

JOINT INTERPOLATION, MOTION AND PARAMETER ESTIMATION FOR IMAGE SEQUENCES WITH MISSING DATA

Simon J. Godsill and Anil C. Kokaram

Signal Processing and Communications Laboratory, University of Cambridge
Cambridge CB2 1PZ, England

Tel: +44 1223 332767; fax: +44 1223 332662

e-mail: {s.jg,ack}@eng.cam.ac.uk

ABSTRACT

This paper presents a new scheme for interpolation of missing data in image sequences, an important problem in many areas including archived motion picture film and digital video. A unified framework for image data modelling and motion estimation is adopted which is based on 3-dimensional autoregressive (3DAR) models with motion correction. A fully Bayesian methodology is implemented using the Gibbs Sampler, a method which allows for joint estimation with respect to all of the unknowns, including the motion field.

1 INTRODUCTION

Missing data occurs in many areas involving image sequences, including archived motion picture film, high speed photography and digital video. The problem is of particular importance to companies such as Kodak and Quantel who expend much effort in manual correction of defects in film material.

Previous approaches to the problem have employed a two stage process involving motion estimation followed by interpolation [1]. These schemes perform well in general but can suffer from errors in motion estimation in the presence of corrupted data. Recent work [2] has gone some way towards correcting this problem by introducing a probabilistic motion field correction which takes into account the corrupted data. In this paper a new method is presented which employs a Bayesian methodology to solve for both the image data and motion estimate at the same time.

A Markov Chain Monte Carlo (MCMC) method called the Gibbs Sampler is used for numerical implementation of the scheme in which all system parameters (including motion vectors, AR coefficients and missing data) are treated as unknown *a priori*.

An example of the method applied to a corrupted movie image sequence demonstrates the high quality of restoration which is possible.

2 3DAR MODEL

The motion-corrected 3D-AR model with coefficients a_k (for $k = 1 \dots P$) can be represented as

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

where $I(\mathbf{x})$ is the pixel intensity at location $\mathbf{x} = [x, y, n]$ (i.e. co-ordinate $[x, y]$ in the n th frame of the sequence). $\mathbf{q}_k = [qx_k, qy_k, qn_k]$ is the 3-dimensional AR support corresponding to coefficient a_k . This support is constrained to be causal for this work. $\mathbf{d}_k = [dx_k, dy_k, 0]$ is the motion offset between the current frame (n) and the frame of support vector \mathbf{q}_k (i.e. $(n - qn_k)$). Note that the motion offset is required only to pixel accuracy in this model. $\epsilon(\mathbf{x})$ is the excitation sequence, assumed here to be white and Gaussian with variance σ_e^2 .

3 GIBBS SAMPLER

We wish to reconstruct a set of missing pels $\mathcal{U} = \{\mathbf{x}_i, i = 1 \dots l\}$. Consider a vector of image intensities \mathbf{i} which includes at least the pels in \mathcal{U} and their immediate 3DAR support. The Bayesian framework (see e.g. [3]) will perform inference using the marginal posterior distribution for the missing pels $\mathbf{i}_{(x)}$ conditional upon the remaining pels in \mathbf{i} , denoted by $\mathbf{i}_{-(x)}$:

$$p(\mathbf{i}_{(x)} | \mathbf{i}_{-(x)}) \propto \int_{\{\mathbf{a}, \sigma_e^2, \mathbf{d}\}} p(\mathbf{i} | \mathbf{a}, \sigma_e^2, \mathbf{d}) p(\mathbf{a}, \sigma_e^2, \mathbf{d}) d\{\mathbf{a}, \sigma_e^2, \mathbf{d}\} \quad (2)$$

In this expression, which is obtained directly from Bayes Rule and the marginalization identity, the term $p(\mathbf{i} | \mathbf{a}, \sigma_e^2, \mathbf{d})$ is the likelihood for the block of image pels and $p(\mathbf{a}, \sigma_e^2, \mathbf{d})$ is the prior distribution for the unknown parameters. Note that \mathbf{d} contains the motion vectors required to form all the \mathbf{d}_k 's in (1) for a given \mathbf{i} .

For the prior distributions used here it is possible to obtain the marginal distribution (2) in analytical form. However, the resulting distribution is not straightforward to deal with, involving a summation over all candidate motion vectors amongst other complications. Instead we use the Gibbs Sampler [4, 5], a MCMC scheme

which generates a sequence of (dependent) samples from the required marginal distribution. These samples can then be used to obtain Monte Carlo estimates for the missing data.

The Gibbs Sampler requires the full conditional densities for all unknowns, in this case $\{\mathbf{i}_{(x)}, \mathbf{a}, \sigma_e^2, \mathbf{d}\}$. The basic form of the algorithm may then be summarized as:

1. Make an initial guess: $\{\mathbf{i}_{(x)}^0, \mathbf{a}^0, \sigma_e^{20}, \mathbf{d}^0\}$.
2. for $i=1$ to N {
 - (a) $\mathbf{i}_{(x)}^i \sim p(\mathbf{i}_{(x)} | \mathbf{a}^{i-1}, \sigma_e^{2i-1}, \mathbf{d}^{i-1}, \mathbf{i}_{-(x)})$
 - (b) $\mathbf{a}^i \sim p(\mathbf{a} | \mathbf{i}_{(x)}^i, \sigma_e^{2i-1}, \mathbf{d}^{i-1}, \mathbf{i}_{-(x)})$
 - (c) $\sigma_e^{2i} \sim p(\sigma_e^2 | \mathbf{i}_{(x)}^i, \mathbf{a}^i, \mathbf{d}^{i-1}, \mathbf{i}_{-(x)})$
 - (d) $\mathbf{d}^i \sim p(\mathbf{d} | \mathbf{i}_{(x)}^i, \mathbf{a}^i, \sigma_e^{2i}, \mathbf{i}_{-(x)})$

where ‘ \sim ’ denotes drawing a random sample from the distribution to the right. A Monte Carlo estimate may then be made for the missing data using the samples obtained following convergence of the algorithm, e.g. the minimum mean-squared error (MMSE) estimate is the posterior mean, estimated as

$$\hat{\mathbf{i}}_{(x)}^{\text{MMSE}} = \frac{1}{N - M} \sum_{i=M+1}^N \mathbf{i}_{(x)}^i \quad (3)$$

where M is the number of iterations until convergence, the ‘burn-in’ time.

3.1 Priors

We assume a non-informative prior framework [3] for the model parameters, $p(\mathbf{a}, \sigma_e^2, \mathbf{d}) \propto 1/\sigma_e^2$. This assumes that all candidate motion vectors are equiprobable *a priori*. A more sophisticated framework might seek to regularize the motion field in the manner of [2] by assigning a Gaussian random field (GRF) prior which favours motion vectors similar to those in adjacent areas of the image. The modifications required to achieve this are straightforward.

Within such a prior framework the conditional distributions required in step 2. (above) can be formulated directly, see appendix A.

3.2 Modified Gibbs Sampler

The sampling scheme outlined above is likely to converge slowly, due in part to the fact that there is strong interdependence between the unknowns $\mathbf{i}_{(x)}$, \mathbf{a} , σ_e^2 and \mathbf{d} . The convergence can be improved by sampling from several of the dependent unknowns jointly (see [6, 7] for some relevant discussion). There are two possibilities which lead to a practical sampling scheme. Steps 2.(a)-(d) are then replaced by either

- (a) $\mathbf{i}_{(x)}^i \sim p(\mathbf{i}_{(x)} | \mathbf{a}^{i-1}, \sigma_e^{2i-1}, \mathbf{d}^{i-1}, \mathbf{i}_{-(x)})$
- (b) $\{\mathbf{a}^i, \sigma_e^{2i}, \mathbf{d}^i\} \sim p(\mathbf{a}, \sigma_e^2, \mathbf{d} | \mathbf{i}_{(x)}^i, \mathbf{i}_{-(x)})$

or

- (a) $\mathbf{a}^i \sim p(\mathbf{a} | \mathbf{i}_{(x)}^{i-1}, \sigma_e^{2i-1}, \mathbf{d}^{i-1}, \mathbf{i}_{-(x)})$
- (b) $\{\mathbf{i}_{(x)}^i, \sigma_e^{2i}, \mathbf{d}^i\} \sim p(\mathbf{i}_{(x)}, \sigma_e^2, \mathbf{d} | \mathbf{a}^i, \mathbf{i}_{-(x)})$

Appendix B gives the details of these steps. Note that each scheme has its own merits and both might be alternated within the same run to improve convergence properties further.

4 EXAMPLE

The method has been tested using both standard image sequences and real movie footage, giving very high quality results. Here an example is presented for a movie image sequence. One frame from the degraded original frame is shown as figure 1. A standard detection scheme (SDIa, see [1]) flags a large region of pels in the area of the hair fringe as corrupted (highlighted in white in figure 2, in which the relevant area is zoomed in). Figure 3 shows the MMSE estimate obtained using equation (3), while figure 4 shows the final sampled realization \mathbf{i}^N extracted from the Markov chain. Both have performed well, but close examination shows that the sampled interpolation is perceptually better, retaining more of the detailed graininess of the surrounding image. This may be explained by noting that the sampled interpolation is a typical realization for the missing data and thus captures some of the randomness inherent in the AR process (see [2] for further discussion). Implementational details were as follows. The causal 3D AR support used 5 pels from the previous frame and 3 from the current frame. A 50-square block of image pels in the current and previous frame was used to construct \mathbf{i} . The Markov chain was run for $N = 100$ iterations with a burn-in of $M = 10$. Both figures are conservative as the chain was informally observed to converge within 5 iterations on most images.

5 COMMENT

Results obtained so far match or improve upon other methods familiar to the authors. This is as a result of the full probabilistic modelling of all unknowns within a single framework. The possibility for using sampled interpolations rather than MAP or MMSE estimates provides a novel deviation from standard techniques. Further work will extend the framework to allow for occlusion and uncovering of objects in the previous frame as well as automatic detection of corrupted pels.

A Sampling from the conditional distributions

The approximate AR likelihood for image data \mathbf{i} is given by:

$$p(\mathbf{i} | \mathbf{a}, \sigma_e^2, \mathbf{d}) \approx (2\pi\sigma_e^2)^{-N/2} \exp(-E(\mathbf{a}, \mathbf{i}, \mathbf{d})/2\sigma_e^2)$$

where $E(\mathbf{a}, \mathbf{i}, \mathbf{d}) = \sum_{\mathbf{x}} \epsilon(\mathbf{x})^2$ and $\epsilon(\mathbf{x})$ is the excitation obtained from the model equation (1). The summation

is over all pels \mathbf{x} whose 3DAR support lies within \mathbf{i} and N is the total number of such points.

Using similar methods and notation as [2, 1] we form a vector \mathbf{e} containing all the excitation elements within the summation and express in two alternative forms:

$$\mathbf{e} = \mathbf{i} - \mathbf{X}\mathbf{a} = \mathbf{A}_{(x)}\mathbf{i}_{(x)} + \mathbf{A}_{-(x)}\mathbf{i}_{-(x)} \quad (4)$$

Note here that \mathbf{X} depends only upon \mathbf{d} and \mathbf{i} , while both \mathbf{A} -matrices depend only upon \mathbf{d} and \mathbf{a} .

The joint probability of data and parameters is the product of likelihood and prior:

$$p(\mathbf{i}, \mathbf{a}, \sigma_e^2, \mathbf{d}) = p(\mathbf{i}|\mathbf{a}, \sigma_e^2, \mathbf{d})p(\mathbf{a}, \sigma_e^2, \mathbf{d})$$

The conditionals are then obtained by rearranging this expression. Take, for example, the missing data $\mathbf{i}_{(x)}$:

$$p(\mathbf{i}_{(x)}|\mathbf{i}_{-(x)}, \mathbf{a}, \sigma_e^2, \mathbf{d}) = \frac{p(\mathbf{i}, \mathbf{a}, \sigma_e^2, \mathbf{d})}{p(\mathbf{i}_{-(x)}, \mathbf{a}, \sigma_e^2, \mathbf{d})}$$

But the denominator here is constant for any given $\{\mathbf{i}_{-(x)}, \mathbf{a}, \sigma_e^2, \mathbf{d}\}$, so the conditional may be found by rearranging the joint distribution in terms of $\mathbf{i}_{(x)}$ and determining the normalizing constant of the distribution. Following this procedure for each unknown in turn gives

$$\begin{aligned} p(\mathbf{i}_{(x)}|\mathbf{i}_{-(x)}, \mathbf{a}, \sigma_e^2, \mathbf{d}) &= N_l(\hat{\mathbf{i}}_{(x)}, \sigma_e^2(\mathbf{A}_{(x)}^T \mathbf{A}_{(x)})^{-1}) \\ p(\mathbf{a}|\mathbf{i}, \sigma_e^2, \mathbf{d}) &= N_P(\hat{\mathbf{a}}, \sigma_e^2(\mathbf{X}^T \mathbf{X})^{-1}) \quad (5) \\ p(\sigma_e^2|\mathbf{i}, \mathbf{a}, \mathbf{d}) &= \text{IG}(N/2, E(\mathbf{a}, \mathbf{i}, \mathbf{d})/2) \\ p(\mathbf{d}|\sigma_e^2, \mathbf{i}, \mathbf{a}) &\propto p(\mathbf{d}) \exp(-E(\mathbf{a}, \mathbf{i}, \mathbf{d})/2\sigma_e^2) \end{aligned}$$

where $\hat{\mathbf{a}} = (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{i}$,
 $\hat{\mathbf{i}}_{(x)} = -(\mathbf{A}_{(x)}^T \mathbf{A}_{(x)})^{-1} \mathbf{A}_{(x)}^T \mathbf{A}_{-(x)} \mathbf{i}_{-(x)}$

Here $N_l(\boldsymbol{\mu}, \mathbf{C})$ denotes the l -dimensional Gaussian with mean vector $\boldsymbol{\mu}$ and covariance matrix \mathbf{C} ; $\text{IG}(\alpha, \beta) \propto u^{-(\alpha+1)} \exp(-\beta/u)$ is the inverted-gamma distribution (see e.g. [8]);

Drawing samples from the multivariate Gaussian and inverted-gamma distributions are standard procedures (see e.g. [9]). The only complication is the discrete variable \mathbf{d} , whose normalizing constant is only available through direct summation over all \mathbf{d} . If \mathbf{d} is constrained to lie within some small grid of displacements then this will be a feasible sampling strategy. If, however, large motion shifts are anticipated then other strategies must be adopted. One possible solution which uses a more general form of the MCMC method will be presented in a future publication. For now we assume the former strategy with small displacements relative to the starting guess \mathbf{d}^0 .

B Joint sampling

Consider the first strategy, i.e. sampling jointly from $p(\mathbf{a}, \sigma_e^2, \mathbf{d}|\mathbf{i})$. This may be achieved by making the factorization

$$p(\mathbf{a}, \sigma_e^2, \mathbf{d}|\mathbf{i}) = p(\mathbf{a}|\sigma_e^2, \mathbf{d}, \mathbf{i})p(\sigma_e^2|\mathbf{d}, \mathbf{i})p(\mathbf{d}|\mathbf{i}) \quad (6)$$

The joint sample is then obtained by sampling from these three distributions in turn:

$$\mathbf{d}^i \sim p(\mathbf{d}|\mathbf{i}^i), \quad \sigma_e^{2i} \sim p(\sigma_e^2|\mathbf{d}^i, \mathbf{i}^i), \quad \mathbf{a}^i \sim p(\mathbf{a}|\sigma_e^{2i}, \mathbf{d}^i, \mathbf{i}^i)$$

The third term here is simply the full conditional for \mathbf{a} , derived as (5). The second and third terms are obtained by successive integrations of the joint conditional (6) with respect to \mathbf{a} and σ_e^2 using standard integral identities obtained from the Gaussian and gamma distributions, respectively, to give:

$$\begin{aligned} p(\sigma_e^2|\mathbf{d}, \mathbf{i}) &= \text{IG}((N-P)/2, E(\hat{\mathbf{a}}, \mathbf{i}, \mathbf{d})/2) \\ p(\mathbf{d}|\mathbf{i}) &\propto \frac{E(\hat{\mathbf{a}}, \mathbf{i}, \mathbf{d})^{(N-P)/2}}{|\mathbf{X}^T \mathbf{X}|^{1/2}} \end{aligned}$$

Similar working applies to the second scheme for sampling $\{\mathbf{i}_{(x)}^i, \sigma_e^{2i}, \mathbf{d}^i\}$, although the details are omitted here.

References

- [1] A. Kokaram, R. Morris, W. Fitzgerald, and P. Rayner. Detection/interpolation of missing data in image sequences.. *IEEE Image Processing*, pages 1496–1519, Nov. 1995.
- [2] A. Kokaram and S. Godsill. A system for reconstruction of missing data ima image sequences using sampled 3D AR models and MRF motion priors. In *Proc. European Conference on Computer Vision*, pages 613–624, April 1996.
- [3] G. E. P. Box and G. C. Tiao. *Bayesian Inference in Statistical Analysis*. Addison-Wesley, 1973.
- [4] S. Geman and D. Geman. Stochastic relaxation, Gibbs distributions and the Bayesian restoration of images. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 6:721–741, 1984.
- [5] A. E. Gelfand and A. F. M. Smith. Sampling-based approaches to calculating marginal densities. *J. Am. Statist. Assoc.*, 85:398–409, 1990.
- [6] S. J. Godsill and P. J. W. Rayner. Robust treatment of impulsive noise in speech and audio signals. In *IMS Lecture Note Series in assoc. with Second International Workshop on Bayesian Robustness*, Rimini, Italy, June 1995. (*To appear*).
- [7] C. K. Carter and R. Kohn. On Gibbs Sampling for state space models. *Biometrika*, 81(3):541–553, 1994.
- [8] N. L. Johnson and S. Kotz. *Distributions in Statistics: Continuous Univariate Distributions*. Wiley, 1970.
- [9] W. H. Press, B. P. Flannery, S. A. Teukolsky, and W. T. Vetterling. *Numerical Recipes in C, Second Edition*. CUP, 1992.



Figure 1: Original image (256×256).

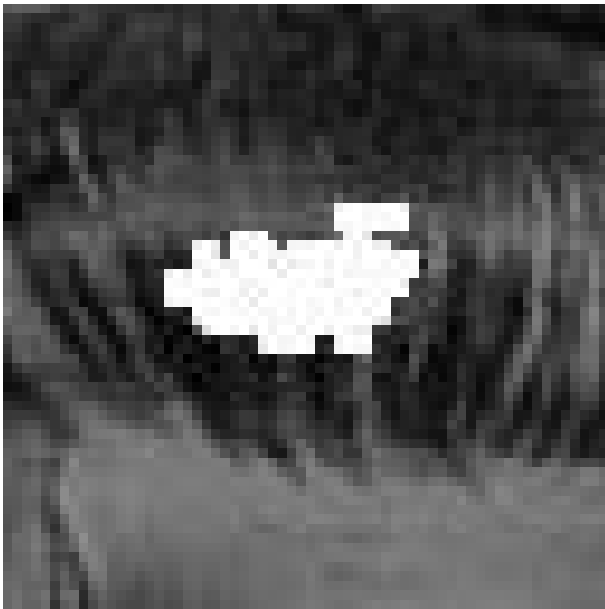


Figure 2: Pixels detected as missing data by the SDIa detector shown *bright white* (zoom to fringe of hair).



Figure 3: Interpolation using Gibbs MMSE estimate (zoom).



Figure 4: Sampled Interpolation from the Gibbs scheme (zoom).