

# Removal of Line Artefacts for Digital Dissemination of Archived Film and Video \*

Anil C. Kokaram

Electrical Engineering Dept., Trinity College, Dublin 2, Ireland.

anil.kokaram@tcd.ie

Tel : +353 1 608 3412, Fax : +353 1 677 2442

## Abstract

*Line scratches are a common form of degradation in archived motion picture film. The automatic restoration of such material has become of increasing interest in the last few years with the rise of consumer digital video applications and the need to supply more programming material of an acceptable quality in a multimedia context. This paper introduces mechanisms for removing line artefacts from degraded images. It also proposes schemes which solve for both the image data and the hidden model parameters simultaneously. The algorithms introduced are based on the use of the 2D-AR process as an underlying image model and employ a Bayesian approach. The Gibbs sampler is used to draw samples from the underlying image texture and so create a convincing interpolation in the region previously obscured by the line feature.*

Line scratches are a common problem in archived film. The problem is transferred to video during the telecine transfer process. The artefact is easily visible as a vertical line of bright or dark intensity, oriented more or less vertically over much of the image. It may be caused when material from some particle is smeared vertically over the film material in the projector or by the abrasion of the film as it passes over some particle caught in the mechanism.

As digital video broadcasters and those wishing to release material on DVD are faced with the prospect of restoring previously released, but degraded material, it has become more attractive to seek automatic methods for restoration. The EU project AURORA, which terminates at the end of 1998 was set up in 1995 to address the problem of creating real time hardware for achieving good quality automatic restoration. Line scratch removal was one of the problems addressed, although the solution developed will not appear in hardware just yet.

The general idea has been to split the line removal process into two stages, an initial detection step followed by an interpolation or reconstruction step. Previous work [1, 3] has introduced deterministic and stochastic methods for the detection of this artefact. This paper concentrates on the interpolation problem. The removal of the defect is treated as a missing data problem. The following material explores stochastic varieties of reconstruction and presents a fast scheme for model based interpolation. In particular the introduced ideas are presented based on the requirement to solve for all the parameters *jointly*. The algorithms presented here therefore mark a distinct departure from the usual ideas of treating model parameter estimation separate from image data interpolation.

## 1 Line removal

The 2D Autoregressive (AR) model can be used as the basis for image interpolation in the same manner as that introduced in [2, 3] for image sequences using the 3D (spatiotemporal) AR model. However the region to be interpolated now extends over much of the image and so the underlying data is statistically non-stationary. Thus those interpolation ideas must be modified for use here. Furthermore, because the artefact often persists in the same location across several frames, it is not possible to reliably use the temporal redundancy in image sequences to help with reconstruction. In what follows, the worst case scenario is assumed in which the same location in consecutive frames is corrupted and there is no motion in the region of the defect. The problem is then essentially a purely spatial one and the use of a 2D AR process is considered for interpolating across such huge connected regions.

## 2 Missing data interpolation

It is assumed that the underlying image generation process is autoregressive. Thus a pixel in the original, clean image,  $I(\mathbf{x})$ , where  $\mathbf{x}$  is a spatial position vector, can be modelled as

$$I(\mathbf{x}) = \sum_{k=1}^P a_k(\mathbf{x})I(\mathbf{x} + \mathbf{q}_k) + e(\mathbf{x}) \quad (1)$$

A particular pixel in frame  $n$  (of an image sequence) at site  $\mathbf{x}$  is therefore predicted by a linear combination of pixels in the current frame plus an added excitation or residual error  $e(\mathbf{x}) \sim N(0, \sigma_e^2(\mathbf{x}))$ . The  $P$  coefficients of the model, which define the linear combination, are denoted  $a_k$  for  $k = 1 \dots P$ . The pixels used in the prediction are called the support or neighbourhood of the model and are mapped by the  $P$  spatial offset vectors  $\mathbf{q}_k$ . Note that the model parameters are non-stationary in order to capture the statistical non-stationarity in interesting images.

The observed degraded sequence  $G_n$  may be modelled by a switching process such that the data that is the corruption (i.e. the line defect) is switched into the original clean data at specific sites. This switching process is monitored by the field (or image)  $b(\mathbf{x})$ . At a corrupted site  $b(\mathbf{x}) = 1$ , otherwise the site is 0. The missing pixels in the image are therefore indicated at sites where  $b(\mathbf{x}) = 1$ . It is assumed that the indicator field has already been determined by a detection process and it now delineates vertical swathes of material which is corrupted. The problem then is to design a process for interpolating the missing data at the sites indicated, using the model defined in equation 1.

### 2.1 A Bayesian framework for joint estimation of missing data and model parameters

Given the corrupted data  $G_n$  it is required to solve for the unknown data  $I(\mathbf{x})$  at sites indicated by  $b(\mathbf{x}) = 1$ . The other hidden unknowns are the non-stationary model coefficients  $\mathbf{a}(\mathbf{x})$  (where  $\mathbf{a}$  is a vector of  $P$  coefficients at location  $\mathbf{x}$ ) and the non-stationary residual variance  $\sigma_e^2(\mathbf{x})$ . To effect this solution it is therefore prudent to manipulate  $p(\theta|G(\mathbf{x}), b(\mathbf{x}))$  which is the probability of all the unknowns (denoted by the vector  $\theta$ ) given the observed degraded data and the missing regions delineated by  $b(\mathbf{x})$ .

It will emerge in later discussion that the model parameters are assumed to be fixed on a block basis. Define the unknown pixels (at sites where  $b(\mathbf{x}) = 1$ ) within such a block as  $\mathbf{i}_u$  and the known pixels as

$\mathbf{i}_k$ . The model parameters within the block are thus  $\mathbf{a}(\mathbf{x}), \sigma_e^2(\mathbf{x})$ . Proceeding in a Bayesian fashion, the conditional may be rewritten in terms of a product of a likelihood and a prior as follows:

$$p(\mathbf{i}_u, \mathbf{a}(\mathbf{x}), \sigma_e^2(\mathbf{x})|\mathbf{i}_k) \propto p(\mathbf{i}_k|\mathbf{i}_u, \mathbf{a}(\mathbf{x}), \sigma_e^2(\mathbf{x})) p(\mathbf{a}(\mathbf{x}))p(\sigma_e^2(\mathbf{x})) \quad (2)$$

#### 2.1.1 The likelihood

The data likelihood  $p(\mathbf{i}_k|\cdot)$  can be derived from the pseudo-likelihood of the data within a block. Note that within a block, the residuals  $e(\mathbf{x})$  will be a sample from a multivariate Gaussian distribution. Scanning the residuals in a raster fashion within this block allows the formation of an excitation vector  $\mathbf{e}$ . Consider a block of pixels of size  $M \times M$  (denote this block by  $m$ ) and allow for a border of pixels around the edges of this block (so that  $(\mathbf{x} + \mathbf{q}_k)$  will never result in a location outside the  $M \times M$  block). An equation for the error at every pixel within a centred  $B \times B$  block inside  $m$  ( $B < M$ ) can be written as below.

$$\mathbf{e} = \mathbf{A}\mathbf{i} \quad (3)$$

where  $\mathbf{i}$  represents an  $M^2 \times 1$  column vector of row ordered pixels from the  $M \times M$  block,  $\mathbf{e}$  is a  $B^2 \times 1$  column vector of errors (made up of prediction error equations in frames), and  $\mathbf{A}$  a matrix of coefficients satisfying the model equation at all the considered points. This coefficient matrix is of size  $B^2 \times M^2$ . The vector  $\mathbf{i}$  contains intensities of both known and unknown pixels.

Because the joint distribution for the residuals  $\mathbf{e}$  inside a block is normal, the data likelihood (for both known and unknown data) can therefore be written

$$p(\mathbf{i}|\mathbf{a}(\mathbf{x}), \sigma_e^2) \propto \frac{1}{\sqrt{2\pi\sigma_e^2}^N} \exp\left(-\frac{\mathbf{i}^T \mathbf{A}^T \mathbf{A} \mathbf{i}}{2\sigma_e^2}\right) \quad (4)$$

where  $N = B \times B$ .

To generate alternate forms for the likelihood the vector  $\mathbf{i}$  can be separated into two vectors  $\mathbf{i}_u$  ( $u$  for unknown) and  $\mathbf{i}_k$  ( $k$  for known), which represent the known and unknown pixel intensities (marked by  $b(\mathbf{x}) = 0, 1$  respectively) then equation 3 can be written as

$$\mathbf{e} = \mathbf{A}_k \mathbf{i}_k + \mathbf{A}_u \mathbf{i}_u \quad (5)$$

Here,  $\mathbf{A}_k, \mathbf{A}_u$  are the coefficient matrices corresponding to the known and unknown data vectors. They are submatrices of the  $\mathbf{A}$  matrix, made by extracting the relevant columns.

### 2.1.2 The priors

For simplicity a uniform prior is assigned to  $\mathbf{a}$  although it is acknowledged that a prior encouraging stable models would be more appropriate. The variance  $\sigma_e^2$  is assigned a non-informative prior  $p(\sigma_e^2) \propto 1/\sigma_e^2$ , following [5].

## 2.2 Solving for the unknowns

The solution is generated by manipulating  $p(\mathbf{i}_u, \mathbf{a}(\mathbf{x}), \sigma_e^2(\mathbf{x}) | \mathbf{i}_k)$ . 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. The Gibbs sampler is used here to generate the required estimates numerically.

## 2.3 The Gibbs sampler

Consider that  $\mathbf{i}$  contains at least the missing pixels (indicated at  $b(\mathbf{x}) = 1$ ) and their immediate AR support. Recall that the missing pixels are denoted  $\mathbf{i}_u$  and the remaining pixels as  $\mathbf{i}_k$ . The Gibbs Sampler then operates iteratively, given some starting guess for the unknowns, by drawing random samples from the conditional posterior distribution for each unknown in turn:

$$\mathbf{a} \sim p(\mathbf{a} | \mathbf{i}, \sigma_e^2, \mathbf{b}); \sigma_e^2 \sim p(\sigma_e^2 | \mathbf{a}, \mathbf{i}, \mathbf{b}); \mathbf{i}_u \sim p(\mathbf{i}_u | \mathbf{a}, \mathbf{i}_k, \sigma_e^2, \mathbf{b})$$

where  $\mathbf{b}$  is a vector containing the indicator variables which are the result of the detection stage, and the location argument  $\mathbf{x}$  has been dropped for simplicity. These conditionals can be derived by manipulation of the joint posterior given in equation 2.

The convergence of the Gibbs Sampler is generally improved if several unknowns are sampled jointly [4]. This is possible using the method of composition [5]. The reader is directed to [3] for details. The result of this manipulation is a recipe for a joint draw for the model parameters within each block as follows.

$$\begin{aligned} p(\mathbf{a} | \mathbf{i}, \mathbf{b}, \sigma_e^2) &= \mathcal{N}_P(\hat{\mathbf{a}}, \sigma_e^2 (\mathbf{I}^T \mathbf{I})^{-1}) \\ p(\sigma_e^2 | \mathbf{i}, \mathbf{b}) &= \text{IG}((N - P)/2, E(\hat{\mathbf{a}}, \mathbf{i})/2) \end{aligned} \quad (6)$$

where  $E(\cdot)$  is the sum squared prediction error in a block using the Least Squares estimate of the coefficients  $\hat{\mathbf{a}}$ . It is defined as follows:

$$E(\hat{\mathbf{a}}, \mathbf{i}) = \sum_{i=1}^N e^2(\mathbf{x}_i) \quad (7)$$

where  $\mathbf{x}_i$  indexes the locations of each of the  $N$  pixels in the inner  $B \times B$  block within  $m$ . The derivation of these expressions can be found in [3]. The draws are undertaken *given* the **current** state of the image data in the current iteration of the Gibbs sampler.

A separate sampling step can then be employed for drawing the unknown pixels using  $\mathbf{i}_u \sim p(\mathbf{i}_u | \mathbf{i}_k, \mathbf{b}, \mathbf{a}, \sigma_e^2)$ . In practice the draw for  $\mathbf{i}_u$  can be performed either on a pixel by pixel basis, or on a block basis. The use of block based draws helps to speed up convergence of the sampler, but does require computationally expensive manipulations in particular a matrix inversion. The pixel by pixel draw requires only simple point wise manipulations, but takes longer to converge.

Manipulation of an alternate form for the likelihood allows an expression for the conditional  $p(\mathbf{i}_u | \mathbf{b}, \mathbf{i}_k, \mathbf{a}, \sigma_e^2)$  to be written as follows:

$$p(\mathbf{i}_u | \mathbf{b}, \mathbf{i}_k, \mathbf{a}, \sigma_e^2) \propto \mathcal{N}(\hat{\mathbf{i}}_u, \sigma_e^2 [\mathbf{A}_u^T \mathbf{A}_u]^{-1}) \quad (8)$$

$$\text{where } \hat{\mathbf{i}}_u = [\mathbf{A}_u^T \mathbf{A}_u]^{-1} \mathbf{A}_u^T \mathbf{A}_k \mathbf{i}_k \quad (9)$$

Thus the Maximum Likelihood estimate, for instance, for the missing data is  $\hat{\mathbf{i}}_u$ .

The utility of the algorithm is increased if there already exist initial estimates for the missing data in particular (of course). Such estimates may be generated by crude median filtering or by using the least squares interpolant in the missing region.

## 3 The JOMBEI algorithm

The steps for interpolating the locations indicated by  $b(\mathbf{x}) = 1$  are enumerated here. The algorithm is called JOMBEI for **JO**int **MO**del **BE**stimation and **IE**nterpolation, to reinforce that both the model parameters *and* the missing data are estimated jointly. Preamble It is assumed that some detection process has delineated regions to be interpolated by entering  $b(\mathbf{x}) = 1$  in the indicator image,  $\mathbf{b}$ . The  $P$  vectors  $\mathbf{q}_k$  are defined by the user to cover some region around a predicted pixel location. A typical choice would be the 8 pixel *causal* neighbourhood. The image is divided into blocks ( $M \times M$ ) which overlap by at least the size of the model support. A *kick start* for the process can be made by interpolating the missing regions using either a median filter (crude) or the least squares interpolant.

### The process

1. Each block is scanned and in those blocks in which missing data is indicated, a sample for the model coefficients  $\mathbf{a}(\mathbf{x}), \sigma_e^2(\mathbf{x})$  (where  $\mathbf{x}$  indicates some reference point in each block, the upper left hand position for instance) is drawn. This is done using

the following steps (where the current block is denoted  $\mathcal{M}$ ). The current state of the pixels in  $m$  is denoted  $\mathbf{i}^0$ .

- (a) Generate the least squares estimate for  $\mathbf{a}$  using the Normal equations with the data in the current block,  $\mathbf{i}^0$ .
- (b) Generate the sum squared prediction error  $E(\hat{\mathbf{a}}, \mathbf{i}^0)$  in that block using this set of coefficients.
- (c) A sample for  $\sigma_e^2, \sigma_s^2$  is then drawn from an Inverted Gamma distribution.

$$p(\sigma_e^2 | \mathbf{i}^0) = \text{IG}((N - P)/2, E/2) \quad (10)$$

This is a more difficult draw and is achieved by the recipe reported in [5].

- (d) A sample for the coefficients,  $\mathbf{a}_s$  in block  $\mathcal{M}$  is then drawn from

$$p(\mathbf{a} | \mathbf{i}^0, \sigma_e^2 = \sigma_s^2) = \mathcal{N}_P(\hat{\mathbf{a}}, \sigma_s^2 (\mathbf{I}^T \mathbf{I})^{-1}) \quad (11)$$

See [5, 3] for details. This draw involves calculating the square root of  $\mathbf{I}^T \mathbf{I}$ .

- (e) Repeat this step at all blocks of interest.
2. Generate a sample for the unknown data  $\mathbf{i}_u$  in each of the block by drawing from  $\mathcal{N}(\hat{\mathbf{i}}_u, \sigma_s^2 [\mathbf{A}_u^T \mathbf{A}_u]^{-1})$ , where  $\hat{\mathbf{i}}_u$  is the least squares (or ML) estimate for the missing data given in equation 9 and the model parameters are set to the samples drawn in that block during the previous step i.e.  $\sigma_s^2, \mathbf{a}_s$ . This draw proceeds in a similar fashion as that outlined with the draw for  $\mathbf{a}_s$  in the first step, since both distributions are multidimensional Normal.
  3. **Alternative:** To reduce computation it is possible to sample at each pixel site in turn. The modifications required are straightforward and result in  $\mathbf{A}_u$  becoming a vector which makes the necessary inverse simple to calculate. Note that for single site sampling, sites should be visited in a checkerboard fashion so that the support used in the next site does not overlap with the support used in the current site. For an 8 pixel causal neighborhood (lag of 2 pixels in each direction) this implies a checker pattern spaced by 5 pixels.
  4. Iterate this pair of draws, the joint sample for the model parameters followed by the sample for the missing data across the entire image until convergence.

## 4 Results: Maximal versus Sampled or Stochastic estimation

Because of the sometimes large gaps that need to be interpolated and the subsequent increase in non-stationarity over these large regions, the least squares or ML interpolant is somewhat blurred (the ‘maximal’ interpolant). This ‘blurring’ or lack of texture is perhaps not so easily noticed in a still image, but in a sequence of images the rather flat interpolant stands out against the otherwise grainy film texture, especially if there is motion. In fact the main contributor to the annoying visibility of the reconstructed region is that it remains in nearly the same place in each frame. In some respects this is a moot point since the line removal is one stage in a chain of restoration modules which will also involve noise reduction. After noise reduction it is harder to see the difference between the texture in the interpolated regions as compared to the surroundings. However, to affect a complete solution it is better to employ samples from the posterior distribution for the unknown data rather than employ the least squares interpolant. The samples tend to have a closer match with the textural behaviour of the film in the surrounding regions simply because they represent the stochastic texture in the real image better than a single ‘maximal’ estimate. Figure 1 illustrates this phenomenon, including the observation that both types of interpolation are superior to the result using subtraction of an estimated line cross section.

Figure 2 shows results from a video sequence contributed by the RTP archives in Lisbon, one of the partners in AURORA. The interpolation was achieved here by using the fully sampled process in which both the missing data and the model parameters were estimated over 200 iterations using a block size of  $16 \times 16$  pixels and the same model as used for the previous example. The reconstructions show good performance. Note that the restoration is incomplete, there are still blotches (Dirt and Sparkle) and noise to be removed from this example, however this is the first step in the chain of events that would result in a clip suitable for DVD release for instance.

## 5 Summary comments

Reconstruction of this kind of persistent artefact remains a difficult problem. The visibility of artefacts due to possible errors in the reconstruction is potentially high. It is particularly important to capture the stochastic nature of the underlying image texture, even if it is just to resynthesise the film grain noise effect

in relatively low activity regions. This paper has presented just such a technique which employs Bayesian analysis to allow the exploration of the space of possible interpolants. The reader is directed to detailed examination of the general interpolation problem in [3] which derives a fast algorithm for calculating the least squared estimate for the missing data. Pictures of a better quality can be found at

[www.mee.tcd.ie/~ack/cd/lines/lines.htm](http://www.mee.tcd.ie/~ack/cd/lines/lines.htm)

## References

- [1] A. Kokaram. Detection and removal of line scratches in degraded motion picture sequences. In *Signal Processing VIII*, volume I, pages 5–8, September 1996.
- [2] A. Kokaram, R. Morris, W. Fitzgerald, and P. Rayner. Detection of missing data in image sequences. *IEEE Image Processing*, pages 1496–1508, November 1995.
- [3] A. C. Kokaram. *Motion Picture Restoration*. Springer Verlag, 1998.
- [4] J. Liu, W. H. Wong, and A. Kong. Covariance structure of the Gibbs sampler with applications to the comparison of estimators and augmentation schemes. *Biometrika*, 81:27–40, 1994.
- [5] J. J. O. Ruanaidh and W. J. Fitzgerald. *Numerical Bayesian Methods Applied to Signal Processing*. Springer-Verlag, Springer Series in Statistics and Computing, 1996.

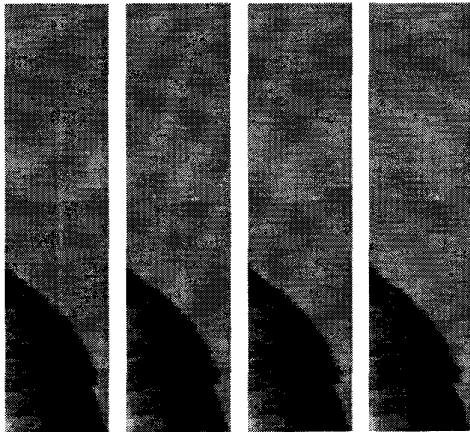


Figure 1. Left to right: Portion of original KNIGHT image; Result of subtracting estimated line profile; Least Squares Interpolation of line region; Sampled Interpolation in line region.

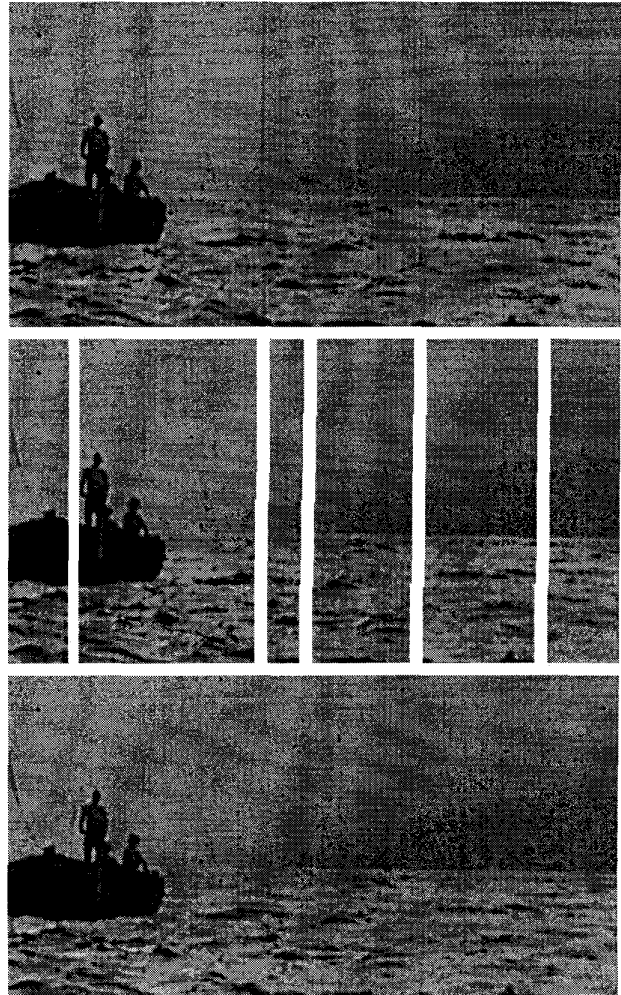


Figure 2. Reconstruction of artefacts in the SIT-DOWN sequence, using sampled 2D AR interpolation with an 8 tap causal model. From top to bottom: Original image ( $256 \times 512$  portion); Location of detected lines in white; Interpolation.