of Dirt and Sparkle, or “Blotches.” Research in archive restoration has been conducted for more than a decade and this paper places that material in context to highlight the advances made during that time. The paper also presents a new and simpler Bayesian framework that achieves joint processing of noise, missing data, and occlusion.

**Index Terms**—Autoregressive models, Bayesian inference, composition sampling, factored sampling, film and video post production, Gibbs sampling, image processing, marginalization, Markov chain Monte Carlo, missing data reconstruction, motion estimation, noise reduction, video processing, video restoration.

## I. INTRODUCTION

WITHIN the last five years, there has been an explosion in the exploitation and availability of digital visual media. Digital television has been widely available in Europe for the last three years, and Internet usage continues to grow as does the availability of MPEG(1, 2, 4), AVI audio/video clips through the increasing use of streaming media. DVD (Digital Video/Versatile Disk) usage is growing faster than CD-audio usage did when it was first introduced. There is now growing interest in Digital Cinema implying that the whole chain from “film” shooting to distribution/projection will be digital.

With all these available digital video “channels,” it is amusing to note that the main concern for broadcasters is the relative unavailability of content. Holders of large video, film, and photograph archives for instance the BBC (U.K.), Institut National de L’Audiovisuel (INA France), Radio Televisão Portuguesa (RTP Portugal) find that archive material is in increasing demand. However, the material is typically degraded due to physical problems in repeated projection or playback or simply the chemical decomposition of the original material. Typical problems with much of the archived film material have been increased level of noise, and Dirt and Sparkle due to the depo-

sition of dust or the abrasion of the material. Of course, there are many more problems specific to the media, e.g., Two Inch Tape Scratches affecting two inch video tape and Vinegar Syndrome, a Moire affecting film and the film scanning process.

In order to preserve and exploit this material, these defects must be removed so that the picture quality can be restored. Because of the large amount of data, manual retouching is impractical. Therefore, automated techniques have become important. Furthermore, it has been recognized that the reduction of *noise* in particular before MPEG compression allows a more efficient usage of the available digital bandwidth [1]–[3].

Hence, the area of automated restoration of image sequences has moved from being principally a Signal Processing research topic to one of more widespread significance. As an example of increased industrial significance, at the International Broadcast Convention<sup>1</sup> held in Amsterdam, September 2002, there were no less than nine companies presenting solutions for automated digital restoration: Philips, Thomson, Diamant [4], DaVinci, Silicon Graphics, Apple, Discreet, Snell and Wilcox [5], and MTI. Projects such as AURORA (Automated Restoration of Original Film and Video Archives) and BRAVA (Broadcast Restoration of Archives by Video Analysis) funded by the EU are further examples of this increased relevance.

This paper provides both a review of the research work that has in part given rise to these new systems, and also considers a unifying framework for missing data treatment. Before moving on to the central artefact addressed this paper, it is educational to briefly consider the wide range of problems that exist.

### A. A Brief Taxonomy of Defects

There has been some effort by the BRAVA consortium (<http://brava.ina.fr>) during 2000–2002 to attempt to standardize or rather educate the community about the names and types of defects that can occur in archived video and film. An exhaustive description is given at [http://brava.ina.fr/brava\\_public\\_impairments\\_list.en.htm](http://brava.ina.fr/brava_public_impairments_list.en.htm). This was compiled by the project leader, J.-H. Chenot with input from the various archives in the project, e.g., BBC, INA, and RTP. This kind of taxonomy is notably missing from the recent literature. Figs. 1–10 present part of this taxonomy of defects in an attempt to educate the IEEE readership as to the names given to the defects as they are used in the archive industry. The methods considered in this paper address the problems illustrated by Figs. 1, 2, 4, and 6. Massive loss of data as in Figs. 3 and 7 are better treated

Manuscript received March 4, 2003; revised October 8, 2003. This work was supported in part by the EU under Projects MOUMIR HPRN-CT1999-00108 and BRAVA IST-1999-11628 as well as Enterprise Ireland Projects MUSEDTV, CASMS, and DYSVIDEO. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Ioannis Pitas.

The author is with the Electrical Engineering Department, Trinity College, Dublin 2, Ireland (e-mail: [anil.kokaram@tcd.ie](mailto:anil.kokaram@tcd.ie)).

Digital Object Identifier 10.1109/TIP.2004.823815

<sup>1</sup>One of the two major television technology conferences in the world.



Fig. 1. *Dirt and Sparkle* occurs when material adheres to the film due to electrostatic effects (for instance) and when the film is abraded as it passes through the transport mechanism. It is also referred to as a *Blotch* in the literature. The visual effect is that of bright and dark flashes at localized instances in the frame. The image indicates where a piece of Dirt is visible.



Fig. 2. *Film Grain Noise* is a common effect and is due to the mechanism for the creation of images on film. It manifests slightly differently depending on the different film stocks. The image shows clearly the textured visible effect of noise in the blue sky at the top left. Blotches and noise typically occur together and are the main form of degradation found on archived film and video. A piece of Dirt is indicated on the image.

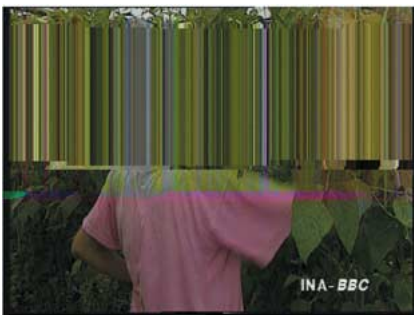


Fig. 3. *Betacam Dropout* manifests due to errors on Betacam tape. It is a missing data effect and several field lines are repeated for a portion of the frame. The repeating field lines are the machine's mechanism for interpolating the missing data.

through temporal frame interpolation and the reader may see [6] and [7] for some treatment of this issue. A possible solution for Kinescope Moire illustrated in Fig. 5 can be found in [8] and [9]. *Two Inch scratches* (caused by scratching of old two inch video tape) are an example of a specialized missing data problem and a treatment can be found in [10] and [11].

Two major defects are missing from the visual description: Shake and Flicker. Those are best viewed as video clips. *Shake* simply refers to unwanted global motion of the picture caused



Fig. 4. *Digital Drop Out* occurs because of errors on digital video tape. This example is drop out from D1 tape.



Fig. 5. *Kinescope Moire* is caused by aliasing during Telecine conversion and manifests as rings of degradation that move slightly from frame to frame.



Fig. 6. *Film Tear* is simply the physical tearing of a film frame sometimes due to a dirty splice nearby.



Fig. 7. *Vinegar Syndrome* often results in a catastrophic breakdown of the film emulsion. This example shows long strands of missing data over the frame.

either by camera movement or problems during scanning. Algorithms for removing shake abound [12]–[17]. This is principally



Fig. 8. *Echoes and Overshoots* manifest as image shadows slightly displaced from each object in the frame. When the effect is severe it is called Echo and when it is just limited to edges it is called Overshoot, as in this case.



Fig. 9. *Color Fading* implies that the picture color is not saturated enough, giving the image a washed out look.

because it is related to the global motion estimation problem that is also important for video compression issues [18]–[23].

*Flicker* manifests as a fluctuation in picture brightness from frame to frame. In stills the effect is very difficult to observe indeed, but as a sequence the effect is often very disturbing. Two different types of degradations result in a perceivable *flicker* artefact. The first realistic de-flicker algorithm was developed by Roosmalen [24], [25] and a real-time hardware version was developed by Snell and Wilcox in the late 1990s. Both changing film exposure (in old silent movies for instance) and varying lighting conditions result in luminance fluctuations. However, a misalignment in the two optical paths in a telecine machine also yield the same visible artefact, called *Twin Lens Flicker*. In that case, the two fields of each interlaced TV frame are incorrectly aligned with respect to the original film frame, and the observed fluctuations are due more to the *shake* between the fields than any real luminance changes. Vlachos *et al.* considered this problem in [26] and a real-time implementation was also developed by Snell and Wilcox in the late 1990s.

Video clips showing serious degradation by shake, flicker, lines, grain and blotches can be seen at <http://www.mee.tcd.ie/~sigmedia/ifc/ifc.html>. The book by Read and Meyer [27] gives an excellent coverage of the physical nature of archive material, and the practices in the film preservation industry.

### B. Missing Data Treatment in General

This paper concentrates on a central issue in automated archive restoration, Missing Data. It manifests as Dirt and Sparkle (Fig. 1), Dropout (Figs. 3 and 4), Film Tear (Fig. 6),

and Dirty Splices, Vinegar Syndrome (Fig. 7) some examples of which have been illustrated in the previous section. Line scratches can sometimes represent missing data, but quite often there is still data in the defect region. Line scratch work can be found in [6] and [28]–[32] and is not considered further here. Assuming that the missing data does not occur in the same location in consecutive frames, it seems sensible that a repair of the damaged region can be achieved by copying the relevant information from previous or next frames. This relies on the heavy temporal redundancy present in the image sequence. Because this redundancy is prevalent only along motion trajectories, motion estimation has become a key component of missing data treatment systems.

Historically, the approach has been to develop a method to detect the defect [33], then to correct it by some kind of spatio-temporal interpolation activity [34]–[37]. The traditional concept in detection is to assume that any set of pixels that cannot be located in next and previous frames, must represent some kind of impulsive defect and should be removed. This requires some kind of matching criterion and could be dealt with via a model-based approach [6], [31] or several clever heuristics [6], [38], [39], [31]. Recently, the interaction between the motion estimation and missing data detection/correction stages has been receiving more attention [40], [41]. A key difference between missing data as it appears in real footage and *speckle* degradation is that blotches are almost never limited to a single isolated pixel.

The review of missing data treatment therefore begins with a consideration of heuristics for treating this problem, then moves on to review and extend the model based approaches. Finally, consideration is given to a major practical shortcoming in the work to date, that of coping with pathological motion.

Note that in all the work done in this area thus far, the model of the clean original image sequence has always been adopted in the form as follows:

$$I_n(\vec{x}) = I_{n-1}(\vec{x} + \mathbf{d}_{n,n-1}(\vec{x})) + e(\vec{x}) \quad (1)$$

where the luminance at pixel site  $\vec{x} = [i, j]$  in frame  $n$  is  $I_n(\vec{x})$ , and the two-component motion vector mapping site  $\vec{x}$  in frame  $n$  into frame  $n - 1$  is  $\mathbf{d}_{n,n-1}(\vec{x})$ . The model error follows a Gaussian distribution  $e(\cdot) \sim \mathcal{N}(0, \sigma_e^2)$ . The model is therefore consistent with luminance conserving, translational motion in the sequence.

## II. HEURISTICS FOR DETECTION

Perhaps the earliest work on designing an automatic system to “electronically” detect Dirt and Sparkle was undertaken by Storey at the BBC [38], [42] as early as 1983. The design was incorporated directly into hardware which was subsequently used in-house for video restoration before broadcast. The idea was to flag a pixel as missing if the forward and backward pixel difference was high. This idea was, of course, beset with problems in parts of the image where motion occurred. The natural extension of this idea was presented by Kokaram around 1993 [6], [33] which allowed for motion compensated differences. That type of detector was called a “Spike Detection Index” (SDI) and the most useful are defined as follows.



Fig. 10. Line Scratches manifest in much archived footage. They also occurs due to accidents in film developing. The color of the scratch depends on which side of the film layer it occurs. It is often the case that not all the image information inside the scratch is lost. They are a challenge to remove because they persist in the same location from frame to frame.

### A. SDIx

The forward and backward motion compensated pixel differences  $E_f, E_b$  of the observed, corrupted image sequence  $G_n(\vec{x})$  are defined as follows.

$$\begin{aligned} E_b &= G_n(\vec{x}) - I_{n-1}(\vec{x} + \mathbf{d}_{n,n-1}(\vec{x})) \\ E_f &= G_n(\vec{x}) - I_{n+1}(\vec{x} + \mathbf{d}_{n,n+1}(\vec{x})). \end{aligned} \quad (2)$$

Note that the previous and next frames are assumed to be *uncorrupted* at the required motion compensated sites, hence  $G_{n-1} = I_{n-1}$ , etc. Two detectors can then be proposed [6] as follows:

$$b_{\text{SDIa}}(\vec{x}) = \begin{cases} 1, & \text{for } (|E_b| > E_t) \text{ AND } (|E_f| > E_t) \\ 0, & \text{otherwise} \end{cases} \quad (3)$$

$$b_{\text{SDIp}}(\vec{x}) = \begin{cases} 1, & \text{for } (|E_b| > E_t) \text{ AND } (|E_f| > E_t) \\ & \text{AND } \text{sign}(E_f) = \text{sign}(E_b) \\ 0, & \text{otherwise} \end{cases} \quad (4)$$

where  $b(\cdot)$  is a detection field variable set to 1 at sites that are corrupted by missing data.  $E_t$  is a user defined threshold for detection of discontinuity. The SDIa is based on thresholding  $E_f, E_b$  only. SDIp additionally applies the constraint that if corruption does not occur in identical locations in consecutive frames and the model in (1) holds,  $I_{n-1} \approx I_{n+1}$ , and one should expect that the sign of the difference signals should be the same. It is now accepted that SDIp is the better detector in almost all situations because of this additional constraint.

### B. ROD

In 1996, Nadenau and Mitra [39] presented another scheme which used a spatio-temporal window for inference: the Rank Order Detector (ROD). It is generally more robust to motion estimation errors than any of the SDI detectors although it requires the setting of three thresholds. It uses some spatial information in making its decision. The essence of the detector is the premise that blotched pixels are outliers in the local distribution of intensity.

Defining a list of pixels as

$$\begin{aligned} p_1 &= I_{n-1}(\vec{x} + \vec{d}_{n,n-1}(\vec{x})) + [0 \ 0] \\ p_2 &= I_{n-1}(\vec{x} + \vec{d}_{n,n-1}(\vec{x})) + [0 \ 1] \\ p_3 &= I_{n-1}(\vec{x} + \vec{d}_{n,n-1}(\vec{x})) + [0 \ -1] \\ p_4 &= I_{n+1}(\vec{x} + \vec{d}_{n,n+1}(\vec{x})) + [0 \ 0] \end{aligned}$$

$$\begin{aligned} p_5 &= I_{n+1}(\vec{x} + \vec{d}_{n,n+1}(\vec{x})) + [0 \ 1] \\ p_6 &= I_{n+1}(\vec{x} + \vec{d}_{n,n+1}(\vec{x})) + [0 \ -1] \\ I_c &= I_n(\vec{x}) \end{aligned} \quad (5)$$

where  $I_c$  is the pixel to be tested, the algorithm may be enumerated as follows.

- 1) Sort  $p_1$  to  $p_6$  into the list  $[r_1, r_2, r_3, \dots, r_6]$  where  $r_1$  is minimum. The median of these pixels is then calculated as  $M = (r_3 + r_4)/2$ .
- 2) Three motion compensated difference values are calculated as follows:  
If  $I_c > M$

$$\begin{aligned} e_1 &= I_c - r_6 \\ e_2 &= I_c - r_5 \\ e_3 &= I_c - r_4 \end{aligned}$$

If  $I_c \leq M$

$$\begin{aligned} e_1 &= r_1 - I_c \\ e_2 &= r_2 - I_c \\ e_3 &= r_3 - I_c \end{aligned}$$

- 3) Three thresholds are selected  $t_1, t_2, t_3$ . If any of the differences exceeds these thresholds, then a blotch is flagged as follows:

$$\begin{aligned} b_{\text{ROD}}(\vec{x}) &= \begin{cases} 1, & \text{if } (e_1 > t_1) \text{ OR } (e_2 > t_2) \text{ OR } (e_3 > t_3) \\ 0, & \text{otherwise} \end{cases} \end{aligned}$$

where  $t_3 \geq t_2 \geq t_1$ . The choice of  $t_1$  is the most important. The detector works by measuring the ‘‘outlierness’’ of the current pixel when compared to a set of others chosen from other frames. The choice of the shape of the region from which the other pixels were chosen is arbitrary.

### C. Image Histograms

Many authors have used the image histogram to detect abrupt changes in image sequences. This idea can be extended to detect very large areas of missing data in degraded archive frames. It is totally unsuitable for detection of all but the largest defects since otherwise the changes in consecutive histograms are not

noticeable. In [43] and [44] Kokaram *et al.* present a mechanism for isolating the image quadrant that contains the defect. The attractive aspect of this proposal is that no motion estimation is necessary, hence the computational cost is extremely low.

Define 5 quadrants in each frame, having a size of  $(N/2) \times (M/2)$  columns and rows<sup>2</sup>, respectively. The top left hand corner of each quadrant (0, 1, 2, 3, 4) is located at positions  $[0, 0], [N/2, 0], [0, M/2], [N/2, M/2], [N/4, M/4]$  in the image. Denote the  $k$ th bin of the histogram of the  $q$ th quadrant of the  $n$ th frame as  $H_n^q(k)$ . Two differences,  $\Delta_{n,n-1}^q = \sum_k |H_n^q(k) - H_{n-1}^q(k)|$  and similarly for  $\Delta_{n,n+1}^q$  can be calculated for each quadrant. When both forward and backward differences are larger than a chosen threshold, that quadrant can be marked as heavily damaged. This is a very coarse detection process indeed and is not evaluated further here. Such a coarse detection process can be used by restoration operators to browse a video sequence for the most severely damaged frames.

#### D. Morphological/Median Filter Approaches

In the one-dimensional (1-D) case, Paisan and Crise [45] were the first to spot that one could use a median filtered signal as a *rough estimate* of the original signal before corruption by impulsive noise. The difference between the observed, degraded signal and this rough estimate would be high at sites corrupted by impulsive defects. This is because the rank order filter removes all outliers, but leaves lower scale trends untouched. This idea can be extended to treat small missing data artifacts in archived video and film, known as *Dust*. These are generally just a few pixels in area ( $3 \times 3$  pixels), and hence only a small median or morphological window need be used. Using a larger window to detect larger artifacts causes problems since more true image detail would then be removed causing an increased number of false alarms. Joyeux, Buisson, Decenciere, Harvey, Tenze, Saito, Boukir *et al.* have been implementing these types of techniques for film restoration since the mid-1990s [31], [46]–[54]. Joyeux [31] points out that these techniques are particularly attractive because of their low computational cost. They perform well when the artefact is small, and surrounded by a relatively low-activity homogenous region. The high resolution of film scans is therefore suitable for these tools.

### III. MISSING DATA RECONSTRUCTION

Having detected an area of missing data, it is necessary to synthesize material to fill the gap. Storey [42], [38] in 1983 used a three tap median operation to interpolate the missing pixel. This was implemented without motion information. In 1993, Kokaram extended the idea to motion compensated filtering, but recognizing that motion estimation errors resulted in very poor performance, introduced a three-dimensional (3-D) median filtering operation on a  $3 \times 3 \times 3$  motion compensated pixel volume around each missing pixel [34]. A model based scheme that used 3-D AR models to synthesize texture in the gap was also presented in [34]. Model based pixel cut and paste operations from previous or next frames then followed that allowed

for occlusion and uncovering. Deterministic frameworks were proposed circa 1996 [55], [56], while a Bayesian cut-and-paste method was proposed by Roosmalen *et al.* [36] in 1999. The essence of all these ideas was to ensure that interpolated pixel data was smooth both in time and space.

Since 2000, there has been an increasing interest in the concept of spatial filling in of gaps in images. The term, *inpainting* was used to describe this idea by Ballester and Bertalmio *et al.* [57], [58]. In 2002, similar ideas emerged based on 2-D AR interpolation [59], [60]. The remarkable results of Efros and Freeman [61], [62] have also received considerable attention. One could consider that these techniques can be used to reconstruct small missing data regions in a degraded image. Bornard [63], [64], in his thesis of 2002, has considered using these techniques as an alternative to volumetric image data reconstruction, because of the problems associated with motion estimation in difficult situations. Remarkable results have been achieved on relatively small corruption without the need for motion interpolation.

Spatial morphological operators as well as deterministic spatial interpolation can be used to good effect to reconstruct missing data for small Blotches (called “dust”). However, these methods tend to work best when the region to be reconstructed is relatively homogenous [31]. For high-resolution film applications in particular, these methods are quite appropriate for handling many of the smaller defects.

In the interest of brevity, an exhaustive comparison of spatial versus spatio-temporal interpolation is not undertaken here. It is important to note though that for large missing regions, e.g., half of a frame, spatial interpolation is unable to cope, while for small regions especially in homogenous areas, spatial interpolation is ideal.

### IV. BAYESIAN APPROACH

One problem with the above approaches to missing data treatment is that they each address single issues separately and do not acknowledge the interaction between their operation. For instance, all of the detectors above cause high-rate false alarms in the presence of poor motion estimation. Yet, the presence of missing data can cause poor motion estimation performance. In addition, reconstruction using any of the volumetric techniques, especially for large pieces of missing data, is poor when motion estimation is poor. Poor detection generally leads to poor reconstruction since the detector is probably flagging areas which cannot be properly reconstructed. To give good performance, therefore, the motion estimate must be robust to the presence of missing data, and the detector should work in tandem with the interpolator in some way to evaluate whether more damage is being introduced than removed.

Another problem is that none of the above approaches incorporate all that is known about the nature of the defects in question. For instance, missing regions would generally be smooth areas that are spatially coherent in some sense. They also usually do not occur in the same place in two consecutive frames. There are indeed many post-processing schemes that can achieve this information insertion via clever filtering operations or morphological dilation/erosion pairs [6], [35], but it is useful to consider

<sup>2</sup>Where there are  $N$  columns and  $M$  rows in an image frame.



how this information could be incorporated in the problem definition itself.

Model-based schemes could provide one solution to this issue, and schemes for missing data detection and interpolation were reviewed and compared with deterministic approaches in [6], [33], and [34]. Morris [65] was the first (1994) to propose using Bayesian approach to resolve this *data fusion* aspect of combining different sorts of *prior* information about Blotches. The Bayesian approach has since been evolving [6], [36], [41], [59], [66], [67] into a unifying framework that treats motion, missing data and noise jointly. The book [6] contains an exhaustive discussion of a joint detection/interpolation scheme for Blotches.

The main problems with the previous approaches presented in [6] and [41] are computational complexity and the failure in cases of motion discontinuity. This can be solved by unifying the problems of motion estimation, picture reconstruction with occlusion estimation. The next sections in the paper present this new formulation of the Bayesian framework for missing data treatment. This new formulation is of lower complexity than that presented in [41] because it is based on pixelwise models instead of a volumetric stochastic processes.

#### A. Quantifying the Problem

There are two effects to be modeled as far as missing data in archives is concerned.

**Replacement noise** completely obliterates the underlying image pixels at certain pixel positions in the image sequence. The replacement process typically occurs in contiguous patches within a single video frame, and will thus be referred to as the “Blotch” process (Figs. 1, 2, and 15, show good examples of blotches in a video sequence). The Blotch process at pixel location  $\vec{x}$  in a particular frame is fully specified by random variables  $b(\vec{x}) \in \{0, 1\}$  and  $c(\vec{x}) \in \mathfrak{R}$ . The first of these,  $b(\vec{x})$ , is a switching process that determines whether replacement noise is present at  $\vec{x}$  (set  $b(\vec{x}) = 1$ ) or absent (set  $b(\vec{x}) = 0$ ). The second,  $c(\vec{x})$ , gives the pixel intensity of the replacement noise at  $\vec{x}$ .

**Random additive noise.** Also present in a typical frame of video is random additive noise  $\mu(\vec{x})$ , modeled here as a Gaussian i.i.d. process with  $\mu(\cdot) \sim \mathcal{N}(0, \sigma_\mu^2)$ .

The model for degradation is a *Replacement* process with additive noise. A binary field  $b(\vec{x})$  is introduced that is 1 at a site of missing data and zero, otherwise. The degradation model can then be written as follows:

$$G_n(\vec{x}) = (1 - b(\vec{x}))I_n(\vec{x}) + b(\vec{x})c(\vec{x}) + \mu(\vec{x}) \quad (6)$$

where  $\mu(\cdot) \sim \mathcal{N}(0, \sigma_\mu^2)$  is the additive noise, and  $c(\vec{x})$  is a field of random variables that cause the corruption at sites where  $b(\vec{x}) = 1$ . This formulation models the *degradation* of the clean images, and therefore  $\mu$  is not the same as  $e$  in (1). The noise in (1) attempts to quantify the uncertainty in the image sequence model that relates *clean* frames to each other. Unfortunately, it is extremely difficult to keep these two effects separate in any solution to this problem.

Three distinct (but interdependent) tasks can now be identified in the restoration problem. The ultimate goal is image

estimation, i.e., revealing  $I_n(\cdot)$  given the observed missing and noisy data. The missing data detection problem is that of estimating  $b(\vec{x})$  at each pixel site. The noise reduction problem is that of reducing  $\mu(\vec{x})$  *without* affecting image details. The replacement model was employed within a nonprobabilistic framework by Kokaram *et al.* in [68] for image sequences, and *implicitly* employed in a two stage Bayesian framework for missing data detection and interpolation by Morris *et al.* [33], [69].

#### B. Image Sequence Model: Pixel States and Occlusion

To incorporate occlusion into the image sequence model in (1), a hidden field of binary variables  $O_b(\vec{x})$  and  $O_f(\vec{x})$  is introduced between frames  $n, n - 1$  and  $n, n + 1$ , respectively. When  $O_b(\vec{x}) = 1$ , this implies that the data at site  $\vec{x}$  in frame  $n$  does not exist in the frame  $n - 1$ . This represents *Occlusion in the backward temporal direction*. A similar situation exists in the forward direction with  $O_f(\vec{x})$ .

Degradation information can be included in the pixel site information by defining a pixel as occupying one of six states. Each of these states  $s(\vec{x}) \in \{S1 \dots S6\}$  is defined as a combination of three binary variables  $[b(\vec{x}), O_b(\vec{x}), O_f(\vec{x})]$  as follows.

001 The pixel is not “missing” and there is occlusion in the forward direction only.

010 The pixel is not “missing” and there is occlusion in the backward direction only.

000 The pixel is not “missing” and there is no occlusion backward or forward.

100 The pixel is corrupted by a Blotch and there is no occlusion backward or forward.

101 The pixel is corrupted by a Blotch and there is occlusion in the forward direction only.

110 The pixel is corrupted by a Blotch and there is occlusion in the backward direction only.

Note that in this framework the [1 1 1] state is not allowed since it would imply that the data is missing and yet there is no temporal information for reconstruction. This is an interesting practical omission, and some comments are made at the end of the paper.

#### C. Bayesian Framework

From the degradation model of (6) the principal unknown quantities in frame  $n$  are  $I_n(\vec{x})$ ,  $s(\vec{x})$ ,  $c(\vec{x})$ , the motion  $\vec{d}_{n,n-1}$ , and the model error  $\sigma_e^2(\vec{x})$ . These variables are lumped together into a single vector  $\theta(\vec{x})$  at each pixel site  $\vec{x}$ . The Bayesian approach presented here infers these unknowns conditional upon the corrupted data intensities from the current and surrounding frames  $G_{n-1}(\vec{x})$ ,  $G_n(\vec{x})$  and  $G_{n+1}(\vec{x})$ . For the purposes of missing data treatment, it is assumed that corruption does not occur at the same location in consecutive frames, thus, in effect  $G_{n-1} = I_{n-1}$ ,  $G_{n+1} = I_{n+1}$ .

Proceeding in a Bayesian fashion, the conditional may be written in terms of a product of a likelihood and a prior as follows:

$$p(\theta | I_{n-1}, G_n, I_{n+1}) \propto p(G_n | \theta, I_{n-1}, I_{n+1}) \times p(\theta | I_{n-1}, I_{n+1}).$$

This posterior may be expanded at the single pixel scale, exploiting conditional independence in the model, to yield

$$\begin{aligned}
& p(\theta(\vec{x}) | G_n(\vec{x}), I_{n-1}, I_{n+1}, \theta(-\vec{x})) \\
& \propto p(G_n(\vec{x}) | \theta(\vec{x}), I_{n-1}, I_{n+1}) p(\theta(\vec{x}) | I_{n-1}, I_{n+1}, \theta(-\vec{x})) \\
& = p(G_n(\vec{x}) | I_n(\vec{x}), c(\vec{x}), b(\vec{x})) \\
& \quad \times p(I_n(\vec{x}) | \sigma_e(\vec{x})^2, \vec{d}(\vec{x}), O_b(\vec{x}), O_f(\vec{x}), I_{n-1}, I_{n+1}) \\
& \quad \times p(b(\vec{x}) | B) p(c(\vec{x}) | C) p(\vec{d}(\vec{x}) | D) p(\sigma_e(\vec{x})^2) \\
& p(O_b(\vec{x}) | O^b) p(O_f(\vec{x}) | O^f)
\end{aligned} \tag{7}$$

where  $\theta(-\vec{x})$  denotes the collection of  $\theta$  values in frame  $n$  with  $\theta(\vec{x})$  omitted and  $B, C, D, O^b, O^f$  and  $I$  denote local dependence neighborhoods around  $\vec{x}$  (in frame  $n$ ) for variables  $b, c, d, O_b, O_f$  and  $I_n$ , respectively. See the sections on prior distributions for details of these neighborhoods. To proceed, precise functional forms for the likelihoods must be assigned.

#### D. Corruption Likelihood

Considering (6),  $p(G_n | \cdot)$  has different forms depending on the state of  $b(\vec{x})$ , as follows (dropping the argument  $\vec{x}$  for clarity):

$$p(G_n | I_n, c, b) \propto \begin{cases} \exp - \left( \frac{(G_n - I_n)^2}{2\sigma_\mu^2} \right) & b = 0 \\ \exp - \left( \frac{(G_n - c)^2}{2\sigma_\mu^2} \right) & b = 1. \end{cases} \tag{8}$$

#### E. Original (Clean) Data Likelihood

This expression is derived directly from the image sequence model and also involves interaction with the occlusion variables

$$\begin{aligned}
& p(I_n(\vec{x}) | I_{n-1}, I_{n+1}) \\
& \propto \exp - \left( \frac{(1 - O_b(\vec{x}))(I_n(\vec{x}) - I_{n-1}(\vec{x} + \vec{d}_{n,n-1}))^2}{2\sigma_e^2} \right) \\
& \quad \times \exp - \left( \frac{(1 - O_f(\vec{x}))(I_n(\vec{x}) - I_{n+1}(\vec{x} + \vec{d}_{n,n+1}))^2}{2\sigma_e^2} \right) \\
& \quad \times \exp(-\alpha O_b(\vec{x})) \exp(-\alpha O_f(\vec{x})).
\end{aligned} \tag{9}$$

The likelihood is therefore proportional to the Displaced Frame Difference (DFD) =  $I_n(\vec{x}) - I_{n-1}(\vec{x} + \vec{d}_{n,n-1})$  between the image frames in both temporal directions. This expression encourages smoothness in intensities along a motion trajectory. However, when occlusion occurs (e.g.,  $O_b(\vec{x}) = 1$ ), the DFD is turned off because there is no valid data in the preceding (or next) frame.

This expression alone would yield the degenerate solution of  $O_b = O_f = 1.0$  everywhere in the image because that would maximize the impact of this likelihood on the posterior. Therefore, it becomes necessary to introduce a *penalty* for setting occlusion to the ON state. This is represented by  $\alpha O_b(\vec{x})$  where  $\alpha = 2.76^2/2.0$ . This value is chosen to represent *roughly* the 90% confidence limit for a Gaussian distribution. See [6] for further justification.

It is now necessary to specify priors for the variables  $\theta$ .

#### F. Priors

The remaining distributions encode the prior belief about the values of the various unknowns. The variance  $\sigma_e^2$  is assigned a noninformative prior  $p(\sigma_e^2) \propto 1/\sigma_e^2$ , following [70]. This anticipates “small” rather than “large” variance in the DFD. The other priors are discussed next.

1) *Motion*: The prior adopted for motion smoothness is a Gibbs Energy prior, for instance as introduced by Konrad and Dubois [71]. To reduce the complexity of the final solution the motion field is block based, with one motion vector being employed for each specified block in the image. The prior for  $\vec{d}_{n,n-1}(\vec{x})$ , the motion vector mapping the pixel at  $\vec{x}$  in frame  $n$  into frame  $n - 1$ , is as follows:

$$\begin{aligned}
& p_d(\vec{d}_{n,n-1}(\vec{x}) | \vec{d}_{n,n-1}(-\vec{x}), \mathbf{V}(\vec{x})) \\
& \propto \exp - \left( \sum_{\vec{x}_v \in \mathbf{V}(\vec{x})} \lambda(\vec{x}_v) [\vec{d}_{n,n-1}(\vec{x}) - \vec{d}(\vec{x}_v)]^2 \right)
\end{aligned} \tag{10}$$

where  $\vec{x}_v$  is one site in the 8-connected block neighborhood represented by  $\mathbf{V}(\vec{x})$ , and  $\lambda(\vec{x}_v)$  is the weight associated with each clique. The same prior is used for  $\vec{d}_{n,n+1}(\vec{x})$ . In order to discourage “smoothness” over too large a distance,  $\lambda(\vec{x}_v)$  is defined with  $\lambda(\vec{x}_v) = \Lambda / |\vec{X}(\vec{x}_v) - \vec{x}_B|$  where  $\vec{X}(\vec{x}_v)$  is the location of the block (in “block” units) providing the neighborhood vector  $\vec{d}(\vec{x}_v)$ ; and  $\vec{x}_B$  is the central block location.  $\Lambda = 2.0$  in the experiments presented later.

2) *Priors for Corruption, Detection and Occlusion*: Since Blotches tend to be “convex” clumps of degradation, the prior for  $b(\vec{x})$  should encourage contiguous areas of  $b = 1$  to form. In practice, blotches tend to have fairly constant intensity (see Figs. 1 and 14). If a texture exists, it is certainly smoother than that in the original image. Thus, the prior for  $c(\vec{x})$  should encourage smoothness of intensity, in much the same way that the prior for  $b(\cdot)$  encourages smoothness in its binary configuration. Therefore, it is reasonable to place a similar energy prior on both the binary field  $b(\vec{x})$  and the blotch value field  $c(\vec{x})$ . A Gaussian Markov Random Field (GMRF) prior is used for  $c(\vec{x})$ . The prior for  $b(\vec{x})$  is similar but operates on a binary variable only; thus, it is related to the Ising model.

The priors are therefore defined as follows:

$$\begin{aligned}
& p_c(c(\vec{x}) | C) \\
& \propto \exp \left( - \sum_{k=1}^8 \lambda_k^c (1 - u(\vec{x}, \vec{x} + \vec{v}_k)) (c(\vec{x}) - c(\vec{x} + \vec{v}_k))^2 \right)
\end{aligned} \tag{11}$$

$$\begin{aligned}
& p_b(b(\vec{x}) | B) \\
& \propto \exp \left( - \sum_{k=1}^8 \lambda_k^b (1 - u(\vec{x}, \vec{x} + \vec{v}_k)) |b(\vec{x}) - b(\vec{x} + \vec{v}_k)| \right)
\end{aligned} \tag{12}$$

where the eight offset vectors  $\vec{v}_k$  define the eight connected *pixel* neighborhood of  $\vec{x}$ , and  $C, B$  represent sets of these values from the respective fields (as previously). A typical set for  $\vec{v}_k$  is therefore  $[0, -1], [-1, -1], [0, -1], [1, -1], [1, 0], [1, 1], [0, 1], [-1, 0]$ .  $u(\vec{x}, \vec{x} + \vec{v}_k)$  is set to 1 if there is a significant zero crossing

between the location  $\vec{x}$  in the image and  $\vec{x} + \vec{v}_k$  from which a neighborhood pixel is extracted. Thus, the smoothness constraint is turned off across significant edges in the image. Note that these priors are defined on the *pixel* resolution image grid, whereas the motion prior discussed previously is defined on a *block* grid. In the results shown later  $u(\cdot)$  (the edge field) was configured using an edge detector employing Difference of Gaussians with the gradient threshold set at 5.0, the variance of the Gaussian filters was 1.0, 1.6 and the filter window sizes were  $11 \times 11$ .

For both these priors,  $\lambda_k^c, \lambda_k^b$  are assigned values such that  $\lambda_k^c = \Lambda^c / |\vec{v}_k|$ ,  $\lambda_k^b = \Lambda^b / |\vec{v}_k|$ . This makes the hyperparameter weighting circularly symmetric. Typical values are  $\Lambda_c = 0.15$ ,  $\Lambda_b = 4.0$ . These are found from experiment to yield good results over a wide cross section of degraded sequences.

The occlusion priors  $p_o(O_b | \mathbf{O}_b), p_o(O_f | \mathbf{O}_f)$  are identical to the prior for the Blotch indicators  $b$ , with  $\Lambda^o = 2.0$ . It encourages organization (clumpiness) in the occlusion surfaces. This is sensible given that the moving objects tend to cause contiguous regions of occlusion and uncovering.

### G. Solving for the Unknowns

The solution is generated by manipulating  $p(\theta | I_{n-1}, G_n, I_{n+1})$ . For instance the MAP estimate is generated by maximizing the distribution with respect to the unknowns. Unfortunately, due to the nonlinear nature of the expression, a closed form solution to the optimization problem is not available. To yield a practical solution, a number of simplifying manipulations will be made.

It is expedient at this stage to note that in fact the variables of this problem can be grouped into a pixel state,  $s(\vec{x}) = [b(\vec{x}), O_b(\vec{x}), O_f(\vec{x})]$  (defined previously), and its associated "clean" image value  $I_n(\vec{x})$ , at each pixel site, and a motion and error variable  $[\vec{d}_{n,n-1}, \vec{d}_{n,n+1}, \sigma_e^2]$  for each block in the image. In other words, the variables can be grouped into a *pixelwise* group  $[s, I_n, c]$  (including the corruption value  $c$ ), that varies on a pixel grid, and a *blockwise* group  $[\vec{d}_{n,n-1}, \vec{d}_{n,n+1}, \sigma_e^2]$ , that varies on a coarser grid.

To solve for these unknowns, the ICM algorithm [72] is used. At each site, the variable set that maximizes the *local* conditional distribution *given* the state of the variables around, is chosen as a suboptimal estimate. Each site is visited with variables being replaced with ICM estimates. By treating each group of variables *jointly* an efficient solution results. Thus, first the group  $[s, I_n, c]$  is estimated *given* the current estimate for motion, then the blockwise motion information is estimated given the current reconstructed image estimate  $I$ . This process is then iterated.

1) *Factoring*: It transpires that  $s, I_n(\vec{x})$  and  $\sigma_e^2, \vec{d}$  can be *jointly* manipulated by factoring the posterior. To illustrate, the p.d.f. for  $s, I_n(\vec{x})$  can be factored by the decomposition

$$\begin{aligned} p(s, I(\vec{x}) | I_{n-1}, I_{n+1}, \mathbf{d}, \sigma_e^2) \\ = p(I(\vec{x}) | s, I_{n-1}, I_{n+1}, \mathbf{d}, \sigma_e^2) p(s | I_{n-1}, I_{n+1}, \mathbf{d}, \sigma_e^2). \end{aligned} \quad (13)$$

The algorithm therefore proceeds by first solving for  $s$  with  $p(s | I_{n-1}, I_{n+1}, \mathbf{d}, \sigma_e^2)$  then using that estimate in the gener-

ation of the clean image data. The various factor terms can be derived by successively *integrating out*  $I_n(\vec{x})$  and  $c(\vec{x})$  from the posterior distribution. See Appendix A for details. Thus, the state  $s$  that maximizes  $p(s | I_{n-1}, I_{n+1}, \mathbf{d}, \sigma_e^2)$  is chosen first then that value is used in estimating  $I(\vec{x})$  from  $p(I(\vec{x}) | s, I_{n-1}, I_{n+1}, \mathbf{d}, \sigma_e^2)$ .

Although suboptimal, the approximation is helped by the observation that there is no *spatial* dependence in the conditional for the clean image data  $p(I_n(\vec{x}) | \cdot)$  and the conditional  $p(\sigma_e^2(\vec{x}) | \cdot)$ . Maximizing the local conditional distribution for  $s$  can be performed very efficiently by evaluating the posterior for the six possible state options and simply choosing that which maximizes the posterior. The derivation of the state conditionals is tedious and is left to the Appendix.

2) *Manipulating Motion*: Recall that in this implementation, motion is handled on a block basis. Integration of the posterior yields the following factorization of  $p(\sigma_e^2, \mathbf{d} | \cdot)$ , using the backward motion  $\mathbf{d}_{n,n-1}$  and backward frame pair

$$\begin{aligned} p(\sigma_e^2 | \mathbf{d}_{n,n-1}, \mathbf{i}) &= \text{IG}(N/2, E(\mathbf{i}, \mathbf{d}_{n,n-1})/2) \\ p(\mathbf{d}_{n,n-1} | \mathbf{i}, D) &\propto E(\mathbf{i}, \mathbf{d}_{n,n-1})^{-N/2} p_d(\mathbf{d}_{n,n-1} | D) \end{aligned} \quad (14)$$

where IG denotes the Inverted Gamma distribution,  $N$  is the number of pixels in the image block,  $D$  (shorthand for  $\mathbf{S}_n(\vec{x})$ ) represents a neighborhood of vectors surrounding the current location, and  $E(\mathbf{i}, \mathbf{d}_{n,n-1})$  is the sum of the square of the DFD's in a block (allowing for occlusion), given the vector  $\mathbf{d}$ . The expressions for the forward frame pair are similar except for the use of  $\mathbf{d}_{n,n+1}$  instead. Note that the distinction between forward and backward temporal directions means that the forward and backward DFD variance is allowed to vary. The derivation of these expressions can be found in [6]. See Appendix B for a concise summary.

Again the local conditional is maximized in parts; first  $\mathbf{d}$  from  $p(\mathbf{d} | \mathbf{i}, D)$ , then  $\sigma_e^2$  from maximizing the IG distribution above. The maximum of the IG distribution can be calculated analytically and is simply the usual estimate for variance given  $\mathbf{d}$  as follows:

$$\hat{\sigma}_e^2 = \frac{1}{N} \sum_{\vec{x} \in B} [\text{DFD}(\vec{x}) - \overline{\text{DFD}(\vec{x})}]^2 \quad (15)$$

where  $\overline{\text{DFD}(\vec{x})}$  is the mean of the DFD in that block of pixels.

The maximization of the motion conditional has no straightforward analytic solution. However, a simple and practical solution is to use fast deterministic motion estimation schemes to yield candidates for motion, for example block matching or optic flow schemes. The marginal conditional probability of these candidates are then evaluated using (14). This idea has not been fully explored in the literature but does yield much practical benefit.

3) *Final Algorithm*: The algorithm begins with the initialization of the fields using deterministic estimates. Motion is initialized using some standard block based motion estimator. The detection field  $b$  is initialized using any one of the deterministic schemes, e.g., SDIp, ROD, or Morphological Operations [31]. A conservative initial estimate for  $c(\vec{x})$  is  $G_n(\vec{x})$ . It then proceeds on the basis of two major iterations as follows.



- 1) Given the current estimates for  $\vec{d}, \sigma_e^2$ , the image is swept on a pixel basis, using a checkerboard scan, generating an estimate for  $s$  and  $I(\vec{x})$  at each site.
- 2) Given the current estimates for  $s$  and  $I(\vec{x})$ , the image is scanned on a block basis. Within each block, the solution for motion is generated by selecting eight candidate motion vectors from surrounding spatial blocks and choosing the one which minimizes the following “energy” [arising from (14)]

$$\frac{N}{2} \log_e \left[ \sum_{\vec{x} \in B} (I_n(\vec{x}) - I_{n-1}(\vec{x} + \vec{d})(1 - O_b(\vec{x})))^2 \right] + \log_e(p_d(\vec{d}|\cdot)) \quad (16)$$

$\sigma_e^2$  is then measured as the variance of the prediction error in the block after the motion is selected. Note that as the occlusion variables are taken into account (removing pixel pairs that are occluded), the normalizing term ( $N/2$ ) changes accordingly. In cases where there is no valid temporal data,  $\log_e(p_d(\vec{d}|\cdot))$  dominates and the smoothest motion field will be selected regardless of DFD.

These two processes are iterated until reasonable pictures are built. It is found that typically no more than ten such iterations are needed, for good pictures to be built. The algorithm for the selection of the best state  $s$  at each pixel is as follows.

4) *Maximizing With Respect to S*: At each pixel site there are just six possibilities for states. Each of these states defines a value for all the pixelwise variables under consideration. By integrating out  $c$  and then  $I_n(\vec{x})$  from the posterior (see [6]), the maximization process manifests as the minimization of six possible energies (log likelihoods) depending on the values of state. Four of these are shown below. Details are in Appendix A. Note that  $g_n, i_n$  is shorthand for a single pixel observation of the degraded and clean image respectively, i.e.,  $G_n(\vec{x}), i_n$ . In the expressions that follow, it is assumed that any data required from frames  $n - 1, n + 1$  are compensated for motion.

$$\begin{aligned} \mathcal{E}(S1) &= \log_e [2\pi\sigma_e^2\sigma_\mu^2] + \alpha + \left[ \frac{(g_n - \hat{i}_b)^2}{2\sigma_\mu^2} \right] \\ &+ \left[ \frac{(i_{n-1} - \hat{i}_b)^2}{2\sigma_e^2} \right] \\ &+ \log_e [p_b(b=0|B)p_o(O_b=0|\mathbf{O}_b)] \\ &\times p_o(O_f=1|\mathbf{O}_f) \\ &+ \log_e [p_c(c=\hat{c}_0|C)] \\ \mathcal{E}(S3) &= \frac{1}{2} \log_e [8\pi^3\sigma_e^4\sigma_\mu^2] + \left[ \frac{(g_n - \hat{i})^2}{2\sigma_\mu^2} \right] \\ &+ \left[ \frac{(i_{n-1} - \hat{i})^2}{2\sigma_e^2} \right] \\ &+ \left[ \frac{(i_{n+1} - \hat{i})^2}{2\sigma_e^2} \right] \\ &+ \log_e [p_b(b=0|B)p_o(O_b=0|\mathbf{O}_b)] \\ &\times p_o(O_f=0|\mathbf{O}_f) \\ &+ \log_e [p_c(c=\hat{c}_0|C)] \end{aligned}$$

$$\begin{aligned} \mathcal{E}(S4) &= \frac{1}{2} \log_e [2\pi\sigma_\mu^2] + \left[ \frac{(g_n - \hat{c})^2}{2\sigma_\mu^2} \right] \\ &+ \log_e [p_b(b=1|B)p_o(O_b=0|\mathbf{O}_b)] \\ &\times p_o(O_f=0|\mathbf{O}_f) \\ &+ \log_e [p_c(c=\hat{c}_1|C)] \\ \mathcal{E}(S6) &= \frac{1}{2} \log_e [2\pi\sigma_\mu^2] + \alpha + \left[ \frac{(g_n - \hat{c})^2}{2\sigma_\mu^2} \right] \\ &+ \log_e [p_b(b=1|B)p_o(O_b=1|\mathbf{O}_b)] \\ &\times p_o(O_f=0|\mathbf{O}_f) \\ &+ \log_e [p_c(c=\hat{c}_1|C)] \end{aligned}$$

where  $i_{n-1}$  denotes a motion compensated pixel in frame  $n - 1$ . The various constants  $\hat{i}_b, \hat{i}_f, \hat{i}, \hat{c}$  are the least squares estimates of the unknowns  $I_n(\vec{x}), c(\vec{x})$  given various temporal situations defined by the state variable. For instance

$$\begin{aligned} \hat{i}_b &= (\sigma_e^2 g_n + \sigma_\mu^2 i_{n-1}) / (\sigma_e^2 + \sigma_\mu^2) \\ \hat{i} &= (\sigma_e^2 g_n + \sigma_\mu^2 (i_{n-1} + i_{n+1})) / (\sigma_e^2 + 2\sigma_\mu^2). \end{aligned}$$

The Energies  $\mathcal{E}(S2, S5)$  are the same as  $\mathcal{E}(S1, S6)$  except calculated using the forward motion compensated frame direction; see (28).

5) *Estimating  $i_n, c$* : Once the state configuration is established at a pixel site, estimates for  $I(\vec{x}), c(\vec{x})$  can be generated. This is done directly from (18) and (21). If  $b(\vec{x}) = 1$  then  $I(\vec{x})$  is interpolated from the previous and/or next frames (depending on  $s(\vec{x})$ , and  $c$  is set to a noise reduced version of the observed image. When  $b(\vec{x}) = 0$ , the image estimate is a temporal average (two or three frames) depending on occlusion, and  $c(\vec{x})$  is interpolated spatially. Further computational savings can be had by exploiting the fact that all the energies involving  $g_n, I_n$  can be pre-computed.

## H. Results

It is difficult to compare this system with previous work in noise reduction or blotch removal since this system treats the two problems jointly. Therefore it makes compromises that sometimes give worse performance in one of the two domains, but overall the output pictures show good improvement. In some attempt to place this algorithm in context, a  $256 \times 256$  subsection of the Mobile and Calendar sequence was corrupted with Blotches that follow the prior for  $c(\vec{x})$  and Gaussian noise of  $\sigma_\mu^2 = 100$  was added. This is in keeping with the degradation model discussed here. A number of experiments were performed to evaluate the behavior as a blotch detector and as a noise reducer. The algorithm introduced here is called *Joint Noise Reduction, Detection and Interpolation*: JONDI.

1) *Blotch Detection Performance*: As discussed previously, many of the currently available simple/fast Blotch detection processes are based on the model in (1), and all employ temporal motion compensated frame differences, e.g., SDIa (Spike Detection Index a), SDIp, ROD (Rank Order Detector). JONDI is using precisely, the same information, but in a unique way.

Fig. 11 shows a Receiver Operating Characteristic (ROC) that compares the performance of JONDI with several detectors with respect to their blotch detection performance. Good performance is indicated by curves that are close to the top left

corner of the graph. To create the characteristics the processes were run with a range of parameter settings. In the case of SDIp,  $T = 5 : 5 : 55$  (Matlab notation) and the performance degrades as the threshold increases. For ROD  $t_1$  was varied  $5 : 5 : 45$ , with the other thresholds held at nominal values of  $t_2 = 39.0$  and  $t_3 = 55.0$ .

The situation is more complicated with JONDI, Morris [33], [65] and JOMBANDI [6], [41], [73] since there are two parameters to be set in each case. However, from top right to bottom left, the points on the JONDI curve shown correspond to the following values for  $(\Lambda^c, \Lambda^b)$  :  $(0.001, 1.0)$ ,  $(0.1, 1.0)$ ,  $(0.1, 4.0)$ ,  $(0.1, 8.0)$ ,  $(.15, 15)$ . For JOMBANDI (see [73], [41]) the parameters (having the same meaning) were  $(0.15, 1.0)$ ,  $(0.1, 1.0)$ ,  $(0.15, 4.0)$ ,  $(0.1, 4.0)$ . The lower curve (dashed) corresponds to these parameters with a 1-tap AR model (identical to the (1) except with a gain parameter in the prediction), while the upper curve (solid) corresponds to a 5-tap 3DAR process. Ten iterations were used for JONDI, JOMBANDI, and Morris on each frame.

The Morris detector has two parameters,  $\Lambda_o$  for Occlusion smoothness (as with JONDI), and  $\alpha$ , the Occlusion penalty (again the same as JONDI). The points on the ROC for this detector from top right to bottom left correspond to settings  $(\Lambda_o, \alpha) = (4, 1)$ ,  $(1, 5)$ ,  $(4, 5)$ ,  $(4, 8)$ ,  $(4, 10)$ ,  $(8, 10)$ ,  $(8, 20)$ .

A multiresolution gradient based block motion estimator was used to generate candidates for motion [6]. A block size of  $9 \times 9$  pixels was employed with a 2-pixel overlap between blocks. The SDIp detector was used to initialise the  $b(\vec{x})$  field for JONDI, using a threshold of 25 grey levels. The algorithm was run recursively, i.e., restored frames were used in the processing of the following frames.

The correct detection rate is measured as the fraction of pixels (out of the total number of missing sites) correctly set to 1 in  $b(\vec{x})$ . The false alarm rate is measured as the fraction of pixels incorrectly flagged as missing out of all the possible uncorrupted sites. Fig. 11 shows clearly that JONDI outperforms SDIp, ROD, Morris, JOMBANDI across a wide range of parameters. It is interesting to note that ROD and JOMBANDI (with a 1 tap AR process) perform almost the same at high false alarm rates. The Morris detector and JONDI have the same basic image sequence model in common, yet JONDI is able to continue giving reasonable performance at low false alarm rates ( $<.001$ ) while the Morris performance drops off sharply. This is because JONDI also incorporates priors on the blotches, and also seeks to improve the motion estimate as its iterations proceed. The 5-tap JOMBANDI gets close to the JONDI performance but does not reach it. JONDI is therefore gaining much improvement from the incorporation of the Occlusion fields.

Fig. 13 shows the result of JONDI on three frames from the corrupted sequence. The bottom row shows  $(I(\vec{x}) - G(\vec{x})) + 128$  and so illustrates more clearly what has been removed from the dirty image. The combined blotch rejection and noise reduction features are clear. Also important is that the rotating ball is not damaged. Note as well that the corruption level in the test sequence is very high, and in fact corruption at the same site in consecutive frames does occur.

2) *Noise Reduction Performance:* Fig. 12 shows the dB improvement in SNR after processing with JONDI, the Temporal

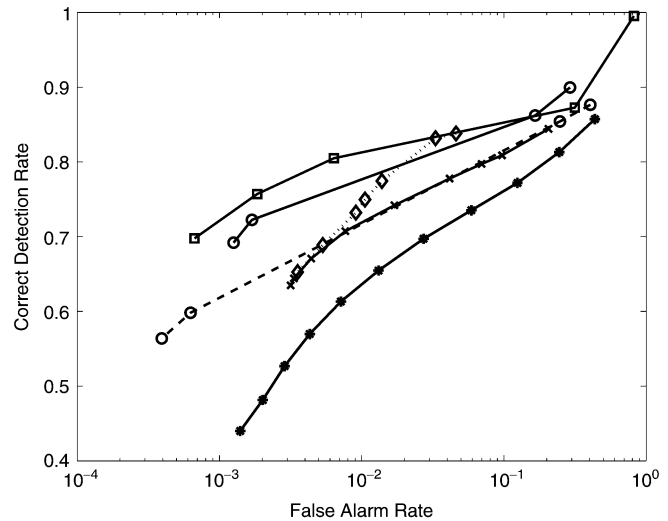


Fig. 11. Performance of JONDI ( $\square$ ), JOMBANDI ( $- \circ -$  1 tap,  $- \circ -$  5 tap), Morris ( $\diamond$ ), ROD ( $\times$ ), SDIp ( $*$ ) at  $\sigma_\mu^2 = 100$ .

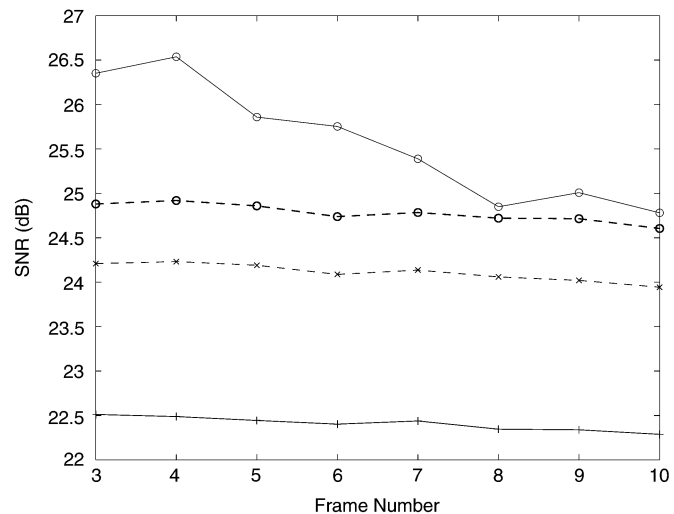


Fig. 12. Noise reduction performance for  $\sigma_\mu^2 = 100$ . JONDI ( $\circ -$ ), Temporal Wiener ( $\times - -$ ) and Recursive Filtering ( $\circ - -$ ). Algorithm settings are  $\Lambda^c = 0.15$ ,  $\Lambda^b = 4.0$  ( $\circ - -$ ). Noisy sequence is represented by the bottom line.

Wiener Filter [74] (discussed above) and Temporal Recursive Frame Averaging [75]. To separate out the noise reduction component of JONDI from the missing data treatment component, the measurement of SNR was made only in those regions not corrupted by missing data. This does not however, totally separate the two components since Blotches can have an effect on processing for some distance *outside* their area.

The lowest curve shows the SNR of the degraded sequence at about 22 dB, and the top curve shows that JONDI at  $\Lambda^c = 0.15$ ,  $\Lambda^b = 4.0$  performs best. Changing  $\Lambda^b$  to 1.0 makes JONDI perform somewhere between the two temporal filters as far as noise reduction goes. This is sensible since a reduction in  $\Lambda^b$  implies that it is expected that Blotches are less “convex,” which is not the case. Note that as a noise reducer, JONDI acts as a kind of automatically controlled recursive process. It should therefore be similar in performance to [75]. The main difference for noise reduction is that JONDI incorporates a motion refinement step. Thus initially, when there is

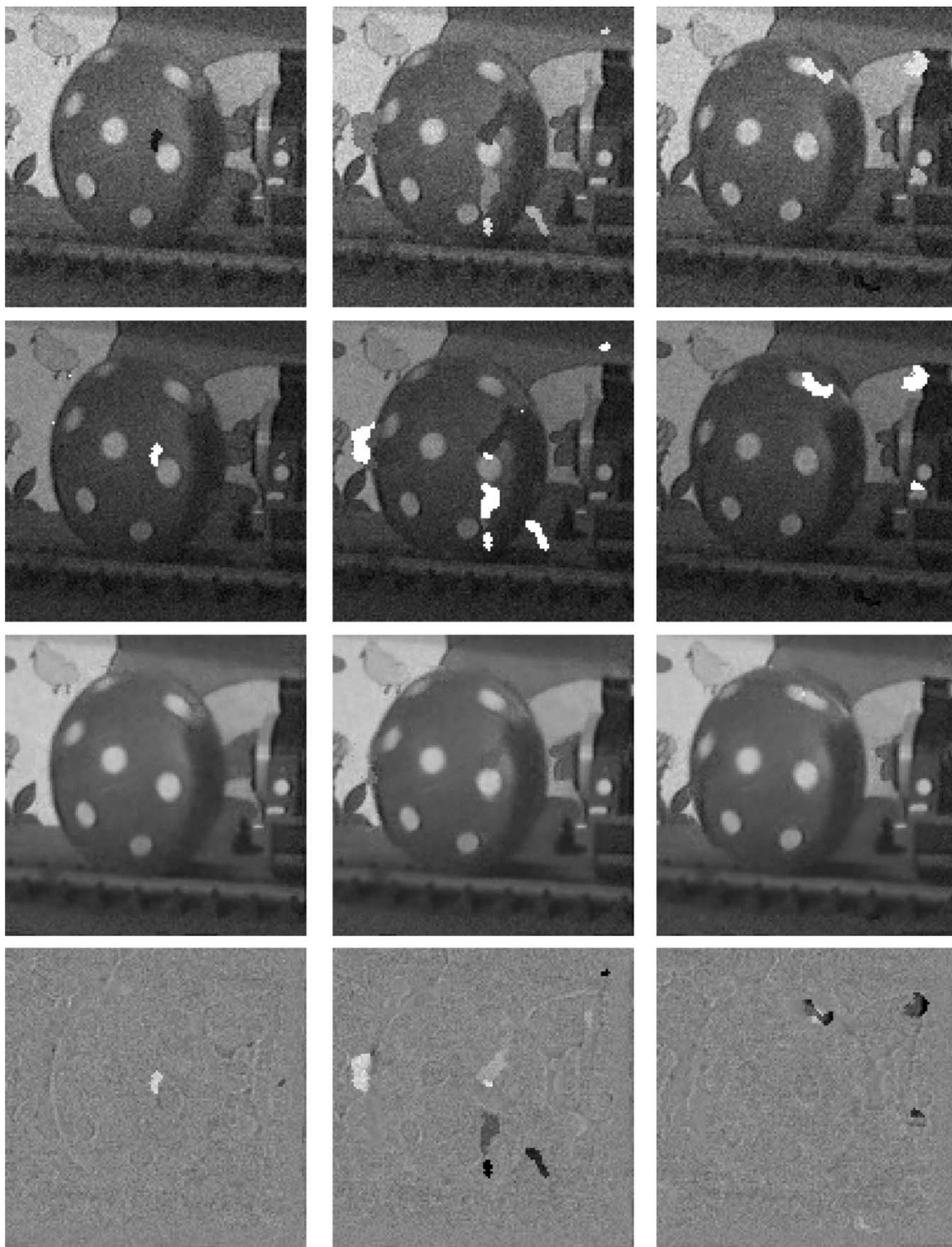


Fig. 13. Section of Mobile and Calendar sequence frames 6, 7, 8. Top: Corrupted with blotches and  $\sigma_{\mu}^2 = 100$ . Second Row: Blotch detection ( $b(\vec{x})$ ) superimposed in bright white. Third Row: Algorithm result  $\Lambda^c = 0.15$ ,  $\Lambda^b = 4.0$ . Bottom: Difference between images in top and third row offset by 128.

little pathological motion in the sequence, JONDI does well. However as the mobile in the scene starts to move rapidly, even the motion processes in JONDI begin to fail and performance degrades to the normal recursive noise reducer level.

3) *Real Pictures*: The top row of Fig. 14 shows a zoom on part ( $300 \times 300$  pixels) of a typical real degraded sequence

showing damage with both missing data and noise. The motion is roughly upwards. The damage in this case is caused by a *tear* at two parts in the frame. The next row shows a corresponding restoration for the last two frames with the algorithm described in this paper. The restoration is good, and the noise reduction is effective despite the fact that there is no spatial



Fig. 14. Top row: Degraded frames 1, 2, 3 of real sequence. Second Row: Frame 2,3 restored with algorithm described in this paper. Third Row:  $b(\vec{x})$  for frames 2, 3 superimposed on darkened original frames. Bottom Row: Initial motion estimate (left) superimposed on  $c(\vec{x})$  for frame 2; Motion field after 5 iterations of algorithm together with  $O_b$  ('o'),  $O_f$  ('x').

constraint on the image data. Five iterations were used, and  $\Lambda^b = 4.0, \Lambda^o = 2.0; \Lambda^c = 0.15$ . The same multiscale gradient based motion estimator [6] as was used for the artificial case above, was employed to initialize the motion field and the block size was  $17 \times 17$  pixels.  $\sigma_\mu = 20.0$  for these results. This is a user defined parameter since noise reduction is highly subjective. The next rows show the final configured detection, corruption, motion, and occlusion fields.

The fourth row shows that the motion is initially severely affected by the tear (in the region of the tear), but the algorithm corrects this as shown on the right. The estimated corruption field is seen to be roughly equal to the degraded image where a blotch is detected, and is very "flat" otherwise. This is sensible given the GMRF used as a prior. The occlusion fields are

seen to be able to prevent distortion in one frame from affecting the next, since the forward occlusion field is correctly set in frame 2 in the region that is damaged by the tear in frame 3. Previous work did not account for occlusion; and distortion therefore tended to "leak" between frames. Note however, that the occlusion indicators are only linked to intensity and not to motion. This implies that they do not infer occlusion correctly, but only to the extent that it helps the restoration.

What is interesting is that even though the estimated pictures (second row) show a definite reduction in artefacts, especially noise, not all the damaged pixels *appear* to have been detected. See the top right corner of Fig. 3. This is because the system is accounting for some of the low contrast missing data through the noise process and *not* through the blotch corruption process.



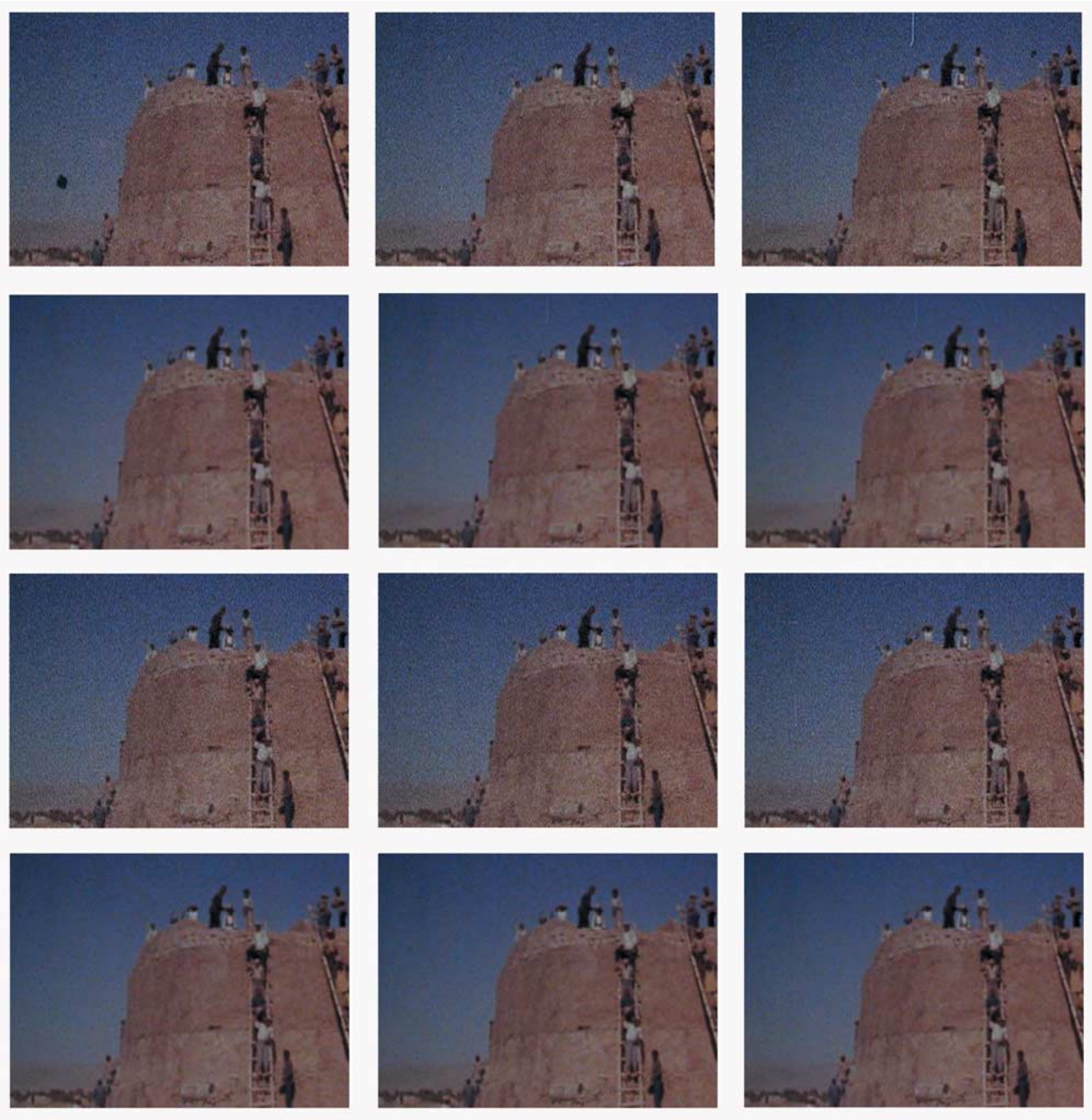


Fig. 15. Color restoration. Top two rows show, respectively, three frames from a real archived sequence, and the corresponding restoration. Bottom two rows show more images from the same sequence.

This kind of interaction is interesting and is under current investigation. The same observation could be made about the pictures in Fig. 13.

4) *Color*: Fig. 15 shows a series of images from a color sequence that have been processed with JONDI. The YUV color space was used. The motion manipulation, blotch detection and noise reduction is performed only on the image Luminance component, while the blotch reconstruction process was performed on all three color planes. Noise reduction and blotch detection is restricted to Luminance only because most of the signal energy lies in that component. It is true that

sometimes blotches can occur in the color planes only (especially in the case of digital dropout) but this can be dealt with by performing the detection across all three planes. In this case limiting detection and noise reduction to the luminance component, adds the benefit of keeping computation low while still doing an effective job across the majority of degraded material. The identical parameter settings as for the previous example were used here. The reduction in noise in particular is remarkable. Note also that a large dark blotch on the left of the first image is correctly removed. There is generally a lower level of corruption by Blotches in real pictures

as compared with the Mobile/Calendar experiment discussed above.

### I. Relationships

This new pixelwise relaxation approach that incorporates occlusion, shows some interesting links with previous works. These are considered next in chronological order.

Morris [33], [34], [65] considered that detection of blotches could be articulated by the detection of sites at which both forward and backward discontinuities occur. In the framework just presented, the Morris problem is then one of detecting when  $O_f = 1$  and  $O_b = 1$ . To allow this in JONDI, the pixel states would have to be altered to consider *only* occlusion, such that each pixel state  $s' = [O_b, O_f]$  instead of  $[b, O_b, O_f]$ .  $b$  is then no longer a variable of interest. The clean data likelihood in (9) can then be changed such that  $I_n = G_n, I_{n-1} = G_{n-1}, I_{n+1} = G_{n+1}$ . Then, the problem would be to estimate  $s'$  by maximizing  $p(S' | G_{n-1}, G_n, G_{n+1})$ . The motion, detection, and corruption priors are then superfluous, since  $b, I$  are no longer variables of interest, and the same algorithm results. In essence, the Morris detector is detecting temporal discontinuities based on Displaced Frame Difference just like the SDI detectors; but also including a prior to encourage spatial smoothness in the discontinuity output. This process relies heavily on reasonable motion estimation, and is poor at motion edges. In  $\approx 1997$ , [76] improved upon this idea by suppressing the flagging of occlusion at motion edges. A pre-process was required to configure a rough guess motion edge field, and this was used to configure the prior.

Kokaram [6], [34], [41] *et al.* (1995–2002) proposed that the image sequence can be locally modeled by a 3-D spatiotemporal Autoregressive process, or 3-D linear predictor. This replaces the pixel difference image model in (1) with a linear prediction error based on a spatio-temporal volume. There was no direction associated with this prediction, and this causes difficulties in occluded areas. Deleting the occlusion variable from JONDI, incorporating a linear prediction likelihood, and relinquishing the notion of a pixel state that is other than just luminance, yields JOMBANDI as presented in [41].

Woi Boon *et al.* solved the 3DAR direction problem in a deterministic framework by choosing the direction that gave the lowest prediction error. Roosmalen [35], [36] incorporated a direction indicator into a 3DAR reconstruction framework. This is identical to the use of occlusion in the JONDI framework, if the likelihood is altered, and it is assumed that the detection field  $b(\vec{x})$ , and motion was given.

### J. Why Can JONDI Work?

The unique aspect about JONDI, aside from the obvious joint treatment of several variables, is that it is connecting the picture building process with the discontinuity detection and motion estimation process. The energies that are being used to select pixel states are trading off the luminance deviation from a linear estimate of the current pixel,  $\hat{i}, \hat{i}_b, \hat{i}_f$ , against the smoothness in the local variable field. In other words, if the local luminance is very far away from the luminance in nearby frames, and this behavior also occurs in nearby sites, then there must be some

problem. Viewed in this light, detection using the ROD, and filtering approaches in general, can be seen to be providing an alternative estimate for the current pixel. The ROD is using the idea that a rank-order filter on the local region is a reasonable estimate of the local true pixel intensity. Hence, deviations from that could flag a problem. It would be interesting to generalize this concept within JONDI, by expressing the likelihood in a more general nonlinear fashion, but that is left for future work.

Furthermore, what JONDI explicitly makes clear, is that the need for fully volumetric processing in image sequences is unnecessary **if** motion can be handled properly. In other words, most of the time, each frame in the sequence can be synthesized by rearranging frames in the past or in the future. Therefore, if that rearrangement could be estimated in some robust way, a cut and paste operation is sufficient to synthesise picture material. The need for a 3-D Autoregressive process (for instance) is therefore restricted to cases where the sequence behavior is of a higher order, for instance a *video texture* like flowing water or light reflecting off the ocean. In general perhaps, the need for spatiotemporal reconstruction is motivated more by an attempt to recover from errors in motion estimation itself rather than sequence modeling.

Finally, it is interesting to note that the estimates for  $\hat{i}$  in JONDI are all the optimal Wiener estimates given Gaussian noise. In fact  $\hat{i}$  is very similar indeed to the temporal noise reduction filter first presented by Katsagellos *et al.* [74] in 1989.

### K. Computational Comments

The main advantage of this algorithm is that by adopting simple temporal models, the computation for the state estimation is very low per pixel site. Large parts of the energy expressions can be pre-computed. This is in contrast to previous work that incorporated 3DAR processes [6], that did give better robustness to small motion errors, but at considerably more computational cost. In that process, the solution for the AR parameters alone required the inversion of at least a  $5 \times 5$  matrix, for *every* block of pixels. In JONDI, there is no such estimation step. Without the need for estimating the parameters of a stochastic process, the motion estimation load becomes dominant in JONDI and accounts for 80% of the time spent on one iteration. A single iteration (for both state and motion estimation) takes about 15 s per  $720 \times 576$  frame on a PIII 300 MHz PC (using nonoptimized code) and about 2 s of this is due to the state estimation step. The surprising aspect about JONDI, as far as the ROC tests are concerned however, is that it is able to perform better than other Bayesian detectors at a reasonable computational cost.

## V. FAILURE IN THE MOTION MODEL

Much of work reviewed and presented thus far considers that it is possible to write an image sequence model of the kind in (1). More correctly, it is expected that the use of the Gaussian error term  $e(\vec{x})$  is good enough to account for deviations from the model. This is hardly always the case. Fast motion of objects causes blurring and many interesting objects, e.g., clothing, are not rigid. This means that in some parts of any sequence it will be impossible to model the behavior. In this situation, most of



the current missing data techniques fail. Damage to fast-moving material is typically annoying to most viewers and no processing in those regions is often preferred.

In recent efforts, Bornard [63], [64] and Rares *et al.* [77]–[80] considered these problems from the practical standpoint of processing hours as opposed to seconds of video material. They have acted upon a well-known observation by the users of restoration tools. Whenever motion estimators fail, they tend to cause a high degree of false alarms in roughly the same location in consecutive frames. Recalling that blotches should not occur in the same place in consecutive frames, implies that this behavior is indicating a problem with the image material. Detection of this phenomenon allows the blotch treatment process to be turned off before any damage can be done. By building a five-frame version of the Morris [65] detector, Bornard [63] is able to detect areas where discontinuities are being flagged regularly over the five-frame aperture. Although the process is conservative in what it treats, initial results from Bornard’s work [64] are very encouraging indeed. A classification approach to this problem has been taken by Rares *et al.* [77], [80]. Rares presents a taxonomy of pathological motion in [77] and attempts to classify the regions showing discontinuities as pathological or not. This is a more long term approach as there are implications for video processing in general.

## VI. FINAL COMMENTS AND POSSIBLE FUTURES

Missing data treatment in video and film archives has matured to the point that simple algorithms are currently appearing in industry standard hardware. Although the more complex Bayesian approaches have yet to be implemented in such systems, recent developments have shown that low cost Bayesian inference is possible.

Coping with pathological motion remains an issue. While Rares and Bornard have set the groundwork for considering this issue, the integration of “model failure” into restoration tasks in general is still an open question. It may be thought that allowing the state  $O_f = 1, O_b = 1$  in the Bayesian framework presented here, could be one solution to the integration issue. However this is not clear since that Occlusion state and the state  $b = 1$  would then be heavily correlated, at least in the context of the discussion given in this paper. Further work in this direction is currently being undertaken.

The reader may also notice that most of the work in missing data for degraded video and film has not considered the use of color information. The examples in Figs. 1 and 2 illustrate that in color sequences, black and white dirt against colored backgrounds is common. The reason for ignoring this powerful information is simply that the search is for *generic* tools that would operate equally well for Black and White movies as well as different types of missing data problems, e.g., emulsion or water damage. In the latter case, color content is not necessarily as heavily contrasted. Nevertheless, it should be interesting to consider how color information could best be incorporated when it is advantageous.

It is important to note that not only archived footage suffers from Dirt for instance. Recent footage always contains some level of defects, and the availability of relatively low cost software tools for assisting the clean up process is increasingly in demand. As restoration systems continue to be deployed, either for real time video broadcast [5], or film scanning and post production [81] for digital cinema or DVD, users will increasingly have more say in the requirements for good systems. That will ensure the need for further research and development in this area. It is encouraging that this area of signal and image processing, which for so long had been on the periphery of mainstream applications, is slowly but surely growing in prominence.

## APPENDIX I

### FACTORING THE CONDITIONAL FOR PIXEL STATES, IMAGE AND CORRUPTION DATA

To simplify the algebra, express the priors for  $s = [b, O_f, O_b]$  as  $p_s(s | S)$ , defined as follows:

$$\begin{aligned} p_s(s = [b, o_b, o_f] | S = [B, O_b, O_f]) \\ = p_b(b | B)p_o(O_f | \mathbf{O}_f)p_o(O_b | \mathbf{O}_b) \end{aligned} \quad (17)$$

where  $s = [b, O_f, O_b]$  refers to the state at a single pixel site; and  $S = [B, \mathbf{O}_b, \mathbf{O}_f]$  refers to the states in the local eight connected neighborhood in the same frame.  $i_n$  refers to a single *clean* pixel value at that same site.

In the remainder of the Appendix, it is assumed that all image pixel data required in frames  $n - 1$  and  $n + 1$  has *already* been compensated for motion. The expression for the conditional for  $s, i_n, c$  GIVEN motion and noise variance is then (see (18) at the bottom of the page).

$$p(s, i_n, c | \mathbf{I}, S, \mathbf{C})$$

$$\propto \begin{cases} \exp - \left( \frac{(g_n - i_n)^2}{2\sigma_\mu^2} \right) \exp - \left( \frac{(i_n - i_{n-1})^2}{2\sigma_e^2} \right) p_s(s = [0, 0, 1] | S) p_c(c | C) \exp - \alpha & s = [0, 0, 1] \\ \exp - \left( \frac{(g_n - i_n)^2}{2\sigma_\mu^2} \right) \exp - \left( \frac{(i_n - i_{n+1})^2}{2\sigma_e^2} \right) p_s(s = [0, 1, 0] | S) p_c(c | C) \exp - \alpha & s = [0, 1, 0] \\ \exp - \left( \frac{(g_n - i_n)^2}{2\sigma_\mu^2} \right) \exp - \left( \frac{(i_n - i_{n-1})^2 + (i_n - i_{n+1})^2}{2\sigma_e^2} \right) p_s(s = [0, 0, 0] | S) p_c(c | C) & s = [0, 0, 0] \\ \exp - \left( \frac{(g_n - c)^2}{2\sigma_\mu^2} \right) \exp - \left( \frac{(i_n - i_{n-1})^2 + (i_n - i_{n+1})^2}{2\sigma_e^2} \right) p_s(s = [1, 0, 0] | S) p_c(c | C) & s = [1, 0, 0] \\ \exp - \left( \frac{(g_n - c)^2}{2\sigma_\mu^2} \right) \exp - \left( \frac{(i_n - i_{n-1})^2}{2\sigma_e^2} \right) p_s(s = [1, 0, 1] | S) p_c(c | C) \exp - \alpha & s = [1, 0, 1] \\ \exp - \left( \frac{(g_n - c)^2}{2\sigma_\mu^2} \right) \exp - \left( \frac{(i_n - i_{n+1})^2}{2\sigma_e^2} \right) p_s(s = [1, 1, 0] | S) p_c(c | C) \exp - \alpha & s = [1, 1, 0] \end{cases} \quad (18)$$

Given two Normal distributions for the random variable  $x$ ,  $\mathcal{N}(x | x_1, \sigma_1^2)$  and  $\mathcal{N}(x | x_2, \sigma_2^2)$ ; with means and variances  $x_1, \sigma_1^2; x_2, \sigma_2^2$ , respectively, it can be shown that

$$\int_x \mathcal{N}(x | x_1, \sigma_1^2) \mathcal{N}(x | x_2, \sigma_2^2) dx = \frac{1}{2\pi\sqrt{\sigma_1^2\sigma_2^2}} \exp - \left[ \frac{(\bar{x} - x_1)^2}{2\sigma_1^2} + \frac{(\bar{x} - x_2)^2}{2\sigma_2^2} \right] \quad (19)$$

where  $\bar{x} = (\sigma_2^2 x_1 + \sigma_1^2 x_2) / (\sigma_2^2 + \sigma_1^2)$ . See [70] for some background material on integrating Gaussians.

Integrating out  $i_n$ , and using the above where necessary, yields the following expression for  $p(s, c | \mathbf{I}, S, \mathbf{C})$  (see (20) and (21) at the bottom of the page). The various  $(\hat{\cdot})$  terms have values as follows:

$$\hat{i}_b = \frac{\sigma_e^2 g_n + \sigma_\mu^2 i_{n-1}}{\sigma_e^2 + \sigma_\mu^2} \quad (22)$$

$$\hat{i}_f = \frac{\sigma_e^2 g_n + \sigma_\mu^2 i_{n+1}}{\sigma_e^2 + \sigma_\mu^2} \quad (23)$$

$$\hat{i} = \frac{\sigma_e^2 g_n + \sigma_\mu^2 (i_{n-1} + i_{n+1})}{\sigma_e^2 + 2\sigma_\mu^2} \quad (24)$$

$$\hat{i}_1 = \frac{i_{n-1} + i_{n+1}}{2}. \quad (25)$$

Finally integrating out  $c$  from the above expression yields the required conditional for  $s$ ,  $p(s | S, \cdot)$  as shown in (27) and (28)

at the bottom of the page. The various  $\hat{c}$  terms have values as follows:

$$\hat{c}_0 = \frac{\sum_{k=1}^8 \lambda_k^c (1 - u(\vec{x}, \vec{x} + \vec{v}_k)) c(\vec{x} + \vec{v}_k)}{\sum_{k=1}^8 \lambda_k^c (1 - u(\vec{x}, \vec{x} + \vec{v}_k))} \quad (29)$$

$$\hat{c}_1 = \frac{\sigma_\mu^2 \hat{c}_0 + \sigma_c^2 g_n}{\sigma_c^2 + \sigma_\mu^2} \quad (30)$$

$$\sigma_c^2 = \frac{1}{2 \sum_{k=1}^8 \lambda_k^c (1 - u(\vec{x}, \vec{x} + \vec{v}_k))}. \quad (31)$$

The terms  $\hat{i}, \hat{c}$  are in fact the least squares estimates of the marginalized variables [70]. Note that the constants of proportionality are derived from the normalizing factors for the Gaussian distributions for the corruption likelihood and clean data likelihood (where it exists) in each state case. None of the other normalizing factors are important since all the priors exist in each state option.

#### A. Solution

To maximize the conditional  $p(s, i_n, c | \mathbf{I}, S, \mathbf{C})$  each factor is maximized recursively. Thus the first maximization is over the six possible states of  $s$  using (28). The second maximization uses this value of  $s, \hat{s}$  in choosing  $c$  by maximizing the (21).

Because the value of  $\hat{s}$  immediately selects one of the options in the set of six, and the prior for  $c$  is GMRF, the maximization with respect to  $c$  is analytic. Thus when  $\hat{s}$  is such that  $b = 0$ ,

$$\int_{i_n} p(s, i_n, c | \mathbf{I}, S, \mathbf{C}) = p_s(s, c | \mathbf{I}, S, \mathbf{C}) \quad (20)$$

$$\propto \begin{cases} \exp - \left( \frac{(g_n - \hat{i}_b)^2}{2\sigma_\mu^2} \right) \exp - \left( \frac{(i_{n-1} - \hat{i}_b)^2}{2\sigma_e^2} \right) p_s(s = [0, 0, 1] | S) p_c(c | C) \exp - \alpha & s = [0, 0, 1] \\ \exp - \left( \frac{(g_n - \hat{i}_f)^2}{2\sigma_\mu^2} \right) \exp - \left( \frac{(i_{n+1} - \hat{i}_f)^2}{2\sigma_e^2} \right) p_s(s = [0, 1, 0] | S) p_c(c | C) \exp - \alpha & s = [0, 1, 0] \\ \exp - \left( \frac{(g_n - \hat{i})^2}{2\sigma_\mu^2} \right) \exp - \left( \frac{(i_{n-1} - \hat{i})^2 + (i_{n+1} - \hat{i})^2}{2\sigma_e^2} \right) p_s(s = [0, 0, 0] | S) p_c(c | C) & s = [0, 0, 0] \\ \exp - \left( \frac{(g_n - c)^2}{2\sigma_\mu^2} \right) \exp - \left( \frac{(i_n - \hat{i}_1)^2 + (i_n - \hat{i}_1)^2}{2\sigma_e^2} \right) p_s(s = [1, 0, 0] | S) p_c(c | C) & s = [1, 0, 0] \\ \exp - \left( \frac{(g_n - c)^2}{2\sigma_\mu^2} \right) p_s(s = [1, 0, 1] | S) p_c(c | C) \exp - \alpha & s = [1, 0, 1] \\ \exp - \left( \frac{(g_n - c)^2}{2\sigma_\mu^2} \right) p_s(s = [1, 1, 0] | S) p_c(c | C) \exp - \alpha & s = [1, 1, 0] \end{cases} \quad (21)$$

$$\int_c p(s, c | \mathbf{I}, S, \mathbf{C}, \cdot) = p_s(s | \mathbf{I}, S, \mathbf{C}, \cdot) \quad (27)$$

$$\propto \begin{cases} \exp - \left( \frac{(g_n - \hat{i}_b)^2}{2\sigma_\mu^2} \right) \exp - \left( \frac{(i_{n-1} - \hat{i}_b)^2}{2\sigma_e^2} \right) p_s(s = [0, 0, 1] | S) p(c = \hat{c}_0 | C) \exp - \alpha & s = [0, 0, 1] \\ \exp - \left( \frac{(g_n - \hat{i}_f)^2}{2\sigma_\mu^2} \right) \exp - \left( \frac{(i_{n+1} - \hat{i}_f)^2}{2\sigma_e^2} \right) p_s(s = [0, 1, 0] | S) p(c = \hat{c}_0 | C) \exp - \alpha & s = [0, 1, 0] \\ \exp - \left( \frac{(g_n - \hat{i})^2}{2\sigma_\mu^2} \right) \exp - \left( \frac{(i_{n-1} - \hat{i})^2 + (i_{n+1} - \hat{i})^2}{2\sigma_e^2} \right) p_s(s = [0, 0, 0] | S) p(c = \hat{c}_0 | C) & s = [0, 0, 0] \\ \exp - \left( \frac{(g_n - \hat{c}_1)^2}{2\sigma_\mu^2} \right) \exp - \left( \frac{(i_{n-1} - \hat{i}_1)^2 + (i_{n+1} - \hat{i}_1)^2}{2\sigma_e^2} \right) p_s(s = [1, 0, 0] | S) p(c = \hat{c}_1 | C) & s = [1, 0, 0] \\ \exp - \left( \frac{(g_n - \hat{c}_1)^2}{2\sigma_\mu^2} \right) p_s(s = [1, 0, 1] | S) p(c = \hat{c}_1 | C) \exp - \alpha & s = [1, 0, 1] \\ \exp - \left( \frac{(g_n - \hat{c}_1)^2}{2\sigma_\mu^2} \right) p_s(s = [1, 1, 0] | S) p(c = \hat{c}_1 | C) \exp - \alpha & s = [1, 1, 0] \end{cases} \quad (28)$$

there is only one p.d.f. involving  $c$  and therefore the estimate that maximizes the log posterior is  $\hat{c}_0$ . Otherwise the estimate is  $\hat{c}_1$ . This is sensible since it implies that in regions that there is *no corruption* the corrupting data field is *hidden* and so must be interpolated from nearby values. When there is corruption,  $c$  is directly observed in the image and so the estimate is a noise reduced version of the observed image.

In a similar manner, with (18), the estimate for  $i$  is generated using  $\hat{s}$  and the estimate for  $c$ . When there is no corruption, the estimate is a noise reduced version of the observed sequence allowing for occlusion,  $\hat{i}$ . When the image is corrupted the estimate uses image data either in the previous or next frames or both previous and next frames i.e.,  $\hat{i}_f, \hat{i}_b$  or  $\hat{i}_1$ .

## APPENDIX II

### FACTORIZING THE CONDITIONAL FOR MOTION

Consider now operating on a blockwise basis. Within each block, the motion parameters  $\mathbf{d}_{n,n-1}, \mathbf{d}_{n,n+1}, \sigma_e^2$  are constant. The occlusion parameters are considered alongside the blotch detection indices and are not estimated with the motion information. It is required therefore, to manipulate  $p(\mathbf{d}_{n,n-1}, \mathbf{d}_{n,n+1}, \sigma_e^2 | I_n, I_{n-1}, I_{n+1})$ . Recall from Section G that the idea is to group together the pixelwise parameters  $[s, I, c]$  and the motion parameters  $[d, \sigma_e^2]$  and iterate between the two processes. This is possible because it is possible to handle each group *jointly*. The motion is handled by using the factorization as follows (considering backward motion only)

$$\begin{aligned} & p(\mathbf{d}_{n,n-1}, \sigma_e^2 | I_n, I_{n-1}, O_b, D) \\ & \propto p(\sigma_e^2 | \mathbf{d}_{n,n-1}, I_n, I_{n-1}, O_b, D) \\ & p(\mathbf{d}_{n,n-1} | I_n, I_{n-1}, O_b, D) \end{aligned} \quad (33)$$

Note that the *current estimate* of the clean image  $I_n$  is being used here. The first term in the factorization is

$$p(\sigma_e^2 | \mathbf{d}_{n,n-1}, I_n, I_{n-1}, O_b, D) = \frac{p(\sigma_e^2, \mathbf{d}_{n,n-1} | \cdot)}{p(\mathbf{d}_{n,n-1} | \cdot)}. \quad (34)$$

This requires integrating out  $\sigma_e^2$  from the posterior  $p(\sigma_e^2, \mathbf{d}_{n,n-1} | \cdot)$ . Given  $I_n$  the posterior is written as follows, using (1) and incorporating priors for  $\mathbf{d}$  and  $\sigma_e^2$  shown in (35) at the bottom of the page where the sum is over  $N$  pixels in a block. To simplify matters define  $E(\mathbf{i}, \mathbf{d})$  as follows:

$$\begin{aligned} E(\mathbf{i}, \mathbf{d}_{n,n-1}) &= \sum_{\vec{x}} (I_n(\vec{x}) - I_{n-1}(\vec{x} + \mathbf{d}_{n,n-1}))^2 \\ & \times (1 - O_b(\vec{x})) + O_b(\vec{x})\alpha. \end{aligned} \quad (36)$$

The posterior can now be written

$$p(\sigma_e^2, \mathbf{d}_{n,n-1} | \cdot) \propto \left\{ \exp - \left[ \frac{E(\mathbf{i}, \mathbf{d})}{2\sigma_e^2} \right] \right\} p(O_b)p(\mathbf{d} | D)p(\sigma_e^2). \quad (37)$$

There are just two terms involving  $\sigma_e^2$  and given  $p(\sigma_e^2) = 1/\sigma_e^2$  the posterior can now be written

$$p(\sigma_e^2, \mathbf{d}_{n,n-1} | \cdot) \propto \frac{1}{\sigma_e^2} \frac{1}{\sqrt{2\pi\sigma_e^2}^{N/2}} \exp - \left[ \frac{E(\mathbf{i}, \mathbf{d})}{2\sigma_e^2} \right]. \quad (38)$$

To proceed with integration, the expression can be rearranged to expose a relationship with the Inverted Gamma (IG) distribution as follows:

$$p(\sigma_e^2, \mathbf{d}_{n,n-1} | \cdot) \propto (\sigma_e^2)^{-(\frac{N}{2}+1)} \frac{1}{\sqrt{2\pi}^N} \exp - \left[ \frac{E(\mathbf{i}, \mathbf{d})}{2\sigma_e^2} \right]. \quad (39)$$

For some random variable,  $r$ , the IG distribution with parameters  $\beta_1, \beta_2$  is as follows [70]:

$$p(r | \beta_1, \beta_2) = \frac{\beta_2^{\beta_1}}{\Gamma(\beta_1)} r^{-(\beta_1+1)} \exp - \left[ \frac{\beta_2}{r} \right]. \quad (40)$$

Letting  $\beta_1 = N/2$  and  $\beta_2 = E(\mathbf{d})/2$  and  $r = \sigma_e^2$

$$p(\sigma_e^2, \mathbf{d}_{n,n-1} | \cdot) \propto \frac{1}{\sqrt{2\pi}^N} \frac{\Gamma(\beta_1)}{\beta_2^{\beta_1}} \frac{\beta_2^{\beta_1}}{\Gamma(\beta_1)} (r)^{-(\beta_1+1)} \exp - \left[ \frac{\beta_2}{r} \right]. \quad (41)$$

Thus, the required conditional p.d.f.  $p(\sigma_e^2 | \cdot)$  is in fact an Inverted Gamma distribution and since  $\int IG(r | \beta_1, \beta_2) dr = 1$  the required conditional for  $\mathbf{d}$  is as follows:

$$\begin{aligned} p(\mathbf{d} | \cdot) &= \int p(\sigma_e^2, \mathbf{d} | \cdot) d\sigma_e^2 \\ &= \int \frac{1}{\sqrt{2\pi}^N} \frac{\Gamma(\beta_1)}{\beta_2^{\beta_1}} \frac{\beta_2^{\beta_1}}{\Gamma(\beta_1)} (\sigma_e^2)^{-(\beta_1+1)} \exp - \left[ \frac{\beta_2}{\sigma_e^2} \right] d\sigma_e^2 \\ &= \frac{1}{\sqrt{2\pi}^N} \frac{\Gamma(\beta_1)}{\beta_2^{\beta_1}} \int IG(r | \beta_1, \beta_2) dr \\ &= \frac{1}{\sqrt{2\pi}^N} \frac{\Gamma(\beta_1)}{\beta_2^{\beta_1}} \\ & \propto E(\mathbf{i}, \mathbf{d})^{-N/2}. \end{aligned} \quad (42)$$

Hence, the two results are obtained as follows:

$$\begin{aligned} p(\sigma_e^2 | \mathbf{d}, \cdot) &= IG \left( \sigma_e^2 \left| \frac{N}{2}, \frac{E(\mathbf{i}, \mathbf{d})}{2} \right. \right) \\ p(\mathbf{d} | \mathbf{i}, D) &= E(\mathbf{i}, \mathbf{d})^{-N/2} p_d(\mathbf{d} | D). \end{aligned} \quad (43)$$

This allows the factorization  $p(\sigma_e^2, \mathbf{d} | \cdot) = p(\sigma_e^2 | \mathbf{d}, D, \cdot) p(\mathbf{d} | D, \cdot)$  to be used.

## ACKNOWLEDGMENT

The author would like to thank the many members of the AURORA and BRAVA EU consortia who created an educational atmosphere over several years for this kind of work to

$$p(\sigma_e^2, \mathbf{d}_{n,n-1} | \cdot) \propto \left\{ \exp - \left[ \frac{\sum_{\vec{x}} (I_n(\vec{x}) - I_{n-1}(\vec{x} + \mathbf{d}_{n,n-1}))^2 (1 - O_b(\vec{x}))}{2\sigma_e^2} \right] \right\} p(O_b)p(\mathbf{d} | D)p(\sigma_e^2) \quad (35)$$

arise. It is important as well to acknowledge the input from several archives and workers in archive houses, including the BBC (R&D Unit), INA (Jean-Hugues Chenot, Raphael Bornard and Louis Laborelli), RTP (Joao Sequeira) both for material and the understanding of how defects develop. It is clear that this kind of work cannot be well informed without communication with experienced archivists and signal processing researchers working within archive houses.

## REFERENCES

- [1] P. V. M. Roosmalen, A. Kokaram, and J. Biemond, "Noise reduction of image sequences as preprocessing for mpeg2 encoding," in *Eur. Conf. Signal Processing (EUSIPCO'98)*, vol. 4, Sept. 1998, pp. 2253–2256.
- [2] F. Chen and D. Suter, "Motion estimation for noise reduction in historical films: Mpeg encoding effects," in *Proc. 6th Digital Image Computing: Techniques and Applications (DICTA2002) Conf.*, 2002, pp. 207–212.
- [3] D. Suter and S. Richardson, "Historical film restoration and video coding," in *Proc. Picture Coding Symp.*, 1996, pp. 389–394.
- [4] P. Schallauer, G. Thallinger, and M. J. Addis, "Diamant—Digital film manipulation platform," in *Proc. IEE Seminar on Digital Restoration of Film and Video Archives (Ref. no. 2001/049)*, Jan. 2001, pp. 3/1–3/3.
- [5] S. M. Sommerville, "Archangel—Automated real-time archive restoration," in *Proc. IEE Seminar on Digital Restoration of Film and Video Archives (Ref. no. 2001/049)*, Jan. 2001, pp. 5/1–5/30.
- [6] A. C. Kokaram, *Motion Picture Restoration: Digital Algorithms for Artefact Suppression in Degraded Motion Picture Film and Video*. Berlin, Germany: Springer-Verlag, 1998.
- [7] —, "Reconstruction of severely degraded image sequence," in *Image Analysis and Processing*. Berlin, Germany: Springer-Verlag, Sept. 1997, vol. 2, pp. 773–780.
- [8] D. Sidorov and A. Kokaram, "Suppression of moiré patterns via spectral analysis," *Proc. SPIE*, vol. 4671, pp. 895–906, Jan. 2002.
- [9] —, "Removing moire from degraded video archives," in *Proc. 11th Eur. Conf. Signal Processing (EUSIPCO 2002)*, Sept. 2002.
- [10] P. J. W. Rayner, S. Armstrong, and A. C. Kokaram, "Restoring video images taken from scratched 2-inch tape," in *Workshop on Non-Linear Model Based Image Analysis, NMBIA'98*, S. Marshall, N. Harvey, and D. Shah, Eds. Berlin, Germany: Springer-Verlag, Sept. 1998, pp. 83–88.
- [11] S. Manhall and N. R. Harvey, "Film and Video Archive Restoration Using Mathematical Morphology," Jan. 2001.
- [12] S.-J. Ko, S.-H. Lee, S.-W. Jeon, and E.-S. Kang, "Fast digital image stabilizer based on gray-coded bit-plane matching," *IEEE Trans. Consumer Electron.*, vol. 45, pp. 598–603, Aug. 1999.
- [13] K. Uomori, A. Morimura, H. Ishii, T. Sakaguchi, and Y. Kitamura, "Automatic image stabilizing system by full-digital signal processing," *IEEE Trans. Consumer Electron.*, vol. 36, pp. 510–519, Aug. 1990.
- [14] J. Tucker and A. de Sam Lazaro, "Image stabilization for a camera on a moving platform," in *Proc. IEEE Pacific Rim Conf. Communications, Computers, Signal Processing*, vol. 2, May 1993, pp. 734–737.
- [15] S.-J. Ko, S.-H. Lee, and K.-H. Lee, "Digital image stabilizing algorithms based on bit-plane matching," *IEEE Trans. Consumer Electron.*, vol. 44, pp. 617–622, Aug. 1998.
- [16] K. Ratakonda, "Real-time digital video stabilization for multimedia applications," in *Proc. IEEE Int. Symp. Circuits and Systems*, vol. 4, Monterey, CA, May 1998, pp. 69–72.
- [17] T. Vlachos, "Simple method for estimation of global motion parameters using sparse translational motion vector fields," *Electron. Lett.*, vol. 34, no. 1, pp. 60–62, Jan. 1998.
- [18] J.-M. Odobez and P. Bouthémy, "Robust multiresolution estimation of parametric motion models," *J. Vis. Commun. Image Representat.*, vol. 6, pp. 348–365, 1995.
- [19] L. Hill and T. Vlachos, "On the estimation of of global motion using phase correlation for broadcasting applications," in *Proc. 7th Int. Conf. Image Processing Applications*, vol. 2, July 1999, pp. 721–725.
- [20] —, "Global and local motion estimation using higher-order search," in *Proc. 5th Meeting on Image Recognition and Understanding (MIRU 2000)*, vol. 1, July 2000, pp. 18–21.
- [21] F. Dufaux and J. Konrad, "Efficient, robust and fast global motion estimation for video coding," *IEEE Trans. Image Processing*, vol. 9, pp. 497–501, 2000.
- [22] A. Smolic and J.-R. Ohm, "Robust global motion estimation using a simplified m-estimator approach," in *Proc. IEEE Int. Conf. Image Processing*, Vancouver, BC, Canada, Sept. 2000.
- [23] W. Qi and Y. Zhong, "New robust global motion estimation approach used in MPEG-4," *J. Tsinghua Univ. Sci. Technol.*, 2001.
- [24] P. M. B. Van Roosmalen, R. L. Lagendijk, and J. Biemond, "Flicker reduction in old film sequences," in *Time-Varying Image Processing and Moving Object Recognition*. Amsterdam, The Netherlands: Elsevier, 1997, vol. 4, pp. 9–17.
- [25] —, "Correction of intensity flicker in old film sequences," *IEEE Trans. Circuits Syst. Video Technol.*, submitted for publication.
- [26] T. Vlachos and G. A. Thomas, "Motion estimation for the correction of twin-lens telecine flicker," in *IEEE Int. Conf. Image Processing*, vol. 1, Sept. 1996, pp. 109–112.
- [27] P. Read and M.-P. Meyer, *Restoration of Motion Picture Film*. New York: Butterworth Heinemann, 2000.
- [28] V. Bruni and D. Vitulano, "Scratch detection via underdamped harmonic motion," in *Proc. Int. Conf. Pattern Recognition*, vol. III, 2002, pp. 887–890.
- [29] L. Joyeux, S. Boukir, and B. Besserer, "Film line scratch removal using kalman filtering and bayesian restoration," in *Proc. 5th IEEE Workshop on Applications of Computer Vision*, 2000, pp. 8–13.
- [30] L. Joyeux, S. Boukir, B. Besserer, and O. Buisson, "Detection and removal of line scratches in motion picture films," in *IEEE Int. Conf. Computer Vision and Pattern Recognition*, June 1999, pp. 548–553.
- [31] —, "Reconstruction of degraded image sequences. Application to film restoration," *Image Vis. Comput.*, no. 19, pp. 503–516, 2001.
- [32] A. Kokaram, "Detection and removal of line scratches in degraded motion picture sequences," *Signal Process. VIII*, vol. I, pp. 5–8, Sept. 1996.
- [33] A. Kokaram, R. Morris, W. Fitzgerald, and P. Rayner, "Detection of missing data in image sequences," *IEEE Trans. Image Processing*, pp. 1496–1508, Nov. 1995.
- [34] —, "Interpolation of missing data in image sequences," *IEEE Trans. Image Processing*, pp. 1509–1519, Nov. 1995.
- [35] P. M. B. van Roosmalen, "Restoration of Archived Film and Video," Ph.D. dissertation, Tech. Univ. Delft, Delft, The Netherlands, Sept. 1999.
- [36] P. V. M. Roosmalen, A. Kokaram, and J. Biemond, "Fast high quality interpolation of missing data in image sequences using a controlled pasting scheme," in *Proc. IEEE Conf. Acoustics Speech and Signal Processing (ICASSP'99)*, vol. IMDSP 1.2, Mar. 1999, pp. 3105–3108.
- [37] G. R. Arce, "Multistage order statistic filters for image sequence processing," *IEEE Trans. Signal Processing*, vol. 39, pp. 1146–1161, May 1991.
- [38] R. Storey, "Electronic detection and concealment of film dirt," *SMPTE J.*, pp. 642–647, June 1985.
- [39] M. J. Nadenau and S. K. Mitra, "Blotch and scratch detection in image sequences based on rank ordered differences," in *Proc. 5th Int. Workshop on Time-Varying Image Processing and Moving Object Recognition*, Sept. 1996.
- [40] A. C. Kokaram, "Advances in the detection and reconstruction of blotches in archived film and video," in *IEE Seminar on Digital Restoration of Film and Video Archives (Ref. no. 2001/049)*, Jan. 2001, pp. 7/1–7/6.
- [41] A. Kokaram and S. J. Godsill, "MCMC for joint noise reduction and missing data treatment in degraded video," *IEEE Trans. Signal Processing*, vol. 50, pp. 189–205, Feb. 2002.
- [42] R. Storey, "Electronic Detection and Concealment of Film Dirt," 1984.
- [43] A. Kokaram, R. Bornard, A. Rares, D. Sidorov, J.-H. Chenot, L. Laborelli, and J. Biemond, "Robust and automatic digital restoration systems: Coping with reality," in *Proc. Int. Broadcasting Convention (IBC)*, 2002.
- [44] A. Kokaram, R. Bornard, A. Rares, D. Sidorov, J.-H. Chenot, L. Laborelli, and J. Biemond, "Automatic digital restoration systems: Coping with reality," *J. Soc. Motion Picture Television Eng.*, to be published.
- [45] F. Paisan and A. Crise, "Restoration of signals degraded by impulsive noise by means of a low distortion, nonlinear filter," *Signal Process.*, vol. 6, pp. 67–76, 1984.
- [46] E. D. Ferrandière, *Motion Picture Restoration Using Morphological Tools*. Norwell, MA: Kluwer, May 1999, pp. 361–368.
- [47] E. D. Ferrandière and J. Serra, "Detection of local defects in old motion pictures," in *Proc. VII Nat. Symp. Pattern Recognition Image Analysis*, Apr. 1997, pp. 145–150.
- [48] E. Ferrandière, *Mathematical Morphology and Motion Picture Restoration*. New York: Wiley, 2001.
- [49] O. Buisson, "Analyse Deséquences d'images Haute résolution, Application à la Restauration Numérique de Films Cinématographiques." Ph.D. dissertation, Univ. de La Rochelle, Rochelle, France, Dec. 1997.
- [50] E. D. Ferrandière, "Restauration Automatique de Films Anciens." Ph.D. dissertation, Ecole des Mines de Paris, Paris, France, Dec. 1997.

- [51] O. Buisson, B. Besserer, S. Boukir, and F. Helt, *Deterioration Detection for Digital Film Restoration*. New York: IEEE Press, June 1997, vol. 1, pp. 78–84.
- [52] L. Tenze, G. Ramponi, and S. Carrato, “Blotches correction and contrast enhancement for old film pictures,” in *IEEE Int. Conf. Image Processing*, 2000, p. TP06.05.
- [53] —, “Robust detection and correction of blotches in old films using spatio-temporal information,” *Proc. SPIE*, Jan. 2002.
- [54] T. Saito, T. Komatsu, T. Ohuchi, and T. Seto, “Image processing for restoration of heavily-corrupted old film sequences,” in *Proc. Int. Conf. Pattern Recognition 2000*, vol. 3, 2000, pp. 17–20.
- [55] G. W. Boon, M. N. Chong, S. Kalra, and D. Krishnan, “Bidirectional 3D autoregressive model approach to motion picture restoration,” in *Proc. IEEE Int. Conf. Acoustics and Signal Processing*, Apr. 1996, pp. 2275–2278.
- [56] S. Kalra, M. N. Chong, and D. Krishnan, “A new autoregressive (AR) model based algorithm for motion picture restoration,” in *Proc. IEEE Int. Conf. Acoustics and Signal Processing*, Apr. 1997, pp. 2557–2560.
- [57] M. Bertalmio, G. Sapiro, V. Caselles, and C. Ballester, “Image inpainting,” in *Proc. SIGGRAPH*, 2000.
- [58] M. Bertalmio *et al.*, “Filling-in by joint interpolation of vector fields and gray levels,” *IEEE Trans. Image Processing*, vol. 10, Aug. 2001.
- [59] A. Kokaram, “Parametric texture synthesis for filling holes in pictures,” in *Proc. IEEE Int. Conf. Image Processing 2002*, Sept. 2002.
- [60] —, “A statistical framework for picture reconstruction using ar models,” in *Proc. Eur. Conf. Computer Vision Workshop on Statistical Methods in Video Processing (ISBN 0-9581044-0-9)*, Australia, June 2002, pp. 73–78.
- [61] A. A. Efros and T. K. Leung, “Texture synthesis by nonparametric sampling,” in *Proc. IEEE Int. Conf. Computer Vision (ICCV)*, vol. 2, Sept. 1999, pp. 1033–1038.
- [62] A. A. Efros and W. T. Freeman, “Image quilting for texture synthesis and transfer,” in *Proc. SIGGRAPH*, 2001, pp. 341–346.
- [63] R. Bornard, “Approches Probabilistes Appliquées À La Restauration Numérique D’Archives Télévisées,” Ph.D. dissertation, École Centrale Des Arts Et Manufactures, École Centrale Paris, Paris, France, 2002.
- [64] R. Bornard, E. Lecan, L. Laborelli, and J.-H. Chenot, “Missing data correction in still images and image sequences,” *ACM Multimedia*, Dec. 2002.
- [65] R. D. Morris, “Image Sequence Restoration Using Gibbs Distributions,” Ph.D. dissertation, Cambridge Univ., Cambridge, U.K., 1995.
- [66] A. Kokaram and S. Godsill, “Joint detection, interpolation, motion and parameter estimation for image sequences with missing data,” in *Proc. IEEE Int. Conf. Image Processing*, Oct. 1997, pp. 191–194.
- [67] A. Kokaram, “Practical MCMC for missing data treatment in degraded video,” in *Proc. Eur. Conf. Computer Vision Workshop Statistical Methods Video Processing*, Australia, June 2002, pp. 85–90.
- [68] A. Kokaram and P. Rayner, “A system for the removal of impulsive noise in image sequences,” *Proc. SPIE*, pp. 322–331, Nov. 1992.
- [69] 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 (IPTA’93)*, June 1993.
- [70] J. J. O. Ruanaidh and W. J. Fitzgerald, *Numerical Bayesian Methods Applied to Signal Processing*. Berlin, Germany: Springer-Verlag, 1996.
- [71] J. Konrad and E. Dubois, “Bayesian estimation of motion vector fields,” *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 14, Sept. 1992.
- [72] J. Besag, “On the statistical analysis of dirty pictures,” *J. R. Statist. Soc. B*, vol. 48, pp. 259–302, 1986.
- [73] A. Kokaram and S. J. Godsill, “Joint noise reduction, motion estimation, missing data reconstruction and model parameter estimation for degraded motion pictures,” *Proc. SPIE Int.*, July 1998.
- [74] A. Katsagellos, J. Driessen, S. Efstratiadis, and R. Legendijk, “Spatio-temporal motion compensated noise filtering of image sequences,” in *Proc. SPIE VCIP*, 1989, pp. 61–70.
- [75] E. Dubois and S. Sabri, “Noise reduction in image sequences using motion compensated temporal filtering,” *IEEE Trans. Commun.*, vol. COM-32, pp. 826–831, July 1984.
- [76] M. N. Chong, P. Liu, W. B. Goh, and D. Krishnan, “A new spatio-temporal MRF model for the detection of missing data in image sequences,” in *IEEE Int. Conf. Acoustics Signal Processing*, vol. 4, Apr. 1997, pp. 2977–2980.
- [77] J. Biemond, A. Rares, and M. J. T. Reinders, “Statistical analysis of pathological motion areas,” in *Proc. 2001 IEE Seminar on Digital Restoration of Film and Video Archives*, London, U.K., Jan. 2001.
- [78] —, “Complex event classification in degraded image sequences,” in *Proc. IEEE Int. Conf. Image Processing*, Thessaloniki, Greece, Oct. 2001.
- [79] —, “Image sequence restoration in the presence of pathological motion and severe artifacts,” in *Proc. IEEE Int. Conf. Acoustics, Speech, Signal Processing*, Orlando, FL, May 2002.
- [80] —, “A spatiotemporal image sequence restoration algorithm,” in *Proc. IEEE Int. Conf. Image Processing 2002*, New York, Sept. 2002.
- [81] I. McLean and S. Witt, “Telecine noise reduction,” in *Proc. IEE Seminar Digital Restoration of Film and Video Archives*, Jan. 2001, pp. 2/1–2/6.



**Anil C. Kokaram** (M’92) was born in Trinidad. He graduated in electrical and information sciences from the University of Cambridge, Cambridge, U.K., in 1989. He received the the Ph.D. in engineering from the same university for his dissertation entitled “Motion Picture Restoration” in 1993.

He then joined the Signal Processing Group at the Cambridge University Engineering Department. He was then a Research Associate in that group and later appointed to a Fellowship of Churchill College, Cambridge. From 1993 to 1998, he worked on the EU project entitled “Automated restoration of original film and video archives” (AURORA). In 1998, he was appointed to a Lectureship in the Electronic and Electrical Engineering Department, Trinity College, Dublin, Ireland. His research interests lie principally in the areas of image and video processing including motion estimation, texture synthesis, and video and film restoration/post-production in particular. His fundamental research areas include Bayesian methods in signal processing, and fast (practical) algorithms for MCMC used for video material. He has recently diversified into such application areas as multimedia over wireless and multimedia information retrieval. He is Coordinator of the EU project entitled “Models for Unified Multimedia Information Retrieval” (<http://www.moumir.org>). He has published over 40 papers in refereed journals and conference proceedings and the book *Motion Picture Restoration* (New York: Springer-Verlag, 1998).

Dr. Kokaram was recently awarded a Fellowship from Trinity College. He was appointed Associate Editor of the IEEE TRANSACTIONS ON IMAGE PROCESSING in 2002.