A Bayesian approach to simultaneous autofocus and super-resolution

Richard O. Lane, Keith D. Copsey and Andrew R. Webb

QinetiQ Ltd, St Andrews Road, Great Malvern, Worcestershire, UK, WR14 3PS

ABSTRACT

This paper presents a numerical Bayesian approach to the autofocus and super-resolution of targets in radar imagery. An ill-posed inverse problem is studied in which the known linear imaging operator is subject to an unknown degree of distortion (defocusing). The goal is simultaneously to reconstruct a high-resolution representation of a target based on noisy lower resolution image measurements and to estimate the degree of defocus. We present a Markov chain Monte Carlo algorithm for parameter estimation, illustrate the approach on an explanatory example and compare our technique with a maximum likelihood approach. Given a model for the sensor measurement process, this technique may be applied to any type of radar image such as those produced by a synthetic aperture radar (SAR), inverse SAR (ISAR) or a real beam imaging radar. The proposed approach fits into a larger set of procedures aiming to exploit targeting information from different radar sensors.

Keywords: Radar image restoration, Bayesian methods, autofocus, super-resolution, Markov chain Monte Carlo

1. INTRODUCTION

A wide variety of radar sensors are available to image vehicles or structures of interest. Automatic target recognition (ATR) systems typically use a set of target data recorded using a certain sensor to train a classifier, which can then be used to recognise targets in unseen data. For good classifier performance, the unseen data must have similar characteristics to the training set. However, if the unseen data comes from a different radar sensor with a different resolution, the target data from the two sensors are not directly comparable – it is necessary to transform from one sensor representation to another. One approach is to determine an underlying high-resolution representation. ATR algorithms could then be trained and operated in this high-resolution data space independent of sensor type.

The cross-range resolution of a noncoherent radar is approximately $R\lambda/D$, where R is the distance from the radar to the target, λ is the wavelength, and D is the antenna aperture dimension. Thus, the cross-range resolution of a noncoherent radar is limited by the physical size of the antenna: the larger the antenna, the better the resolving power. In synthetic aperture radar (SAR) the radar is mounted on a platform, flown along a specified path and successive pulses are coherently combined. This improves the cross-range resolution of the radar[1] to $R\lambda/(2L_{SA})$ for stripmap SAR with a synthetic aperture length L_{SA} or $\lambda/(2\psi)$ for spotlight SAR with a maximum squint angle of ψ.

Super-resolution is the use of signal processing techniques to increase the resolution beyond the physical limits discussed in the previous paragraph by exploiting knowledge of the antenna radiation pattern or point spread function (PSF). A feature of super-resolution techniques is that the PSF must accurately be known for the results to be reliable. In the case of a noncoherent radar the PSF is nominally constant – it does not vary with platform motion. The PSF may be estimated to a reasonable degree of accuracy by measuring the response of the radar to a point scatterer, under controlled conditions. Richards[2] has proposed a constrained iterative

Further author information:
R.O.L.: E-mail: rlane1@qinetiq.com, Telephone: +44(0) 1684 896954
K.D.C.: E-mail: kkcopsey@qinetiq.com, Telephone: +44(0) 1684 896926
A.R.W.: E-mail: arwebb@qinetiq.com, Telephone: Tel: +44(0) 1684 894490

Algorithms for Synthetic Aperture Radar XI, E.G.Zelnio, F.D.Garber, Eds., Proc. SPIE Vol. 5427 April 2004

deconvolution approach to solving the noncoherent super-resolution problem. However, in SAR signal processing the precise PSF is dependent on the relative motion between the radar and the target. Accurate knowledge of this motion is therefore required before super-resolution can be performed.

Usually for airborne SAR, the radar is assumed to travel in a perfectly straight line at constant velocity. However, any deviations from this assumed trajectory will introduce phase errors in the complex data resulting in a defocused image. This is equivalent to a broadening of the cross-range PSF. Sensors on board the aircraft, such as an inertial navigation sensor, can be used measure the motion of the aircraft and focus the system, but the information provided by these sensors is not necessarily available with sufficient accuracy. Under these conditions the system must be focused by estimating phase errors from the data itself – *i.e.* autofocus.

There are several autofocus techniques, which work in a variety of ways. Contrast maximisation[3] assumes a model for platform motion and varies parameters of the model until the image is focused. Other techniques use entropy as a focus criterion[4] instead of contrast. Multilook registration[3] examines the difference in Doppler shift between two successive image swaths to give an estimate of cross-track acceleration and then corrects for this. The Wigner-Ville transform has also been used for a time-frequency analysis of the problem.[5] A widely used and powerful method is phase gradient autofocus (PGA),[6] which is robust because it is not dependent on a specific motion model. It is not clear with any of these techniques, however, precisely how the PSF and hence image of a target changes with minor variations in defocus parameters.

The down-range resolution of a radar must also be considered. The resolution is limited to $c/(2B)$, where B is the bandwith of the transmitted pulse and c is the speed of radio wave propagation.[1] The higher the bandwidth the better the resolving power of the radar. However, the resolution does not vary with platform motion and in principle down-range super-resolution could be performed without autofocus. In this paper we present cross-range super-resolution only, the extension to two-dimensional super-resolution incorporating down-range information is obvious. Also, super-resolution for noncoherent radar may be considered a special case of our method, where the defocus parameters of the PSF are known.

In addition to the above considerations there may be non-linear effects due to self-shadowing of the target, which would interfere with the super-resolution technique. In practice, other physical processes such as multi-path and multi-bounce could also occur. These issues are not addressed in this paper.

The problem of simultaneous autofocus and super-resolution is an inverse one. When the imaging process is well understood it is relatively simple to apply an imaging operator to the high-resolution representation of a target to obtain the equivalent lower resolution radar image. However, obtaining this representation, given an image, is not trivial and there may be many possible realisations that give rise to the same image. A Bayesian approach to the inversion process provides a principled way to recover an estimate of the high-resolution target signature. Bayesian methods are probabilistic, able to cope with the possibility of multiple solutions and can readily incorporate other available information as prior knowledge.

The rest of the paper is organised as follows. Section 2 outlines the model used to describe the scattering and imaging processes. Section 3 describes the Metropolis-Hastings algorithm, which is used to perform Bayesian autofocus and super-resolution. Results are presented in section 4 and compared with a maximum likelihood technique using an explanatory synthetic data example. Finally, conclusions are drawn in section 5.

2. THE MODEL

Due to the nature of the radar scattering process a target will reflect energy by different amounts depending on the geometry of the imaging platform and target, and the frequency of the radar. Each type of target is assumed to have a two-dimensional plan view scattering cross section per unit area. This cross section is a function of imaging geometry and frequency; however the dependency is omitted in this paper for clarity.

Following Luttrell,[7] a diagram of the overall model is shown in Figure 1, using the notation

$$\begin{aligned} \sigma &= \text{scattering cross section} \\ f &= \text{scattered field} \\ \theta &= \text{defocus parameters} \\ g &= \text{image.} \end{aligned}$$

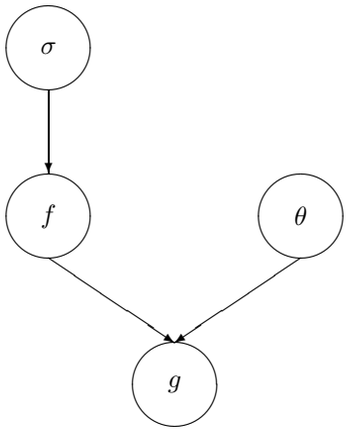

Figure 1. Overall model

The scattering model generates a field f given the cross section σ. The imaging model generates an image g given f and the defocus parameters θ. These two models are now described in more detail.

2.1. Scattering model

The scattered field f is generated according to a scattering model $p(f|\sigma)$. It is assumed here that each cross section element σ_i produces a large number of wavelets that combine coherently to produce an element of complex scattered field f_i. If there are a sufficient number of wavelets then the central limit theorem applies and we may assume Gaussian statistics *. For m cross sections

$$\begin{aligned} p(f|\sigma) &= \prod_{i=1}^{m} \frac{\exp(-|f_i|^2/\sigma_i)}{\pi\sigma_i} \\ &= \frac{\exp(-f^\dagger\sigma^{-1}f)}{\det(\pi\sigma)}, \end{aligned} \tag{1}$$

where

$$\sigma \equiv \begin{pmatrix} \sigma_1 & 0 & \dots & 0 \\ 0 & \sigma_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \sigma_m \end{pmatrix} \tag{2}$$

$$f^\dagger\sigma^{-1}f \equiv \sum_{i,j=1}^{m} f_i^*(\sigma^{-1})_{ij}f_j. \tag{3}$$

Note that f is a complex vector, which explains the unusual normalisation in equation 1 and also that $f^\dagger$ represents the Hermitian (complex conjugate) transpose $f^\dagger = f^{T*}$.

2.2. Imaging model

The radar image may be thought of as being produced by a point spread function applied to each scattered field element and sampled at the image resolution with the addition of thermal noise. If the bandwidth, temperature and noise figure of the radar transceiver are known then the noise power ν can be calculated. For uncorrelated thermal noise, the noise covariance matrix N is diagonal: $N = \nu I$.

*This is an alternative formulation of the case 1 and case 2 Swerling fluctuation models. Other more general models, such as the gamma distribution are possible.[8]

The imaging model is defined as

$$p(g|f,\theta) = \frac{\exp\left[-(g-T(\theta)f)^\dagger N^{-1}(g-T(\theta)f)\right]}{\det(\pi N)}, \tag{4}$$

where $T(\theta)$ is a matrix with each row containing a shifted version of the point spread function. θ represents defocus parameters that alter the PSF – a scalar θ could represent cross-track acceleration, for example. It is important to note here that the numerical Bayesian technique described in the next section is fully able to cope with non-linear and non-Gaussian systems by replacing equation (4) with the appropriate form.

3. THE BAYESIAN SOLUTION

The Bayesian approach to autofocus and super-resolution is a probabilistic way of modelling uncertainty in the cross section σ and defocus parameters θ. It is possible that different combinations of σ and θ could give rise to the same image due to the interaction of several scatterers in a single pixel, defocus and the addition of noise. This uncertainty is described by the joint probability density $p(\sigma,\theta|g)$ of the cross section and defocus parameters, conditional on the image under consideration. Bayes' theorem gives the joint density as

$$p(\sigma,\theta|g) = \frac{p(g|\sigma,\theta)p(\sigma)p(\theta)}{p(g)}. \tag{5}$$

Under certain simple models it may be possible to calculate the density $p(\sigma,\theta|g)$ analytically.[9] However, in general this will not be the case as the form of the point spread function could be non-trivial and indeed non-linear.[10] For the general case a numerical technique such as Markov chain Monte Carlo (MCMC) is required. The algorithm used here is the Metropolis-Hastings (M-H) algorithm.[11]

The M-H algorithm is an iterative method for generating samples of a probability distribution. For the case considered here the samples represent the joint probability density $p(\sigma,\theta|g)$. The likelihood of the image is given by

$$\begin{aligned} p(g|\sigma,\theta) &= \int p(g|f,\sigma,\theta)p(f|\sigma,\theta)df \\ &= \int p(g|f,\theta)p(f|\sigma)df, \end{aligned} \tag{6}$$

where we have noted from Figure 1 that g does not depend directly on σ, and f does not depend on θ. With the linear convolution model described in the previous section this can be shown to be[7]

$$p(g|\sigma,\theta) = \frac{\exp(-g^\dagger M^{-1}g)}{\det(\pi M)}, \tag{7}$$

where

$$M \equiv T\sigma T^\dagger + N. \tag{8}$$

One advantage of the M-H algorithm is that it is necessary only to know the shape of the distribution – there is no need to calculate the normalising factor $p(g)$. With choice of suitable priors $p(\sigma)$ and $p(\theta)$ for the cross section and defocus parameters the quantity of interest is then

$$\pi(\sigma,\theta|g) = p(g|\sigma,\theta)p(\sigma)p(\theta). \tag{9}$$

At each iteration it is possible to update one variable at a time or all variables in one go. Here we update each element of σ in turn and then perform a θ update. During the σ update at the ith iteration, a proposed new sample for a single cross section element is generated from a proposal distribution $q(\sigma^{i+1}|\sigma^i)$. The proposal distribution may take a wide variety of forms, each having their advantages and disadvantages – see Ref. 11 for details. The proposed sample is accepted with a probability $\alpha(\sigma^i,\sigma^{i+1})$, where

$$\alpha(\sigma^i, \sigma^{i+1}) = \min\left[\frac{\pi(\sigma^{i+1}, \theta^i|g)q(\sigma^i|\sigma^{i+1})}{\pi(\sigma^i, \theta^i|g)q(\sigma^{i+1}|\sigma^i)}, 1\right]. \tag{10}$$

For ease of notation we have omitted the dependence on cross section element number so that while updating the jth element of σ we have in fact

$$\sigma^i = \left[\sigma_1^{i+1}, \sigma_2^{i+1}, ..., \sigma_{j-1}^{i+1}, \sigma_j^i, \sigma_{j+1}^i, ...\sigma_m^i\right]^T \tag{11}$$

and

$$\sigma^{i+1} = \left[\sigma_1^{i+1}, \sigma_2^{i+1}, ..., \sigma_{j-1}^{i+1}, \sigma_j^{i+1}, \sigma_{j+1}^i, ...\sigma_m^i\right]^T. \tag{12}$$

In other words, at each step a new sample is generated; if it is more likely (including the effect of the proposal distribution) than the current one it is always accepted but less likely ones are also accepted with a certain probability. This avoids the problem of getting trapped in local maxima and is analogous to simulated annealing[12] †. If the proposed sample is rejected then the current sample is used in the next iteration step. The same process is then repeated for θ with

$$\alpha(\theta^i, \theta^{i+1}) = \min\left[\frac{\pi(\sigma^{i+1}, \theta^{i+1}|g)q'(\theta^i|\theta^{i+1})}{\pi(\sigma^{i+1}, \theta^i|g)q'(\theta^{i+1}|\theta^i)}, 1\right]. \tag{13}$$

If θ is multidimensional then θ^i and θ^{i+1} follow the same notation convention as equations 11 and 12.

Initial samples generated during the so-called burn-in period depend on the starting position and must be discarded if they are not reasonable values of the distribution. The remaining samples are distributed from $p(\sigma, \theta|g)$ as required.

4. A NUMERICAL EXAMPLE

4.1. Constant acceleration model

The model studied here is that of Luttrell[7] and is equivalent to the radar platform undergoing constant cross-track acceleration in proportion to a scalar θ. The point spread function is given by

$$\begin{aligned} T(\theta) &= \frac{1}{2c}\int_{-c}^{+c} \exp(ikx + i\theta k^2 x^2)dk \\ &\approx \frac{1}{2c}\int_{-c}^{+c} \exp(ikx)(1 + i\theta k^2 x^2)dk \\ &= T_0 + \theta T_1, \end{aligned} \tag{14}$$

where

$$\begin{aligned} T_0 &\equiv \frac{\sin(cx)}{cx} \\ T_1 &\equiv i\left[cx\sin(cx) + 2\cos(cx) - \frac{2\sin(cx)}{cx}\right]. \end{aligned} \tag{15}$$

To accommodate higher order motion one would replace the quadratic in the first line of equation 14 with a polynomial of the appropriate degree.

† However, it should be emphasised that simulated annealing is a technique to find the global optimum of a function (*i.e.* a single value), whereas here we are trying to find a probability density and the methods do not "converge" in the same manner.

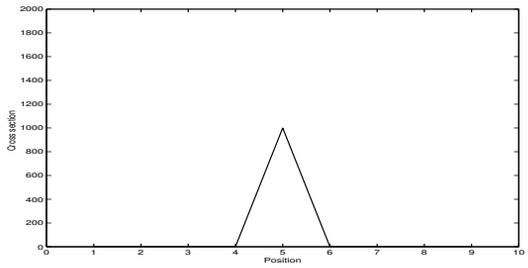

Figure 2. Single point scatterer scene

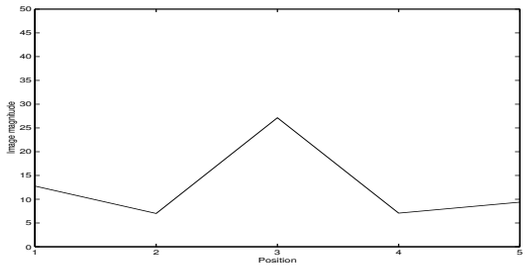

Figure 3. Image of single point scatterer

4.2. Implementation details

The Metropolis-Hastings algorithm described in the previous section was applied to the synthetic data example used by Luttrell[7] to demonstrate autofocus/super-resolution, which was generated under the constant acceleration model. For comparison, Luttrell's results are also reproduced here independently. The data consisted of an idealised point target embedded in nine pixels of clutter with a signal-to-clutter ratio of 30dB. The scattered field was generated according to (1). The point spread function in (14) was then applied to the field with the resultant image sampled at double the spacing used to sample the cross section, to give five image pixels. Gaussian noise was added to the image at a signal-to-noise ratio of 20dB. The original cross section is shown in Figure 2 and the magnitude of the resulting complex image is shown in Figure 3.

A uniform prior distribution was used for θ and all the cross section elements. The σ proposal distribution was a uniform random walk centred on the current sample with a width that achieved a 90% acceptance rate. The uniform distribution was used to aid calculation of the proposoal ratio – if a Gausian is used extra computing power has to be expended on calculating the error function at every iteration, because negative cross section values are not permitted. The θ proposal distribution was a Gaussian random walk centred on the current sample, with a width that achieved a 75% acceptance rate. The random walk width was not found to be critical, Ref. 11 contains a discussion on random walk proposal distribution choice. The initial σ sample was set to an interpolated version of the image. This cross section value was considered a reasonable value to have in the probability distibution so no burn-in period was required. The initial θ value was set to zero and burn-in was not an issue in this case either. One million samples were produced and of these one in every 100 was used for subsequent analysis – *i.e.* a total of 10,000. The main reason for performing this step was to reduce the computer memory and processing requirements during results analysis. Also, the original set of samples are correlated and leaving a "decorrelation gap" reduces this effect.

In an intial run the sample with the highest probability was selected as an estimate of the maximum *a posteriori* (MAP) solution. Ten more runs of the algorithm were performed, each time all the variables were fixed to the MAP estimate except one. This enabled the formation of a histogram for the variable under test, which is a one-dimensional slice through the multivariate distribution characterised by the samples in the initial run.

4.3. Results – single point scatterer

A histogram of samples representing the fifth cross section element, corresponding to the point scatterer, is shown in Figure 4. The true scatterer cross section value in arbitrary units is 1000; it can be seen from the figure that the peak is close to this value, but the distribution is skewed and there are a significant number of high cross section values. A histogram of samples for the first clutter cross section element is shown in Figure 5. The peak of the histogram is at a cross section value of 1 and the histogram has a long tail, although not as "fat" as the one for the point scatterer cross section element - note the different scales on the two graphs. The histograms of the other clutter elements were similar to the first. A graph of the MAP solution with error bars showing where the probability density falls to half its maximum value is shown in Figure 6. The MAP solution is close to the true cross section (Figure 2) but the error bars show considerable variability.

Luttrell's maximum likelihood solution is shown in Figure 7. It can be seen that this solution tends to agree with the MAP solution shown in Figure 6. Convergence of the σ values over time is shown for the point

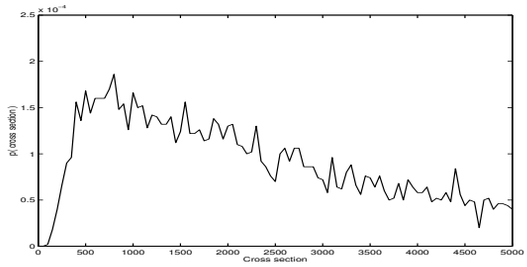

Figure 4. Histogram for σ (point scatterer)

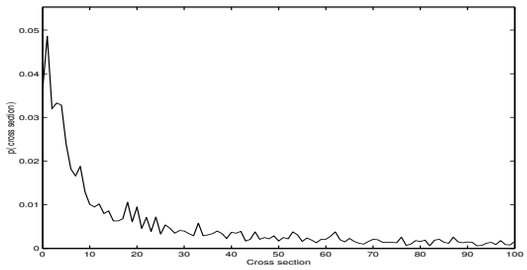

Figure 5. Histogram for σ (clutter)

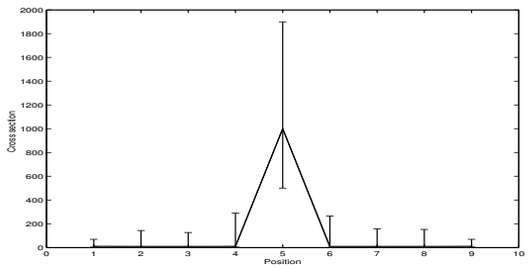

Figure 6. M-H Solution for σ – single scatterer

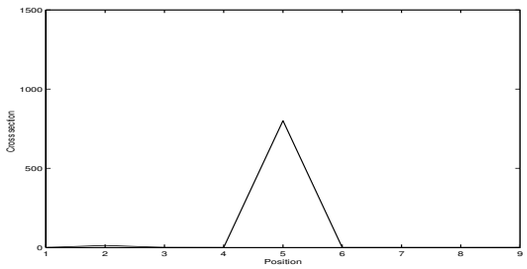

Figure 7. Luttrell solution for σ – single scatterer

scatterer element and first clutter element in Figure 8. Both cross section elements converge to a stable value after about 120 iterations. Convergence graphs for the other clutter elements were similar to the first clutter element, although they did not reach as high values in the intermediate 10-40 iterations.

Figure 9 shows the histogram for θ obtained with the M-H algorithm. The shape of the histogram is approximately Gaussian with a standard deviation of about 0.01. The peak of the distribution is approximately equal to the true value of 0.1. Figure 10 shows how the Luttrell algorithm converges to a θ solution over time with the true value denoted by a dotted line. After about 120 iterations the algorithm has converged to a stable value of 0.075. This is an underestimate of the true value.

The advantage of the full Bayesian approach over a maximum likelihood approach is apparent – there are many cross section values that have a reasonable probability associated with them, an observation not possible with the maximum likelihood solution. In addition, more complicated targets are likely to introduce local maxima in the posterior distribution and the maximimum likelihood approach is sensitive to local maxima.

4.4. Results – two scatterers

The experiment was repeated for a two scatterer case. The scene containing two scatterers is shown in Figure 12. The image of the scene, which has a single peak, is shown in Figure 13. Note that it is not possible to tell from the image alone that two scatterers are present.

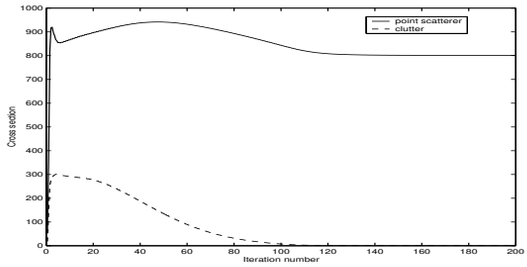

Figure 8. Convergence of σ in Luttrell algorithm

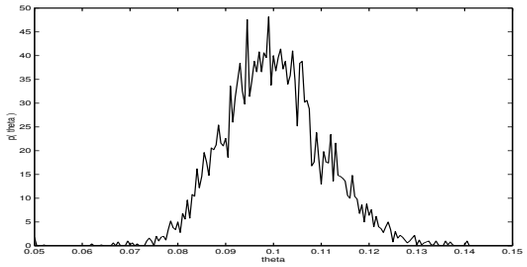

Figure 9. Histogram for θ

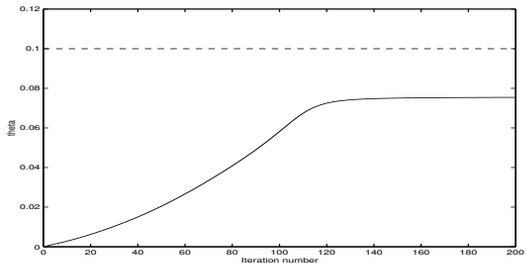

Figure 10. Convergence of θ in Luttrell algorithm

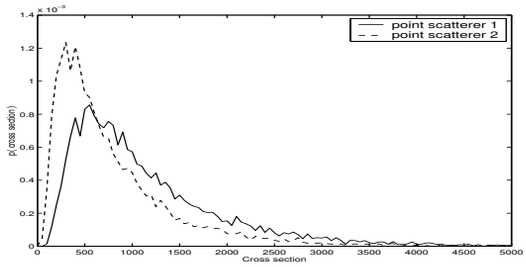

Figure 11. Histogram for σ – two scatterers

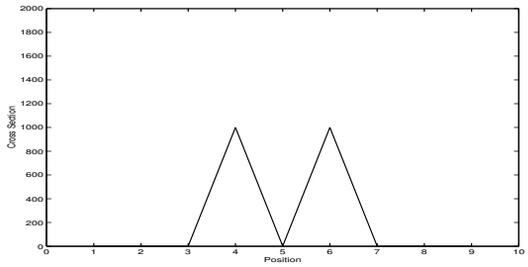

Figure 12. Scene containing two point scatterers

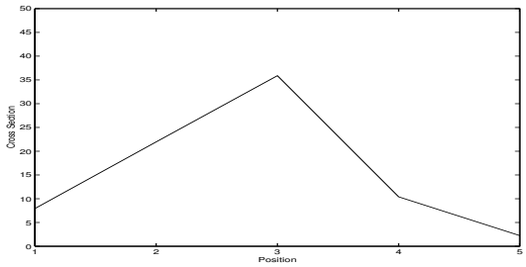

Figure 13. Image of two point scatterers

Histograms for the two point scatterers are shown in Figure 11. These are similar in shape to the single point scatterer case (Figure 4). The peak of the first scatterer is at a higher cross section value than the second. This is because the field generated in this particular example had a low magnitude at the second scatterer. The ability of the field to flutuate in this way is a consequence of the Gaussian scattering model. The MAP solution for the whole scene containing two scatterers is shown with error bars in Figure 14. For comparison the Luttrell solution is shown in Figure 15, which again agrees well with the MAP solution. The θ solution was similar, using both methods, to the single scatterer case.

4.5. Discussion

It was found with initial runs of the M-H algorithm that when all cross section elements were allowed to vary simultaneously, while using a uniform prior for σ, the tail of the point scatterer cross section element histogram was very large. The algorithm spent a long time generating samples in this tail and spent little time at more "sensible" values. This affected the histogram of the other cross section elements and θ. One way to stop this was to introduce a decaying exponential prior for the cross section. There was no need to adjust the θ prior in this case but this procedure tended to bias the peak of the cross section histogram in favour of small values. The best way to avoid this problem seemed to be a method whereby cross section values were limited to a range spanning the expected histogram peak in an intial run before being allowed to fully explore the distribution when only one variable at a time was allowed to change.

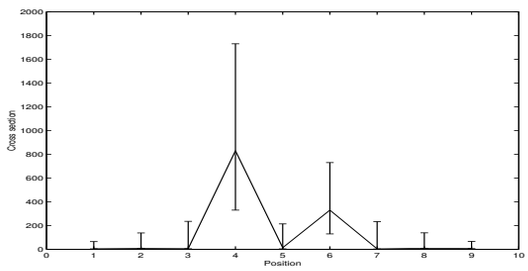

Figure 14. M-H solution for σ – two scatterers

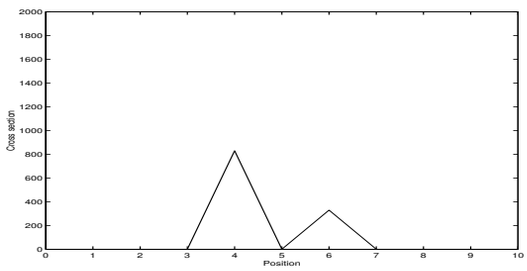

Figure 15. Luttrell solution for σ – two scatterers

The large histogram tail demonstrates the ill-posedness of the problem. The data do not contain sufficient information to estimate parameters of interest with a great degree of precision. If several independent images were to be used then the problem would not be so ill-posed, the tail would naturally fall of more rapidly and the peak would be sharper.

Luttrell found no need to introduce prior information but in a sense that solution was fortunate because the scattered field was similar in form to the original cross section. In general, it is quite possible that the field could have a different form to the cross section. If this were the case then the peak of the histogram would not correspond to the true cross section value and a single maximum likelihood solution would be inappropriate – a situation evident in the two scatterer example. However, the correct value should have a reasonable probability because the histogram is wide. Any use of good prior information would help move the peak towards reasonable values.

5. CONCLUSION

A Bayesian algorithm for simultaneously super-resolving targets in radar imagery and estimating defocus parameters introduced by non-ideal motion of the radar platform has been implemented. The algorithm is based on a numerical Markov chain Monte Carlo technique and produces a series of samples that characterise the probability distribution of the target scattering cross section and defocus parameters. A problem involving synthetic data was considered where an idealised point target was imaged using a linear system with additive Gaussian noise. The cross section distribution was found to be broad and highly skewed but the most probable value was close to the true value. The distribution of the defocus parameter was found to be approximately Gaussian, centred on the true value and with a small spread. A second problem involving two point scatterers was considered. The algorithm was able successfully to discern the presence of two scatterers even though there was a only single peak in the image.

The whole target cross section distribution could be passed on to an automatic target recognition algorithm, which has been trained in a high-resolution data space – *i.e.* data with a resolution independent of the imaging sensor. Each of the samples would be classified by target type and the results combined to give a value of confidence in the classification decision. Thus this technique potentially enables the classification of targets imaged with one radar, when training was performed with a different type of radar of a different resolution. In addition, if it is assumed that the frequencies of two radar sensors are sufficiently close (*i.e.* less than an order of magnitude different) and both sensors exploit the optical scattering regime (as opposed to Rayleigh or resonant scattering), comparisons from a wider set of radars could potentially be made.

Acknowledgements

This work was sponsored by the UK Ministry of Defence Corporate Research Programme "Communications, Information and Signal Processing". The first author would also like to acknowledge support of the EPSRC Engineering Doctorate programme.

REFERENCES

1. C. J. Oliver, "Synthetic-aperture radar imaging," *J. Phys. D: Appl. Phys.* **22**, pp. 871–890, 1989.
2. M. A. Richards, "Iterative noncoherent angular superresolution," *Proceedings of the IEEE National Radar Conference* , pp. 100–105, April 1988.
3. D. Blacknell, A. P. Blake, C. J. Oliver, and R. G. White, "A comparison of SAR multilook registration and contrast optimisation autofocus algorithms applied to real SAR data," *International Radar Conference* , pp. 363–366, October 1992.
4. R. L. Morrison and D. C. Munson, "An experimental study of a new entropy-based SAR autofocus technique," *Proc. of the Int. Conf. on Image Processing* **2**, pp. 441–444, September 2002.
5. T. P. Leonard and M. G. Holden, "The Wigner-Ville transform for the focussing of ISAR images," *IEE Colloquium on Time-Scale and Time-Frequency Analysis and Applications* , February 2000.

6. D. E. Wahl, P. H. Eichel, D. C. Ghiglia, and C. V. Jakowatz, "Phase gradient autofocus - a robust tool for high resolution SAR phase correction," *IEEE Transactions on Aerospace and electronic systems* **30**, pp. 827–835, July 1994.
7. S. P. Luttrell, "A Bayesian derivation of an iterative autofocus/superresolution algorithm," *Inverse Problems* **6**, pp. 975–996, 1990.
8. M. I. Skolnik, *Introduction to Radar Systems*, McGraw-Hill Book Company, 2nd ed., 1980.
9. K. D. Copsey and A. R. Webb, "Classifier design for population and sensor drift," *Submission to IAPR Workshop on Statistical Pattern Recognition, Lisbon, Portugal* , February 2004.
10. S. P. Luttrell, "The complex point spread function of the RSRE SAR," *RSRE Memorandum (Unpublished)* **4079**, September 1987.
11. S. Chib and E. Greenberg, "Understanding the Metropolis-Hastings algorithm," *The American Statistician* **49**(4), pp. 327–335, 1995.
12. W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes in C*, Cambridge University Press, 2nd ed., 1992.