# Non-parametric Bayesian super-resolution

R. O. Lane, QinetiQ, St. Andrews Road, Great Malvern, Worcestershire, United Kingdom, WR14 3PS

rlane1@QinetiQ.com

**Abstract**: Super-resolution of signals and images can improve the automatic detection and recognition of objects of interest. However, the uncertainty associated with this process is not often taken into consideration. This is important because the processing of noisy signals can result in spurious estimates of the scene content. This paper reviews a variety of super-resolution techniques and presents two non-parametric Bayesian super-resolution algorithms that not only take uncertainty into account, but also retain knowledge about the output uncertainty in the form of a full probability distribution. One of the two Bayesian techniques is based on an analytical calculation re-interpreted as super-resolution, and the other is a novel numerical algorithm. Although the algorithms are presented as stand-alone techniques for image analysis, such Bayesian super-resolution algorithms can increase automatic target recognition performance over standard super-resolution.

**Keywords**: Analytic distribution, automatic target recognition (ATR), Bayesian super-resolution, least squares (LS), Markov chain Monte Carlo (MCMC), minimum mean-square error (MMSE), non-parametric models, numerical distribution, point spread function (PSF), radar.

# 1 Introduction

When first introduced, radar systems had low cross-range and range resolutions, and were used primarily for the detection of targets and the estimation of target positions. Modern radar systems, such as synthetic-aperture radar (SAR), have high resolutions of one metre or smaller [1]. With high-resolution radar it becomes possible not only to estimate an object's location but also to recognize the general class of object or type within a class. When target recognition systems employ multiple radar sensors, it is advantageous to directly compare images from the various sensors as images can provide more information than higher level features alone. However, this could prove to be difficult in practice because the images from different sensors have different resolutions.

Rayleigh or classical resolution is defined as the minimum separation required for two closely-spaced equal scattering centres to be distinguished. This is a fundamental limit based on the bandwidth of the transmitted waveform and the physical size of the radar system compared to the transmitted radio wavelength. One way of making signals from different sensors directly comparable is to use super-resolution for the sensors with a poor resolution. Super-resolution is the use of signal processing techniques in high signal-to-noise scenarios to increase the resolution beyond classical limits by using knowledge of a system's point-target response, and by making assumptions about the scene of interest. Once signals from all the sensors have been processed to the same resolution using super-resolution, data from the various sensors can be compared directly and standard pattern recognition techniques can be applied. It should be noted that the resolution improvement factor of super-resolution systems is dependent on the signal-to-noise ratio and it may not be possible to provide a resolution match in cases where the sensors' original resolutions are widely different. However, regardless of a mismatch in resolution, super-resolution is often said to enhance detection and recognition performance

[2][3][4][5][6][7][8][9], and is therefore a technique that is of interest to signal and information practitioners.

A good way of combining data from different sources is to use Bayesian statistics. The main motivation behind a Bayesian approach lies in the unique ability of Bayesian statistics to handle uncertain and possibly conflicting pieces of information in a fully consistent manner [10]. In particular, Bayesian theory provides a consistent mechanism for manipulating probabilities assigned to data. Further advantages in the use of Bayesian techniques include the ability to cope with additional prior information and the production of confidence intervals and other statistics of interest for the parameters estimated. Therefore the combination of Bayesian statistics with super-resolution provides a sound basis for carrying out target recognition using a variety of sensors.

This paper presents Bayesian super-resolution as one processing step that could be used as part of a wider information processing framework that incorporates uncertainty about knowledge of a scene. Although described in terms of radar signal processing and motivated by a radar application, the techniques could be applied to other forms of data that follow the same model, such as found in optical image reconstruction, astronomy, ultrasound, mass spectrometry, or medical image analysis. The remainder of the paper is organised as follows. Section 2 establishes super-resolution concepts and outlines some basic algorithms. Section 3 introduces two new Bayesian super-resolution algorithms based on analytic and numerical solutions to the scene estimation problem. Section 4 presents results of the super-resolution algorithms. Finally, conclusions are drawn and recommendations for further work are made in section 5.

# 2 Super-resolution Theory

## *2.1 Introduction to Super-resolution*

Super-resolution may be defined as the use of signal processing techniques to improve a system's resolution beyond the classical limit. A received signal is considered to be the result of a convolution between a point spread function and a high-resolution representation of the scene of interest. Deconvolution removes the effect of the point spread function and reveals the high-resolution scene. Therefore all deconvolution algorithms are super-resolution algorithms. In radar or communications systems using an array of antennas, weights can be applied to signals received at each antenna to precisely steer nulls in the direction of a nearby strong signal. This allows the direction and power of a second signal to be determined without power leaking from the nulled signal. Since the two targets can be closer together than the Rayleigh limit, this form of processing – sometimes referred to as direction finding – can also be considered to be super-resolution. Radar data is often obtained in the frequency domain and an inverse Fourier transform is applied to obtain data in the spatial domain, which is easier to interpret than frequency data for detection and recognition purposes. The Fourier transform of an ideal point target is a complex sinusoid with a frequency that depends on the position of the target. Therefore any technique that is able to estimate the parameters of multiple sinusoids closely spaced in frequency is equivalent to a super-resolution technique. Accordingly the areas of super-resolution, deconvolution, direction finding, and estimation of sinusoids, are considered to be similar problems.

There are two major classes of super-resolution model. The first class describes the scene as a finite number of point scatterers that can take any position in the scene, and the received continuous radar signal is sampled for digital processing. This is sometimes known as a high-level model [11], the scattering centre model [12], or the parametric model [7]. The advantage of this model is that major

isolated scatterers in the scene, which can often account for the majority of received energy at the radar, are well modelled because of their positional accuracy. The disadvantage is that the number of scatterers in the scene must be estimated as well as their position and amplitude. Without a regularization procedure, algorithms based on this model create a large number of scatterers with small amplitude, which are generally related to noise rather than true structure in the scene. Another disadvantage of this model is that a linear superposition of perfect point scatters may not completely describe a complex target. Examples of algorithms that use the scattering centre model are MUSIC, ESPRIT, RELAX, and IMP. A detailed review of these and other similar techniques is given in [13].

The second class of model assumes the scene is a continuously varying high-resolution function, and during digital processing we consider samples of this function on a regular grid. It is sometimes known as a low-level model [11], the continuum scattering model [12], or the non-parametric model [7]. This model has the advantage that extended targets are better modelled and the number of parameters to be estimated is fixed according to the sample spacing and size of the scene. The disadvantage is that isolated strong scatterers are less well modelled if they are not positioned at a scene sample point. In a similar manner to the first class of model, when noise is present, algorithms will tend to give non-zero amplitude to areas of the scene with no valid target, unless a regularization procedure is used. Since fixed-dimensional algorithms (where the number of parameters to be estimated is known) are easier to implement in a Bayesian framework than trans-dimensional ones, the remainder of this paper concentrates on the non-parametric super-resolution model. The potential for trans-dimensional Bayesian techniques is discussed in section 5.

## *2.2 Non-parametric Super-resolution*

In the non-parametric model, the scene of interest is assumed to have a high-resolution back-scattering coefficient or distributed radar cross section (RCS), which is a function of measurement

geometry. In this context, high-resolution is defined as a signal whose power spectrum has a high bandwidth. For generality, the following discussion refers to images, in which case the distributed RCS is a two-dimensional quantity. However, the same arguments apply to one-dimensional range or cross-range profiles. When imaged by a coherent radar the distributed RCS gives rise to a complex scattered field value. An RCS-to-field scattering model is described in [13]. The image formation process is equivalent to the convolution of a point spread function (PSF) and the high-resolution complex field representation of the target. This effectively places an amplitude-scaled copy of the PSF at each location on a sampled grid. A bulk phase is present in the complex image, which is due to the distance between the radar and the target. In addition to this, the various scattering mechanisms on the target have a phase relative to each other dependent on their precise position and construction, and thus coherently sum to produce a single complex number at each image sampling point. The imaging process is described mathematically by

$$\mathbf{g} = \mathbf{T}\mathbf{f} + \mathbf{n}\,, \tag{1}$$

where $\mathbf{f}$ is a complex vector denoting the raster-scanned high-resolution 2D target representation, $\mathbf{g}$ is a complex vector representing the resultant raster-scanned low-resolution image, and $\mathbf{T}$ is an appropriately formatted Toeplitz convolution matrix that applies the effect of the PSF [13]. A Toeplitz matrix $\mathbf{T}$ has structure such that its *ij*th element $T_{ij}$ is a function of only (*i*-*j*) and thus has identical elements along its main diagonal and sub-diagonals. Clutter is implicitly included in the scene representation through the use of non-zero complex values at positions in the vector $\mathbf{f}$ corresponding to spatial positions near a target but not actually within the target boundary. Thermal noise in the radar receiver is modelled by $\mathbf{n}$, which is a zero-mean circularly complex Gaussian random variable with a diagonal covariance matrix $\mathbf{N}$.

There are many non-parametric super-resolution algorithms in the literature. In the remainder of this section two particular algorithms will be discussed: least squares and minimum mean-square error. Bayesian super-resolution is discussed in more detail in section 3. Other popular super-resolution algorithms not detailed here include singular value decomposition (SVD) [14], super spatially-variant apodization [15], Capon's maximum-likelihood method [16], the amplitude and phase estimation of a sinusoid (APES) algorithm [17], high-definition vector imaging (HDVI) [18], and maximum entropy [19]. A review of these and other non-parametric super-resolution algorithms is given in [13].

## *2.3 Least Squares*

The benchmark solution against which other non-parametric super-resolution algorithms are often compared is the Moore-Penrose matrix pseudoinverse. This is a well-known least-squares (LS) procedure for solving systems of linear equations and is the basis for super-resolution algorithms such as [20]. The approach minimizes, with respect to the recovered scene $\mathbf{f}$, a cost function

$$J = \left\| \mathbf{g} - \mathbf{T}\mathbf{f} \right\|^2 \quad , \tag{4}$$

which is the sum of squared differences between the measured image $\mathbf{g}$ and the noiseless image $\mathbf{T}\mathbf{f}$ that would be produced given the solution $\mathbf{f}$. Differentiation of the cost function with respect to $\mathbf{f}$ and setting the result equal to zero gives the estimate for $\mathbf{g}$ as

$$\hat{\mathbf{f}}_{inv} = \left( \mathbf{T}^H \mathbf{T} \right)^{-1} \mathbf{T}^H \mathbf{g} \, . \tag{5}$$

The $H$ superscript denotes the Hermitian (complex conjugate) transpose of a matrix. The least-squares solution gives good results for general well-conditioned algebraic problems of the form $\mathbf{g} = \mathbf{T}\mathbf{f}$, where small changes due to noise in the measured data $\mathbf{g}$ result in small changes in the solution $\mathbf{f}$. However, the specific super-resolution problem we are interested in is ill-posed as stated due to the structure of the PSF convolution matrix. Small changes in $\mathbf{g}$ result in large changes in $\mathbf{f}$. Super-resolution

performance using the pseudo-inverse algorithm is therefore usually poor due to the ill-conditioned nature of the problem – a large number of spurious scatterers are often placed at positions where there is no valid target. A general discussion of ill-posed and inverse problems is given in [21] and aspects of their applicability to radar super-resolution are discussed in [22].

## 2.4 Minimum Mean-Square Error

A similar approach to least squares is the minimum mean-square error (MMSE) technique. The object of the MMSE approach is to choose the linear operator $\mathbf{R}$, such that the super-resolution solution given by $\hat{\mathbf{f}}_{\mathbf{mmse}} = \mathbf{Rg}$ minimizes the expected norm $J$ of the reconstruction error:

$$J = \left\langle \left\| \hat{\mathbf{f}}_{\mathbf{mmse}} - \mathbf{f} \right\|^2 \right\rangle . \tag{6}$$

If the *a priori* statistical distribution of the estimate of $\mathbf{f}$ is a zero mean Gaussian distribution with covariance matrix $\mathbf{W}$, then the solution is given by [23] as

$$\hat{\mathbf{f}}_{\mathbf{mmse}} = \mathbf{W}\mathbf{T}^H \left( \mathbf{T}\mathbf{W}\mathbf{T}^H + \mathbf{N} \right)^{-1} \mathbf{g} . \tag{7}$$

In this basic form when $\mathbf{W}$ is fixed in advance, the MMSE solution is also known as a Wiener filter. However, in practice $\mathbf{W}$ is not known in advance because prior knowledge of the scene is not well defined. Therefore an iterative scheme that estimates the $\mathbf{W}$ matrix from measured data must be used [24]. In one scheme, $\mathbf{W}$ is set to the identity matrix for the first iteration and in subsequent iterations it is estimated from the $\hat{\mathbf{f}}$ of the previous iteration. Any element of $\hat{\mathbf{f}}$ whose power is below a threshold is assumed to be homogeneous clutter and the combined variance of all these elements is calculated and entered into the appropriate diagonal elements of $\mathbf{W}$. Any element $i$ above the threshold has its variance set to $\hat{f}_i^2$ and entered into the appropriate diagonal element of $\mathbf{W}$. Therefore $\mathbf{W}$ is always diagonal and has a high variance where there is a valid target or a large scatterer. The procedure is

repeated by alternately estimating $\mathbf{W}$ and $\hat{\mathbf{f}}$ until a termination criterion is met when $\hat{\mathbf{f}}$ changes by less than a specified amount between iterations. This particular algorithm is referred to as thresholded minimum mean-square error (MMSE-T) super-resolution. Another algorithm that works in a similar manner to MMSE-T, by partitioning the data into target and clutter pixels, is semi-sparse MMSE [25]. However, that algorithm requires prior knowledge of target and clutter statistics. The dual-model super-resolution technique of [26] also divides the data into target and clutter regions. An alternative formulation of MMSE super-resolution is given in [27] and is categorized as adaptive pulse compression.

If $\mathbf{W}$ is fixed to be the identity matrix and the noise level is set to zero then a single iteration of the MMSE algorithm is identical to the least squares solution. This demonstrates a failing of the LS algorithm – it effectively assumes no noise is present in the measured data, over-fits the data, and results in a very noisy output. In the MMSE algorithm, the noise covariance matrix $\mathbf{N}$ acts as a regularization parameter, allowing there to be a difference between the measured data and modelled data commensurate with the noise level. A disadvantage of the basic MMSE algorithm is the requirement to know the noise power in advance but there are a number of ways to estimate it. One of these is to make a measurement of an area in the scene where there are known to be no targets or clutter present. This could be done by selecting a shadowed area of an image and calculating the variance of the pixels in that area. However, it may be difficult to automatically segment the image into shadow and non-shadow areas. A different way of obtaining noise-only data would be to have the radar point at an area of the sky where there are no targets. Alternatively, rather than measuring it, the noise power could be calculated directly from the formula $kTB_nF_n$, where $k$ is Boltzmann's constant, $T$ is the temperature of the radar receiver, $B_n$ is the noise-equivalent bandwidth and $F_n$ is the noise factor. It would be advantageous if the noise power could be estimated without image segmentation or the need

to make independent measurements because these operations introduce extra complexity into the system. One procedure for automatic noise power estimation integrated with the MMSE technique is given in [28].

# 3 Bayesian super-resolution

## *3.1 Introduction*

Several non-parametric Bayesian super-resolution algorithms for both radar and non-radar applications have been proposed in the literature. One approach in [29] is to find the maximum of the target cross-section posterior probability density function using an expectation maximization algorithm. In the particular problem studied there, the possibility that the image under consideration is defocused is also taken into account. The defocusing is represented as known and unknown parts of the point spread function. This provides some level of robustness with respect to the fact that the PSF is known only to a certain accuracy. In [19] maximum entropy is used to choose the prior probability distributions in a Bayesian model and the conjugate gradient method is used to maximize the posterior probability. In [30] it is stated that standard image restoration using a Tikhonov prior produces overly-smooth solutions, therefore an edge-preserving function and hyper-parameters are introduced in the form of prior information. In [31] an image restoration algorithm is applied to emission computed tomography data. The model follows Poisson statistics and uses a neighbourhood function, which takes into account correlations between adjoining pixels. As with [30], the algorithm simultaneously estimates hyper-parameters and the restored image. The advantage of this approach over other Bayesian methods is that it includes uncertainty in estimates of the hyper-parameter values. In [32] variational methods are used to approximate the true Bayesian posterior distribution in medical and industrial deconvolution problems. This procedure converts an intractable analytic solution to a

tractable one. Although the above algorithms are formulated in a Bayesian framework they often give the mean or most likely image rather than analysing the full uncertainty as is done by the algorithms introduced in this paper.

Recently, there has been a surge of interest in multi-frame image super-resolution in which a group of low-resolution images of a scene are combined to produce a high-resolution image of that scene [33]. Many multi-frame super-resolution algorithms are based on Bayesian statistics. In [34], for example, a Gaussian Markov random field is used as a prior density for the high-resolution image to incorporate knowledge that adjacent pixels in the image are correlated. The maximum *a posteriori* (MAP) solution is then found using a Markov chain Monte Carlo (MCMC) approach based on the Gibbs sampler. The approach is extended in [35], where outlier-sensitive bilateral filtering is used as a post-processing method to suppress image artefacts introduced during super-resolution. In [36] the low-resolution image registration parameters are considered as nuisance parameters and are marginalized to leave a function that can be optimized with respect to the high-resolution image. This procedure outperforms methods where the registration parameters are estimated via a MAP method and are fixed before the high-resolution scene is estimated. Although multi-frame image super-resolution algorithms are powerful, the radar target recognition application considered in this paper would have available only a single image, with a high dynamic range, with which to perform super-resolution. Therefore the advantages of these algorithms, such as the use of multiple frames and a limited dynamic range, are not applicable to this work.

## *3.2 Imaging Model*

The imaging model used for Bayesian super-resolution is the same as that introduced in section 2.2, whereby the image is considered as the convolution of a point spread function with a high-resolution scene representation and the addition of Gaussian noise. However, under a Bayesian model

all information must be represented by a probability density function (PDF). In this case, the PDF of the image **g** is dependent on the target representation **f** and is written as $p(\mathbf{g}|\mathbf{f})$. The imaging model is then defined as

$$p(\mathbf{g} \mid \mathbf{f}) = \frac{\exp\left[-(\mathbf{g}-\mathbf{T}\mathbf{f})^H \mathbf{N}^{-1}(\mathbf{g}-\mathbf{T}\mathbf{f})\right]}{\det(\pi\mathbf{N})}, \tag{9}$$

where the terms are defined in section 2.2. It is important to note here that the numerical Bayesian technique described in section 3.4 is able to cope with non-linear and non-Gaussian systems by replacing equation (9) with the appropriate form.

### *3.3 Analytic Bayesian Solution*

The Bayesian approach to super-resolution is a probabilistic way of modelling uncertainty in our knowledge of the true high-resolution target representation **f**. It is possible that different combinations of **f** and noise could give rise to the same image due to the interaction of several elements of **f** in a resolution width. This uncertainty is described by the probability density $p(\mathbf{f}|\mathbf{g})$ of the target representation, conditional on the image under consideration. Bayes' theorem gives the posterior density as

$$p(\mathbf{f} \mid \mathbf{g}) = \frac{p(\mathbf{g} \mid \mathbf{f})p(\mathbf{f})}{p(\mathbf{g})}. \tag{10}$$

Note that under the model used here the true value of **f** is deterministic, and the expressions $p(\mathbf{f})$ and $p(\mathbf{f}|\mathbf{g})$ respectively refer to the distribution of the estimate of **f** before and after the data have been considered.

If the prior density of **f** is a zero-mean multivariate Gaussian defined by

$$p(\mathbf{f}) = \frac{\exp\left(-\mathbf{f}^H \mathbf{W}^{-1}\mathbf{f}\right)}{\det(\pi\mathbf{W})}, \tag{11}$$

then under this model it is possible to calculate the density $p(\mathbf{f}|\mathbf{g})$ analytically. The solution is also a multivariate Gaussian distribution and is given by [29] as

$$p(\mathbf{f} \mid \mathbf{g}) = \frac{\exp\left[-(\mathbf{f}-\bar{\mathbf{f}})^H \mathbf{C}^{-1}(\mathbf{f}-\bar{\mathbf{f}})\right]}{\det(\pi \mathbf{C})}, \tag{12}$$

where

$$\bar{\mathbf{f}} = \mathbf{C}\mathbf{T}^H \mathbf{N}^{-1}\mathbf{g}, \tag{13}$$

is the mean of the posterior distribution and

$$\mathbf{C}^{-1} = \mathbf{W}^{-1} + \mathbf{T}^H \mathbf{N}^{-1}\mathbf{T} \tag{14}$$

is the inverse covariance matrix. With algebraic manipulation the mean can be re-written as $\bar{\mathbf{f}} = \mathbf{W}\mathbf{T}^H\left(\mathbf{T}\mathbf{W}\mathbf{T}^H + \mathbf{N}\right)^{-1}\mathbf{g}$, which is in fact the same as the basic form of the MMSE solution given in Section 2.4. However, the Bayesian solution provides more information than MMSE in the form of the covariance matrix, which measures uncertainty in the recovered scene.

### *3.4 Numerical Bayesian Solution*

In general, for a non-Gaussian prior distribution or more complicated models, it will not be possible to derive a simple analytic solution similar to that of the previous section. Calculation of the normalization constant $p(\mathbf{g})$ for the posterior distribution in equation (10) is usually not tractable. For most physical and processing models, statistics of interest such as the mean and covariance will not be available analytically either. In such cases, rather than making simplifications to allow analytic inference on the posterior distribution, sophisticated statistical models can be maintained by drawing samples from the posterior distribution $p(\mathbf{f}|\mathbf{g})$. All inferences can then be made through consideration of these samples. In most circumstances it is not possible to sample directly from the posterior distribution, therefore a Markov chain Monte Carlo (MCMC) algorithm is used. The particular algorithm used here is the Metropolis-Hastings (M-H) algorithm [37]. Note that for the Gaussian

imaging model considered in this paper it is unnecessary to use this sampling approach as the analytic solution has already been found – see equation (12). However, the analytic solution allows a comparison with the MCMC approach and aids understanding of the M-H algorithm output. A more advanced scattering and imaging model are used in [13] and [38] where an analytic solution is not available.

The Metropolis-Hastings algorithm is an iterative method for generating samples of a probability distribution. For the case considered here the samples represent the probability density $p(\mathbf{f}|\mathbf{g})$. One advantage of the algorithm is that it is necessary only to know the shape of the distribution $p(\mathbf{f}|\mathbf{g})$ – there is no need to calculate the normalizing factor $p(\mathbf{g})$. The likelihood of the image $p(\mathbf{g}|\mathbf{f})$ was given by equation (9). With choice of a suitable priors for $p(\mathbf{f})$ the quantity of interest is then

$$\pi(\mathbf{f} \mid \mathbf{g}) = p(\mathbf{g} \mid \mathbf{f})\, p(\mathbf{f}) . \tag{15}$$

Each element of the estimated vector $\mathbf{f}$ is considered to be a separate variable, although element values may be correlated. During the update at the $i$th iteration, a proposed new sample for a single element of $\mathbf{f}$ is generated from a proposal distribution $q(\mathbf{f}^{i+1}|\mathbf{f}^{i})$. The proposal distribution may take a wide variety of forms, each having their advantages and disadvantages as discussed in [37]. The proposed sample is accepted with probability

$$\alpha(\mathbf{f}^{i}, \mathbf{f}^{i+1}) = \min\left[1, \frac{\pi\left(\mathbf{f}^{i+1} \mid \mathbf{g}\right) q\left(\mathbf{f}^{i} \mid \mathbf{f}^{i+1}\right)}{\pi\left(\mathbf{f}^{i} \mid \mathbf{g}\right) q\left(\mathbf{f}^{i+1} \mid \mathbf{f}^{i}\right)}\right]. \tag{16}$$

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 sample it is always accepted but less likely samples are also accepted with a certain probability. This avoids the problem of getting trapped in local maxima – the purpose of the algorithm is to explore the entire distribution rather than find a single optimal value. If the proposed sample is rejected then the current sample is copied and used in the next iteration step. Initial samples generated depend on the starting position and must be discarded if they are not

reasonable values of the distribution. These samples form what is known as the bum-in period. The remaining samples are distributed as $p(\mathbf{f}|\mathbf{g})$, as required.

It has been assumed in the above discussion that the variance of the thermal noise is known. If desired, it is possible to include the noise variance as an additional unknown parameter to be estimated as part of the Bayesian framework [39]. Alternatively a heuristic method for estimating the noise variance, such as that suggested in [28], could be used.

# 4 Results

In this section we present the results of Bayesian super-resolution when applied to a simulated data set. A small one-dimensional example has been chosen to illustrate the potential of the super-resolution technique in an easily understood manner. In practice, the size of measured data sets would be larger than that discussed here.

The simulated scene consists of two complex-valued point targets in uncorrelated clutter at a signal-to-clutter ratio of 20 dB. The two targets are in phase, each with a phase of 45 degrees with respect to an arbitrary reference; the clutter has random phase. The power profile of the scene is shown in Figure 1. A point spread function was applied to the scene and thermal noise added at a signal-to-noise ratio (SNR) of 20 dB. The resulting low-resolution signal is shown in Figure 2. The point spread function was a modified sinc function that has previously been used in [29] to demonstrate super-resolution algorithms in systems where there is some level of defocus. The continuous form of the function is $T(x) = T_0(x) + \theta T_1(x)$, where

$$T_0(x) = \frac{\sin(cx)}{cx}, \tag{17}$$

$$T_1(x) = i\left[ cx\sin(cx) + 2\cos(cx) - \frac{2\sin(cx)}{cx} \right]. \tag{18}$$

The focus parameter $\theta$ was set to a value of 0.1, which is roughly equivalent to an uncompensated quadratic phase error of $4 \times 10^{-3}$ $ms^{-2}$ for typical airborne radar system parameters [13][40]. In the simulations carried out for this paper, the parameter $c$ is arbitrary as it scales the $x$ axis. The noisy image, point spread function, and SNR were used as inputs to the Bayesian super-resolution algorithms. The prior covariance matrix used was a diagonal matrix with all elements set to 1000.

The power range profile derived from the mean of the analytic solution is shown in Figure 3. It can be seen that there are some differences between the estimated mean profile and the true scene shown in Figure 1. The existence of the two targets at positions 7 and 11 has been correctly identified and their power is in the correct order of magnitude. However, the mean estimated scene profile demonstrates bias in the complex-field estimates, which is introduced by the noise in the image. The sample values at positions 4, 5 and 6 all have high mean powers, when ideally they should be close to zero. Not much can be done to alleviate this problem – noise places an ultimate limit on super-resolution performance.

Also shown in Figure 3 is the analytically-derived variance as a function of position. Most positions have a low-to-medium variance but the variance of the centre position is very high. To analyze all information about uncertainty in the recovered scene the full covariance must be examined. A graphical representation of this matrix is shown in Figure 4, where the real parts of the $ij$th elements of the matrix are displayed. The variance graph in Figure 3 is a plot of the diagonal values of this matrix. Most values of the matrix are near to zero, which implies a low correlation between the scene values at widely separated parts of the image. However, there is a strong negative correlation between adjacent positions, evident from the dark squares on the sub-diagonals situated one position away from the main diagonal. This is to be expected because the effect of two adjacent scene values on the image tends to get blurred out by the point spread function. Thus when either the first scene value is high and the second is low or vice versa a similar image will result. However, if both scene values are high or

low the image would look different. The negative correlation is especially strong between the central position and its two adjacent positions. This provides an explanation of why the variance of the central position is so high: whether the central value is high or low, its effect is minimized by changes in the value of adjacent positions that result in a modelled image consistent with the measured input image. Note from equation (14) that the posterior covariance matrix is completely determined by parameters known before the image is observed: the prior scene covariance, the noise covariance, and the point spread function. This means the true scene has no influence on the posterior covariance matrix. Since the prior and noise covariance matrices are uniform over the scene, it is the action of the point spread function that creates a high posterior variance at the central position.

The sample-based mean and covariance of the numerical Monte Carlo Bayesian super-resolution algorithm are very similar to those calculated using the analytic algorithm. The MMSE-T algorithm output is also similar to the mean analytic estimated profile. These results are not shown here for reasons of space.

A graphical comparison of the distribution given by the analytic and Monte Carlo Bayesian solutions for three positional values is shown in Figure 5 and Figure 6. In each diagram the analytic result is represented by a square symbol plotted at the mean of the distribution and an ellipse representing the covariance between the two variables is displayed. The size of the ellipse is selected such that its semi-major and semi-minor axes are equal to one standard deviation in a direction along each axis. The major axis of the ellipse is oriented along the line of highest correlation. The Monte Carlo result is represented by the sample values output by the M-H algorithm. In Figure 5, the diagram shows the real and imaginary parts of the recovered scene at positions 10 and 11. For both positions there is a good agreement between the location and spread as determined by the two methods. Figure 6 shows the correlation between the real part of the recovered scene at positions 9 and 10. Agreement between the Monte Carlo and analytic results in both the direction and degree of correlation is also very

good. The main difference between the two methods is that the samples do not fully cover the extreme tails of the distribution so the variance calculated from samples is slightly lower than the analytically derived result.

The above analysis is highly detailed given that the considered scene simply contains two ideal scatterers embedded in uncorrelated clutter. This has only been possible through the use of Bayesian statistics. Other super-resolution algorithms give a single solution to the problem and do not allow uncertainty in the recovered scene to be analyzed. The discovery of correlation between adjacent positions is a particularly useful result because this explains why there are many possible solutions to the super-resolution problem consistent with noisy data. While the existence of correlations could have been conjectured from a qualitative analysis, the Bayesian approach has given precise information as to which values are correlated and by how much. When super-resolution is used without a further specific application, this information is useful in determining why the results are good or bad for a certain scene and SNR. However, the full value of the Bayesian approach is achieved when a further stage of processing, such as automatic target detection or recognition, is used after super-resolution. In this situation uncertainty in the scene can be used as a numerical input to the next processing stage. This could be crucial in a decision-making process, where existence of uncertainty changes the decision boundary. The process of combining super-resolution and automatic target recognition (ATR) is examined in detail in [13].

The analytic solution is relatively easy to implement computationally because it consists simply of calculating a vector and a matrix from other quantities using addition, multiplication and the matrix inverse. In a direct implementation this results in an algorithm with runtime of order $n^3$ where $n$ is the number of samples or pixels in the signal or image. However, since the structure of the PSF matrix $\mathbf{T}$ is Toeplitz, computation could be speeded up by means of a fast Fourier transform (FFT) [41]. The Metropolis-Hastings algorithm is somewhat more difficult to implement due to its additional

complexity. The run time in a direct implementation is of order $mn^3$, where $m$ is the number of samples required to represent the distribution. As before, an FFT could be used to speed up the computation. Although the Metropolis-Hastings algorithm is slower to run than the analytic solution, it is more general in that it can be applied to non-linear and non-Gaussian systems.

# 5 Conclusion

Super-resolution of signals and images can improve the automatic detection and recognition of objects of interest [2][5]. However, the uncertainty associated with this process is not often taken into consideration. This paper has presented two non-parametric Bayesian super-resolution techniques that not only take uncertainty into account, but also retain knowledge about the output uncertainty in the form of a full probability distribution. This knowledge could be used as part of a wider framework for target detection and recognition. Using measured 2D radar data, it has been shown in [13] that such Bayesian super-resolution algorithms can increase automatic target recognition performance over standard super-resolution.

The algorithms presented in this paper are applicable to radar images as well as any other data that follows the same model used, such as optical or ultrasound images. However, radar data is often modelled as a set of discrete point scatterers with certain properties. It is possible that this type of parametric model could be incorporated to a Bayesian super-resolution algorithm. This would require the use of trans-dimensional techniques, such as reversible jump Markov chain Monte Carlo (RJMCMC), to estimate the number of parameters in addition to the parameter values themselves [11]. Such an approach to radar data processing should be the subject of future research.

# 6 Acknowledgments

The author recognizes the assistance of Dr Keith Copsey at QinetiQ Malvern in coding the MCMC Bayesian super-resolution algorithm. The author also thanks Dr Andrew Webb of QinetiQ Malvern and Professor Chris Baker, formerly of University College London, for advice given during the research and development of these algorithms. This research was funded in part by the EPSRC Engineering Doctorate (EngD) scheme.

# 7 References

[1] Doerry, A.W., and Dickey, F.M.: 'Synthetic aperture radar', Optics and Photonics News, November 2004, 15, (11), pp. 28-33

[2] Anstey, C.P., and Hayward, S.D.: 'The synthesis of high resolution pulse-Doppler waveforms using sparsely sampled frequency sets', International Waveform Diversity and Design Conference, Lihue, Hawaii, USA, January 2006

[3] Zhang, X., Shen, R., and Guo, G.: 'Automatic HRR target recognition based on Prony model wavelet and probability neural network', CIE International Conference of Radar, October 1996, pp. 143-146

[4] Nguyen, D.H., Benitz, G.R., Kay, J.H., Orchard, B.J., and Whiting, R.: 'Superresolution HRR ATR with high definition vector imaging', IEEE Transactions on Aerospace and Electronic Systems, October 2001, 37, (4), pp. 1267-1286

[5] Novak, L.M., Owirka, G.J., and Weaver, A.L.: 'Automatic target recognition using enhanced resolution SAR data', IEEE Transactions on Aerospace and Electronic Systems, January 1999, 35, (1), pp. 157-175

[6] Çetin, M.: 'Feature-enhanced synthetic aperture radar imaging', PhD thesis, College of Engineering, Boston University, USA, 2001

[7] Mehra, R.K., Ravichandran, B., and Huff, M.: 'Survey of radar superresolution methods with applications to automatic target recognition', Proceedings of SPIE, April 1998, 3374, pp.186-193

[8] Kim, K.T., Bae, J.H., and Kim, H.T.: 'Effect of AR model-based data extrapolation on target recognition performance', IEEE Transactions on Antennas and Propagation, April 2003, 51, (4), pp. 912-914

[9] Fuller, D.F., Terzuoli, A.J., Collins, P.J., and Williams, R.: 'Approach to object classification using dispersive scattering centres', IEE Proc.-Radar Sonar Navig., April 2004, 151, (2), pp. 85-90

[10] Lee, P.M.: 'Bayesian Statistics, An Introduction' (Arnold, London, 2nd edition, 1997)

[11] Stawinski, G., Doucet, A. and Duvaut, P.: 'Reversible jump Markov chain Monte Carlo for Bayesian deconvolution of point sources', Proceedings of SPIE, July 1998, 3459, pp. 179-190

[12] Luttrell, S.P., and Oliver, C. J.: 'Prior knowledge in synthetic-aperture radar processing', J. Phys. D: Appl. Phys., 1986, 19, pp. 333-356

[13] Lane, R.O.: 'Bayesian super-resolution with application to radar target recognition', EngD thesis, Department of Electronic and Electrical Engineering, University College London, February 2008

[14] Press, W.H., Teukolsky, S.A., Vetterling, W.T., and Flannery B.P.: 'Numerical Recipes in C', (Cambridge University Press, 2nd edition, 1992)

[15] Stankwitz, H.C., and Kosek, M.R.: 'Sparse aperture fill for SAR using super-SVA', IEEE National Radar Conference, May 1996, pp. 70-75

[16] Capon, J.: 'High-resolution frequency-wavenumber spectrum analysis', Proc. IEEE, August 1969, 57, (8), pp. 1408-1418

[17] Li, J. and Stoica, P.: 'An adaptive filtering approach to spectral estimation and SAR imaging', IEEE Transactions on Signal Processing, June 1996, 44, (6), pp. 1469-1484

[18] Benitz, G.R.: 'High-definition vector imaging', The Lincoln Laboratory Journal, 1997, 10, (2), pp. 147-169

[19] Nguyen, M.K., and Mohammad-Djafari, A.: 'Bayesian approach with the maximum entropy principle in image reconstruction from microwave scattered field data', IEEE Transactions on Medical Imaging, June 1994, 13, (12), pp. 254-262

[20] Sarkar, T.K.: 'An ultra-low sidelobe pulse compression technique for high performance radar systems', IEEE National Radar Conference, May 1997, pp. 111-114

[21] Sarkar, T.K., Weiner, D.D., and Jain, V.K.: 'Some mathematical considerations in dealing with the inverse problem', IEEE Transactions on Antennas and Propagation, March 1981, AP-29, (2), pp. 373-379

[22] Dickey, F.M., Romero, L.A., DeLaurentis, J.M., and Doerry, A.W.: 'Super-resolution, degrees of freedom and synthetic aperture radar', IEE Proc.-Radar Sonar Navig, December 2003, 150, (6), pp. 419-429

[23] Luttrell, S.P.: 'Prior knowledge and object reconstruction using the best linear estimate technique', Optica Acta, June 1985, 32, (6), pp. 703-716

[24] Delves, L.M., Pryde, G.C., and Luttrell, S.P.: 'A super-resolution algorithm for SAR images', Inverse Problems, 1988, 4, (3), pp. 681-703

[25] Selén, Y. and Stoica P.: 'Estimation of semi-sparse radar range profiles', Digital Signal Processing, July 2008, 18, (4), pp. 543-560

[26] Candocia, F.M.: 'A unified superresolution approach for optical and synthetic aperture radar images', PhD thesis, University of Florida, USA, May 1998

[27] Blunt, S.D., and Gerlach, K.: 'Adaptive pulse compression via MMSE estimation', IEEE Transactions on Aerospace and Electronic Systems, April 2006, 42, (2), pp. 572-584

[28] Lane, R.O.: 'The effects of Doppler and pulse eclipsing on sidelobe reduction techniques', IEEE National Radar Conference, Verona, NY, April 2006

[29] Luttrell, S.P.: 'A Bayesian derivation of an iterative autofocus / superresolution algorithm', Inverse Problems, 1990, 6, (6), pp. 975-996

[30] Jalobeanu, A., Blanc-Féraud, L. and Zerubia, J.: 'Hyperparameter estimation for satellite image restoration using a MCMC maximum-likelihood method', Pattern Recognition, February 2002, 35, (2), pp. 341-352

[31] Higdon, D.M., Bowsher, J.E., Johnson, V.E., Turkington, T.G., Gilland, D.R., and Jaszczak, R.J.: 'Fully Bayesian estimation of Gibbs hyper-parameters for emission computed tomography data', IEEE Transactions on Medical Imaging, October 1997, 16, (5), pp. 516-526

[32] Zarb Adami, K,: 'Bayesian inference and deconvolution', PhD thesis, University of Cambridge, UK, December 2003

[33] Park, S.C., Park, M.K., and Kang, M.G.: 'Super-resolution image reconstruction: a technical overview', IEEE Signal Processing Magazine, May 2003, 20, (3), pp. 21-36

[34] Tian, J. and Ma, K.-K.: 'A MCMC approach for Bayesian super-resolution image reconstruction', IEEE International Conference on Image Processing, September 2005, 1, pp. 45-48

[35] Tian, J. and Ma, K.-K.: 'Markov chain Monte Carlo super-resolution image reconstruction with artifacts suppression', IEEE Asia Pacific Conference on Circuits and Systems, December 2006, pp. 940-943

[36] Pickup, L.C., Capel, D.P., Roberts, S.J., and Zisserman, A: 'Bayesian image super-resolution, continued', in Schölkopf, B., Platt, J., and Hoffman, T., (Eds.): 'Advances in Neural Information Processing Systems 19', (MIT Press, Cambridge, MA, 2007), pp. 1089-1096

[37] Chib, S. and Greenberg, E.: 'Understanding the Metropolis-Hastings algorithm', The American Statistician, November 1995, 49, (4), pp. 327-335

[38] Lane, R.O., Copsey, K.D., and Webb, A.R.: 'A Bayesian approach to simultaneous autofocus and super-resolution', Proceedings of SPIE, April 2004, 5427, pp. 133-142

[39] Copsey, K.D.: 'Bayesian Approaches for Robust Automatic Target Recognition', PhD thesis, Imperial College London, UK, August 2004

[40] Lane, R.O.: 'Super-resolution and the radar point spread function', Proceedings of the London Communications Symposium, September 2005, pp. 5-8

[41] Jain, A.K.: 'Fast inversion of banded Toeplitz matrices by circular decompositions', IEEE Transactions on Acoustics, Speech, and Signal Processing, April 1978, ASSP-26, (2), pp. 121-126

## Figures

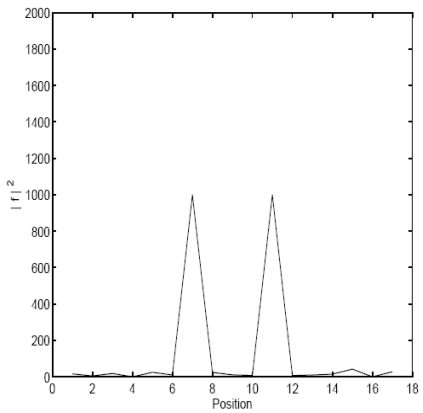


**Figure 1: Power profile of the true-scene complex field**

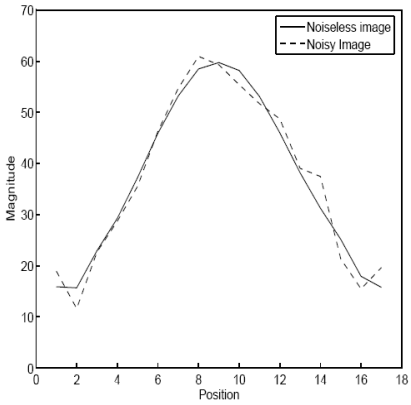


**Figure 2: Magnitude of a low-resolution noisy image with an SNR of 20 dB.**

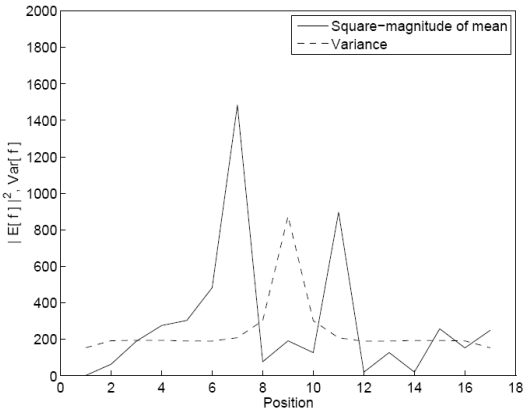


**Figure 3: Analytic Bayesian solution: mean power and variance range profile.**

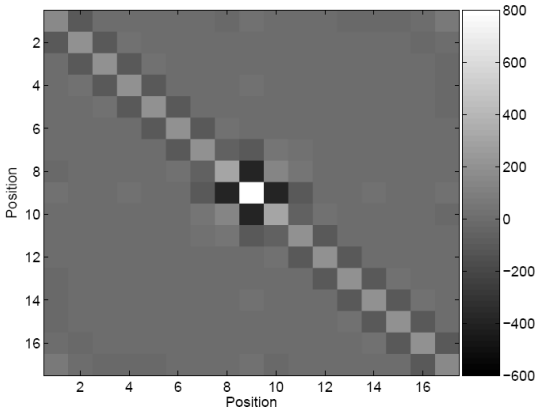


**Figure 4: Analytic Bayesian solution: real part of the covariance matrix.**

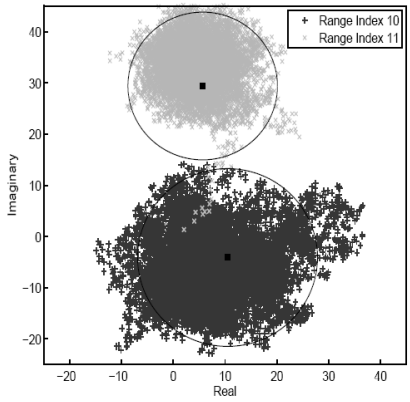


**Figure 5: Comparison of real and imaginary sample values output by the M-H algorithm, and covariance ellipses from the analytic Bayesian solution.**

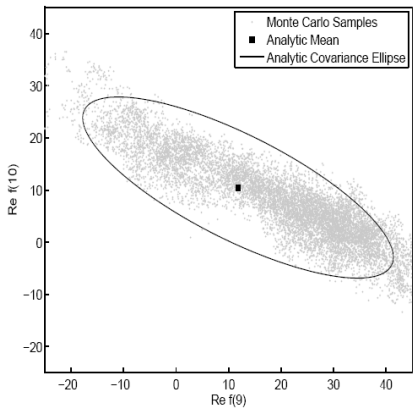


**Figure 6: Comparison of adjacent sample values output by the M-H algorithm, and the covariance ellipse from the analytic Bayesian solution.**