Statistical radar imaging of diffuse and specular targets using an expectation-maximization algorithm

Aaron D. Lanterman

Coordinated Science Laboratory, Univ. of Illinois, 1308 W. Main, Urbana, IL 61801

ABSTRACT

Radar imaging is often posed as a problem of estimating deterministic reflectances observed through a linear mapping and additive Gaussian receiver noise. We consider an alternative view which considers the reflectances themselves to be a realization of a random process; imaging then involves estimating the parameters of that underlying process. Purely diffuse radar targets are modeled by a zero-mean Gaussian process, while specular targets introduce an additive component with fixed amplitude and random uniform phase. When conventional stepped frequency waveforms are employed, the linear mapping amounts to a Fourier transform, and parameter estimation is straightforward. If more complicated waveforms are employed, maximum-likelihood parameter estimates cannot be readily computed analytically; hence, we explore an iterative expectation-maximization algorithm proposed by Snyder, O'Sullivan, and Miller. Although this algorithm was designed for diffuse radar imaging, arguments based on the central limit theorem and computational experiments support its applicability to the specular case.

The resulting estimates tend to be unacceptably rough due to the ill-posed nature of maximum-likelihood estimation of functions from limited data, so some kind of regularization is needed. We explore penalized likelihoods based on entropy functionals, a roughness penalty proposed by Silverman, and an information-theoretic formulation of Good's roughness penalty.

1. INTRODUCTION

In the late 80's, Snyder, O'Sullivan, and Miller[1] proposed an approach to radar imaging which represented a radical departure from traditional radar techniques. They employed statistical models for radar turns which, although well-explored in the *detection* arena, had received relatively little attention in the development of *imaging* algorithms. Instead of pre-supposing a certain kind of processing structure, such as linear filtering, they derived an imaging algorithm based on first-principles estimation-theoretic techniques, resulting in a nonlinear processor consisting of an iterative expectation-maximization[2,3] (EM) algorithm for maximum-likelihood imaging. Intuitively speaking, the algorithm "deconvolves" the ambiguity function associated with traditional matched filtering by fully exploiting statistical knowledge of the underlying data generation mechanism.

Most modern high-resolution radar systems employ stepped frequency waveforms. Such waveforms are not necessarily optimal from any theoretical point of view; instead, their popularity arises from the simplicity of the resulting processing algorithms, which often reduce to two-dimensional fast Fourier transforms. A wide variety of other waveforms have been proposed, including sequences with a varying chirp rate,[4–6] decaying exponential envelopes,[7,8] frequency-hopping pulse trains built from Costas arrays,[9] constant-power, amplitude-modulated pulse trains based on Barker codes or pseudorandom sequences (Ref. 10, Sec. 10.4.2), and random signal radar (RSR) waveforms.[11] In addition to these active waveforms, passive radar applications[12] which employ commercial television or FM radio signals have garnered increasing interest recently. The framework of Snyder, O'Sullivan, and Miller allows easy incorporation of any type of transmitted waveform.

Although the approach and resulting algorithm was presented over a decade ago, surprisingly little work has been reported on it, especially compared with other applications of EM algorithms, such as Poisson imaging in emission tomography,[13,14] astronomy,[15] and optical sectioning microscopy.[16] This is most likely due to its computational complexity. The original 1989 paper which presented the imaging algorithm[1] contained no examples of its execution; examples of it in the literature have been mostly restricted to the special case of stepped frequency waveforms,[17,18] which permit some substantial simplifications that allow great reductions in computational complexity. Considering the rapid advancement of computer power, we feel that it is time to reintroduce this algorithm to the radar imaging

E-mail: lanterma@ifp.uiuc.edu, web: www.ifp.uiuc.edu/~lanterma

community. To our knowledge, this paper is the first to report an implementation of the algorithm for transmitted signals which are not stepped frequency waveforms.

Section 2 outlines our probabilistic formulation of radar imaging, including specialization to the delay-doppler paradigm explored here. Section 3 derives an EM algorithm for maximum-likelihood imaging which is applied to simulated data in Sec. 4. These simulations illustrate the need for regularization as discussed in Sec. 5. Section 6 postulates some directions for future work.

2. PROBLEM FORMULATION

In the statistical approach to high-resolution radar imaging, the received data is modeled as $\mathbf{r} = \mathbf{\Gamma}^\dagger \mathbf{c} + \mathbf{w}$, where $\mathbf{c}$ is the complex target reflectance, $\mathbf{\Gamma}$ represents a linear observation mechanism, and the receiver noise vector $\mathbf{w}$ is i.i.d. complex Gaussian with variance N_0. We will let $\mathbf{\Gamma}$ be rather general. For simple range profiling, the rows of $\mathbf{\Gamma}$ contain samples of the transmitted waveform; for delay-doppler imaging, it will contain the transmitted waveform multiplied by complex sinusoids associated with doppler shifts; for tomographic imaging, $\mathbf{\Gamma}$ may also represent scene rotations. In addition to radar imaging, power spectrum estimation can also be formulated this way, in which case $\mathbf{\Gamma}$ holds complex exponentials. Our algorithm will result from performing optimal statistical inference using this model, and not presuppose any particular kind of processing structure such as matched filtering.

Although statistical models for the radar reflectance $\mathbf{c}$ have enjoyed frequent application in analyzing the performance of target detection systems,[10] their application in radar imaging has been far less frequent; typically, radar imaging is posed as a deterministic reconstruction problem amenable to traditional filtering techniques, with $\mathbf{c}$ treated as a nonrandom parameter.

Shapiro *et al.*[19] have suggested modeling $\mathbf{c}$ as a random process consisting of specular (glint) and diffuse (speckle) components. Diffuse returns may be modeled as $\mathbf{c}_d \sim CN(0, \mathbf{\Sigma})$, where $\mathbf{\Sigma}$ is a diagonal covariance matrix. The vector of its diagonal entries $\boldsymbol{\sigma} = diag(\mathbf{\Sigma})$ is called the *scattering function.*[*] Specular returns are modeled as $\mathbf{c}_s = \mathbf{b} \times exp[j\theta]$, where θ is an i.i.d. uniform random vector over $[0, 2\pi)$, $\mathbf{b}$ is the deterministic *glint reflection coefficient function*, and $\times$ represents elementwise multiplication.

Under a purely diffuse model, $\mathbf{r} \sim CN(0, \mathbf{K_r})$ where $\mathbf{K_r} = \mathbf{\Gamma}^\dagger \mathbf{\Sigma} \mathbf{\Gamma} + N_0 \mathbf{I}$ and estimating the scattering function brings us to the structured covariance estimation waters first charted by Burg.[20] The loglikelihood for the data is

$$L_{id}(\mathbf{\Sigma}) = -\ln\det \mathbf{K_r} - \mathbf{r}^H \mathbf{K_r}^{-1} \mathbf{r}. \tag{1}$$

If $\mathbf{\Gamma}$ is unitary (for instance, in the case of properly chosen stepped-frequency waveforms), then one can orthogonalize the data by computing $\tilde{\mathbf{r}} = \mathbf{\Gamma r}$, and the ML estimate is just $\sigma_i = \max(0, |\tilde{r}_i|^2 - N_0)$; otherwise, there is no closed-form solution for the maximizer of (1). Section 3 reviews an iterative expectation-maximization algorithm for computing ML estimates of $\boldsymbol{\sigma}$ from $\mathbf{r}$ proposed by Snyder *et al.*[1]

When the reflectance process has a specular component, the distribution of $\mathbf{r}$ becomes more complicated. If $\mathbf{\Gamma}$ is unitary, the data can be orthogonalized, and $|\tilde{\mathbf{r}}|$ consists of independent Rician random variables amenable to estimation algorithms described in Refs. 21–23. If no such orthogonalization is available, we are not aware of any closed form for the density of $\mathbf{r}$, which would make pure maximum-likelihood estimation a formidable task. However, note that if the columns of $\mathbf{\Gamma}$ have a sufficient number of non-zero entries, then the elements of $\mathbf{r}$ consist of sums of independent zero-mean random variables; by the central limit theorem, the marginals on $\mathbf{r}$ appear to be Gaussian (although the joint distribution will be more intricate), and $\mathbf{r}$ overall appears "nearly Gaussian" in the sense of Mallows.[24] This observation, along with recent computational experience, has motivated applying the diffuse-reflectance EM algorithm described above in situations where the reflectance process contains specular components.

2.1. Specialization to Delay-Doppler Imaging

Let $s(t)$ denote the continuous-time transmitted signal. We adopt the discretization technique employed by Moulin (Sec. 2.3, Ref. 25). Suppose we want to compute the delay-doppler scattering function on a discrete grid at delay times $\tau_\ell, \ell = 1, \ldots, I_R$, with a spacing of δ_τ, and at doppler frequencies $f_k, k = 1, \ldots, I_{CR}$, with a spacing of δ_f. The vector $\boldsymbol{\sigma}$ represents the scattering function at delay τ_ℓ and doppler f_k with the element $\sigma_{f+\ell I_{CR}}$. One can imagine the I_R by I_{CR} 2-D image being unwrapped and strung along a length $I_R I_{CR}$ 1-D vector.

[*]To avoid some cumbersome double-superscripting in later sections, we use σ to represent variance instead of the traditional σ^2.

Suppose that N samples of the radar return are taken at a spacing of δ_t. The $(n, k+\ell I_{CR})^{th}$ entry of the $N \times I$ matrix $\mathbf{\Gamma}^H$ (where $I = I_R I_{CR}$) is given by (Eq. 16 of Ref. 25)

$$\exp[j2\pi f_k(n\delta_t - \tau_\ell/2)]s(n\delta_t - \tau_\ell). \tag{2}$$

One could develop $\mathbf{\Gamma}$ matrices for other imaging models, such as the tomographic paradigm of Munson *et al.*[26] In this case, $\mathbf{\Gamma}$ would incorporate angle-limited Radon transforms.

3. AN EM ALGORITHM FOR MAXIMUM-LIKELIHOOD IMAGING

Maximum-likelihood imaging requires us to find the $\mathbf{\Sigma} = diag(\boldsymbol{\sigma})$ which maximizes (1). Since no simple solution is evident for this difficult optimization problem, we turn to an expectation-maximization algorithm derived by Snyder, O'Sullivan, and Miller[1]; the algorithm is also reviewed in the overall context of information-theoretic imaging in Sec. IX.C of Ref. 27. In the context of the EM algorithm, $\mathbf{r}$ is called the *incomplete data* and (1) is the *incomplete data loglikelihood*.

To formulate an EM algorithm, one postulates a set of *complete data*[†] which, if available, would make the maximization problem easier. Here we take the complete data to be $\{\mathbf{c}, \mathbf{w}\}$, where $\mathbf{c} \sim CN(0, \mathbf{\Sigma})$ and $\mathbf{w} \sim CN(0, N_0\mathbf{I})$.

Traditional EM formulations requires that there be a many-to-one mapping from the complete data to the incomplete data; this is provided by $\mathbf{r} = \mathbf{\Gamma}^H\mathbf{c} + \mathbf{w}$. The complete data loglikelihood is

$$L_{cd}(\mathbf{\Sigma}) = -\ln\det\mathbf{\Sigma} - \mathbf{c}^H\mathbf{\Sigma}\mathbf{c} = -\sum_{i=1}^{I}\sigma_i - \sum_{i=1}^{I}\frac{|c_i|^2}{\sigma_i}. \tag{3}$$

Let Q denote the expectation of the complete-data loglikelihood given the incomplete data and the estimate from the previous iteration:

$$Q[\mathbf{\Sigma}|\mathbf{\Sigma}^{old}] \stackrel{\text{df}}{=} E[L_{cd}(\mathbf{\Sigma})|\mathbf{\Sigma}^{old}, \mathbf{r}] = -\sum_{i=1}^{I}\ln\sigma_i - \sum_{i=1}^{I}\frac{E[|c_i|^2|\mathbf{\Sigma}^{old}, \mathbf{r}]}{\sigma_i}. \tag{4}$$

At each iteration of the EM algorithm, we get the new estimate by maximizing Q. The derivative of (4) with respect to σ_i is

$$-\frac{1}{\sigma_i} + \frac{E[|c_i|^2|\mathbf{\Sigma}^{old}, \mathbf{r}]}{\sigma_i^2}. \tag{5}$$

Equating (5) with zero yields the simple update

$$\sigma_i^{new} = E[|c_i|^2|\mathbf{\Sigma}^{old}, \mathbf{r}]. \tag{6}$$

Computing the expectation in (6) is a standard problem in estimation theory texts; see, for instance, Eq. 7.112 on p. 303 of Scharf[28] or Eqs. V.B.21 and 7.B.22 on p. 221 of Poor.[29] This results in the explicit update

$$\sigma_i^{new} = \sigma_i^{old} - (\sigma_i^{old})^2[\mathbf{\Gamma}\mathbf{K}^{-1}\mathbf{\Gamma}^H - \mathbf{\Gamma}\mathbf{K}^{-1}\mathbf{S}\mathbf{K}^{-1}\mathbf{\Gamma}^H]_{ii}, \tag{7}$$

where $\mathbf{K} = \mathbf{\Gamma}^H\mathbf{\Sigma}^{old}\mathbf{\Gamma} + N_0\mathbf{I}$ and $\mathbf{S} = \mathbf{r}\mathbf{r}^H$.

Notice that the data only enters into the inference via its empirical covariance $\mathbf{S}$. We have written (7) in a way to show the symmetric structure of the matrix computations. In implementation, it is more efficient to compute the term in brackets by calculating $\Xi = \mathbf{\Gamma}\mathbf{K}_{\mathbf{r}}{}^{old}$ followed by $\Xi[\mathbf{\Gamma}^H - \mathbf{S}\Xi^H]$. Since only the diagonal terms are needed, the final matrix multiplication only requires I inner products.

[†]We call the resulting algorithm "an" EM algorithm instead of "the" EM algorithm since different choices of complete data may result in different algorithms.

4. EXAMPLES

This section explores two different kinds of targets. One consists of three distinct point scatterers. The other is a 10 meter diameter sphere rotating at a rate of 100 revolutions[‡] per second; one can imagine a styrofoam ball wrapped in crinkled aluminum foil. (Such an experiment is described in Ref. 18). The theoretical delay-doppler scattering functions of these targets is shown in Fig. 1. The scattering function of the sphere was calculated as described in Appendix B of Ref. 30. To avoid an aliased appearance, the sphere image was compute on a grid four times denser and then downsampled via averaging.

Figure 1. Two ideal scattering functions used in the simulations.

We will explore data generated under both diffuse and specular models. In the specular experiments, we take $b = \sqrt{\sigma}$. Note that the scattering function for the sphere was derived under the diffuse assumption; we do not necessarily expect that in reality a rotating *smooth* sphere would generate specular reflections based on the diffuse scattering function. We do this to provide a comparison of the overall effect of the different kind of underlying statistical models on an even footing.

Our simulations employ a center frequency of 15 GHz. The transmitted waveform is a 2.13 microsecond long constant-power waveform consisting of 64 equally spaced segments; the amplitude of each segment was independently set to -1 or 1, with equal probabilities. Such waveforms are known to have good ambiguity properties (Ref. 10, Sec. 10.4.2). The waveform and its autocorrelation are shown in Fig. 2.

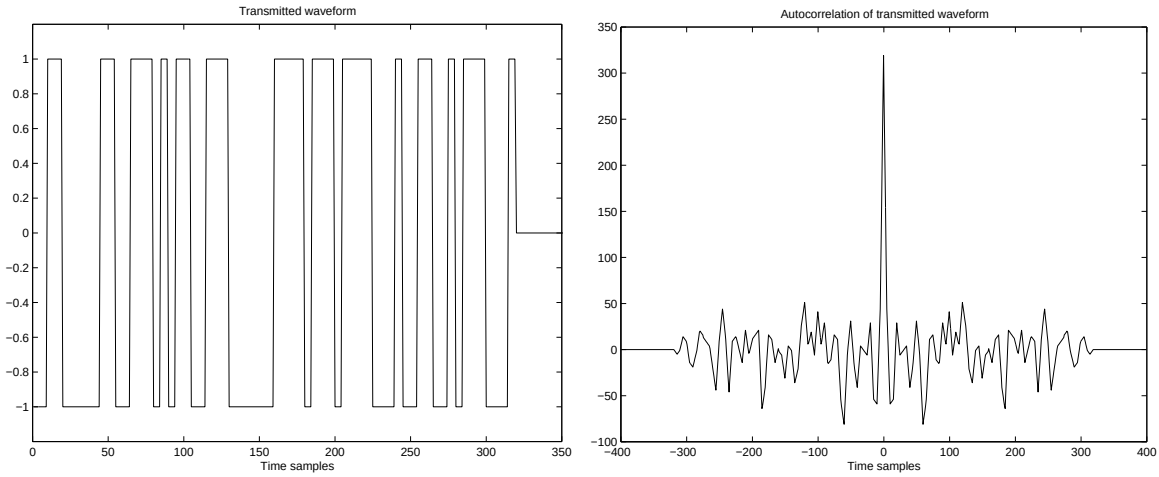

Figure 2. Left panel: Transmitted waveform. Right panel: Autocorrelation of transmitted waveform.

We discretize the delay-doppler space into a 20×20 grid. We desire a range and velocity spacing of $\delta_r = 1$ meter and $\delta_{cr} = 100$ meters per second, so we choose spacings of $\delta_\tau = 2\delta_r/c \approx 6.7$ nanoseconds per pixel in delay and

[‡]Such high rotation rates are unlikely to be encountered in practice. We use it here in our initial studies since it permits the doppler spread of the system ambiguity function to fit reasonably within our small 20 by 20 reconstruction grid for $\mathbf{\Sigma}$. Future experiments on larger grids will allow more realistic velocities.

$\delta_f = 2\delta_{cr} f_c/c \approx 10$ kilohertz per pixel in doppler. The radar return is sampled with a spacing of $\delta_t = \delta_\tau$. At this sampling rate (150 MHz), the transmitted waveform is 319 samples long. We take collect 400 returned samples, so the resulting $\mathbf{\Gamma}$ matrix is square.

Fig. 3 shows the magnitude of data simulated using the three-point scatterer. (We do not present the phases since they are not very informative to a human observer.) A specular model was assumed in which each point was assigned a random uniform phase. The leftmost panel of Fig. 4 shows $|\mathbf{\Gamma r}|^2$, the magnitude squared of the output of a matched filter. Notice that only two points are evident in the matched-filtered output. The remaining panels show the results of successive iterations of the EM algorithm. At 10 iterations, the scatterers are quite well resolved. By 20 iterations, the background energy has been removed.

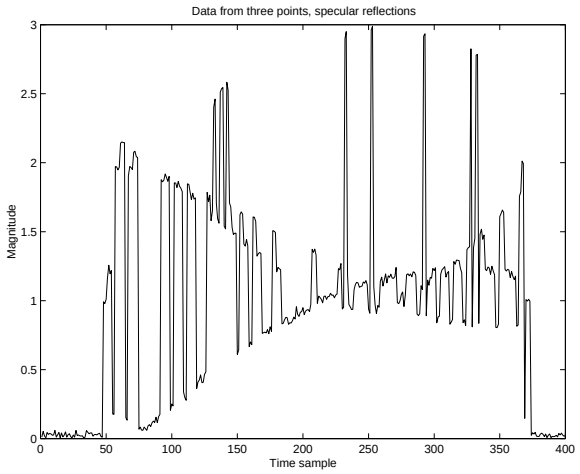

Figure 3. Magnitude of data from a specular realization of the three-point scatterer.

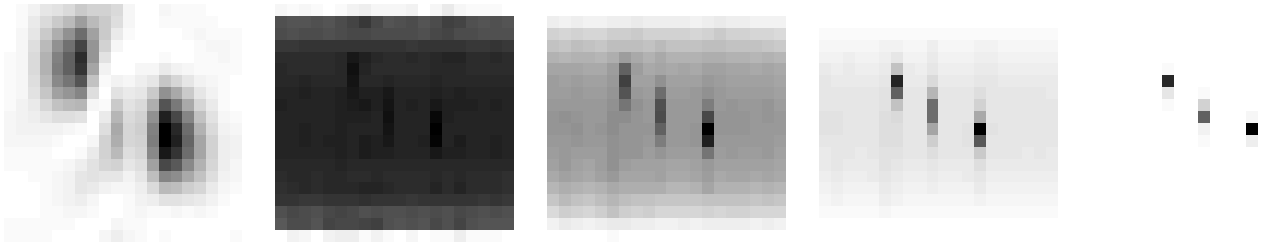

Figure 4. Matched-filtered output and EM iterations for the three-point scatterer. Leftmost panel shows the squared magnitude of the matched-filtered output. From left to right, the remaining panels show results at 1, 5, 10, and 20 iterations.

Figure 5. Squared-magnitudes of two realizations of the underlying diffuse random process $\mathbf{c}$ from the sphere.

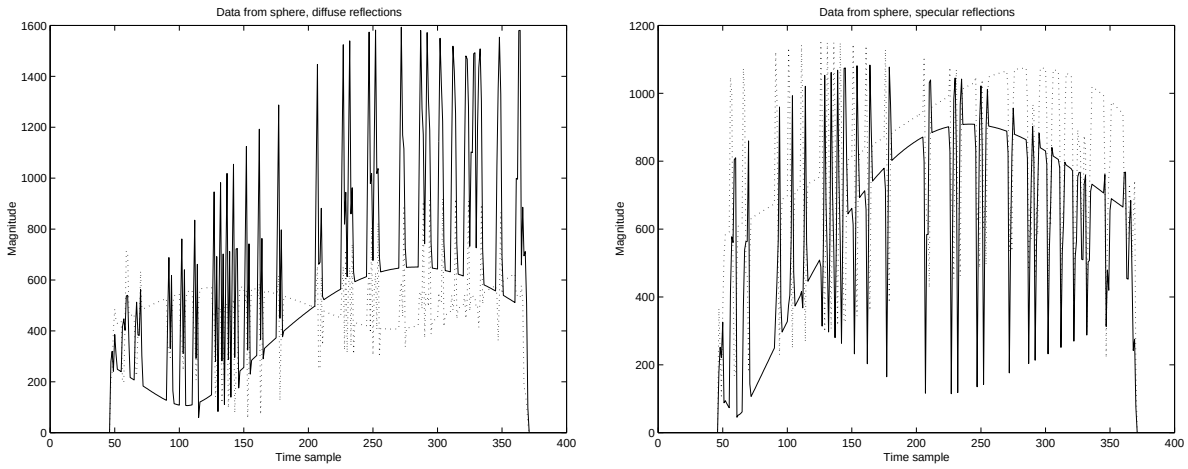

Figure 6. Magnitude of data from two diffuse (left panel) and two specular (right panel) realizations from the sphere. For the left panel, the solid line corresponds to the left panel of Fig. 5; dotted line corresponds to the right panel.

Figure 5 illustrates two realizations of the random process **c** associated with a diffuse radar model. The left panel of Fig. 6 presents the data resulting from these realizations. The different rows of Fig. 7 show the result of processing the two data sets. The top row corresponds to the left panel of Fig. 5, and the bottom row corresponds to the right panel. The leftmost column shows the results of matched filtering; the remaining columns show results of EM iterations. Note that in the top matched-filtered reconstruction, the overall location of the sphere is correctly indicated, albeit blurred. However, in the lower matched-filtered image, the most evident location of the sphere is incorrect. By contrast, in even as few as five iterations, the location of the sphere is obvious (although suffering from blurring in the doppler direction). Further iterations make the overall outline of the scatterer quite evident and sharp, although the reconstruction suffers from a spiky, noisy appearance. Section 5 will revisit this unpleasant behavior and discuss methods of addressing it.

Figure 7. Matched-filtered output and EM iterations for diffuse data generated from the sphere. Leftmost column shows the squared magnitude of the matched-filtered output. From left to right, the remaining columns show results at 1, 5, 10, and 20 iterations. Top and bottom rows correspond to different diffuse data realizations.

Figure 8 reproduces the sphere experiment, except specular instead of diffuse reflections are used. The data is

shown in the right panel of Fig. 6. In this case, only the phases of the underlying process $\mathbf{c}$ are random, so the equivalent of Fig. 5 would look simply like the sphere image of Fig. 1. Although the algorithm was not designed for specular scatterers, it appears to do as well as it did for the diffuse case.

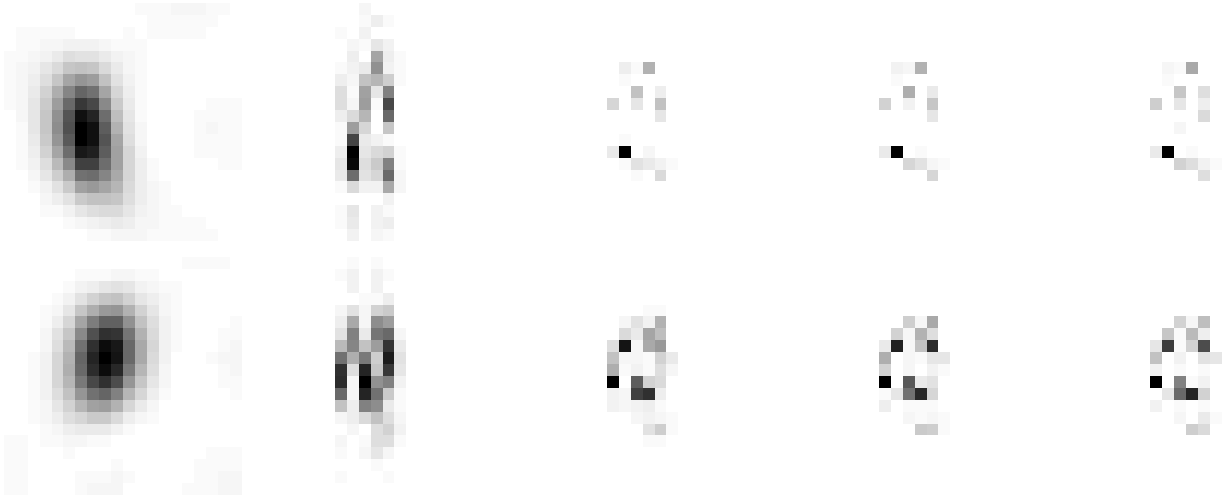

Figure 8. Matched-filter output and EM iterations for specular data generated from the sphere. Leftmost column shows the squared magnitude of the matched-filtered output. From left to right, the remaining columns show results at 1, 5, 10, and 20 iterations. Top and bottom rows correspond to different specular data realizations.

5. REGULARIZATION

The roughness associated with increasing EM iterations observed in Figs. 7 and 8 is indicative of the ill-posed nature of maximum-likelihood estimation of functions from limited data. In most numerical algorithms, accuracy improves as the discretization is refined. Here, the opposite is true; as the grid is refined, the problem becomes increasingly ill-posed and the solutions increasingly ill-behaved. This phenomena, called *dimensional instability* by Tapia and Thompson,[31] would manifest itself in any algorithm which maximizes the loglikelihood, and is not a problem with the EM algorithm itself per se.

These issues have been extensively studied in the context of Poisson intensity estimation in applications such as medical imaging (see Refs. 32, 33, and Chapter 3 of Ref. 34). Several solutions have been proposed for Poisson imaging which we can adapt to our radar imaging problem.

5.1. The Method of Sieves

One approach is to restrict the solution to lie in a restricted subset called a *sieve.*[35] One possibility is to require it to be a sum of weighted basis functions. Gaussian basis yielded much success in emission tomography.[36,33,37] In the radar problem, the EM algorithm of Sec. 3 can be extended to incorporate such a sieve constraint; Refs. 17,18 present the extended algorithm and specific examples employing B-splines. Moulin[38] has also explored wavelet thresholding techniques for denoising diffuse radar imagery when an orthogonalization of the data is available as described in Sec. 2. To avoid difficulties with nonnegativity, the wavelets are used to represent the *logarithm* of the scattering function. The logarithm turns the multiplicative noise inherent in speckle imagery into additive noise amenable to traditional wavelet denoising.

5.2. The Method of Penalties

Another philosophically appealing approach to regularization is to subtract a penalty $\Phi(\mathbf{\Sigma})$ which discourages unacceptably rough estimates and maximize the penalized loglikelihood

$$P(\mathbf{\Sigma}) = L_{id}(\mathbf{\Sigma}) - \alpha\Phi(\mathbf{\Sigma}). \tag{8}$$

Bayesians may think of this as *maximum a posteriori* estimation using a (usually improper) prior proportional to $\exp[-\alpha\Phi(\mathbf{\Sigma})]$. Although the method of penalties has enjoyed much success in Poisson imaging applications, it has not yet been explored in the maximum-likelihood radar imaging context.

The EM algorithm can be easily extended to maximize this penalized loglikelihood. The Q function (4) is generalized to

$$\begin{aligned} Q_P[\mathbf{\Sigma}|\mathbf{\Sigma}^{old}] &\stackrel{\text{df}}{=} E[L_{cd}(\mathbf{\Sigma})|\mathbf{\Sigma}^{old},\mathbf{r}] - \alpha\Phi(\mathbf{\Sigma}) \\ &= -\sum_{i=1}^{I}\ln\sigma_i - \sum_{i=1}^{I}\frac{E[|\mathbf{c}_i|^2|\mathbf{\Sigma}^{old},\mathbf{r}]}{\sigma_i} - \alpha\Phi(\mathbf{\Sigma}). \end{aligned} \tag{9}$$

To maximize (10), we can take derivatives (analogous to (5)) and solve the set of equations:

$$-\frac{1}{\sigma_i} + \frac{E[\mathbf{c}_i^2|\Sigma^{old},\mathbf{r}]}{\sigma_i^2} - \alpha\frac{\partial\Phi(\mathbf{\Sigma})}{\partial\sigma_i} = 0, \tag{10}$$

alternatively written as

$$-\frac{1}{\sigma_i} + \frac{\sigma_i^{uc}(\mathbf{\Sigma}^{old})}{\sigma_i^2} - \alpha\frac{\partial\Phi(\mathbf{\Sigma})}{\partial\sigma_i} = 0, \tag{11}$$

where $\boldsymbol{\sigma}^{uc}(\mathbf{\Sigma}^{old})$ is the result of the *unconstrained* update specified by (7). At each iteration, we take the updated value $\mathbf{\Sigma}^{new}$ to be the $\mathbf{\Sigma}$ which solves (11).

The remainder of this section explores various penalty choices. These are currently being implemented.

5.2.1. Entropy functional

Beginning with Frieden,[39] numerous authors[40–48] have proposed regularizing a variety of inverse problems with nonnegativity constraints via the entropy functional

$$\Phi_E(\mathbf{\Sigma}) = \sum_i \sigma_i \ln\sigma_i. \tag{12}$$

At each iteration, the new $\mathbf{\Sigma}$ is found by finding the zero of

$$-\sigma_i + \sigma_i^{uc}(\mathbf{\Sigma}^{old}) - \alpha\sigma_i^2(1+\ln\sigma_i) \tag{13}$$

for each i. This is convenient since the solution is decoupled from pixel to pixel. Molina and Ripley[49] suggest that entropy "corresponds to a rather peculiar prior, since it depends only on the marginal distribution of greylevels and not on their spatial locations. It is thus surprising that maximum entropy solutions appear smooth in many published examples." Donoho, Johnstone, Hoch, and Stern[50] offer a highly practical discussion of the how entropy regularization operates in practice. They suggest that nonlinearities of the form induced by (13) encourage a "shrinking" of estimates towards a nominal value of $1/e$.

Borden[51] proposed a generalization of this maximum entropy formulation for regularizing estimates of complex reflectances in radar imaging. Since (12) is inappropriate for complex imagery, he instead employs the cross-entropy between densities parameterized by the reflectance.

5.2.2. Good's roughness

Good's roughness penalty[52] was originally formulated for smoothing estimates in nonparametric probability density estimation; a thorough analysis in this context is given by Tapia and Thompson.[31] Following the suggestion of Snyder and Miller (Ref. 32, Sec. II.1), Good's roughness was later applied to closely related problems of Poisson intensity estimation in PET,[53] SPECT,[54–57] and optical sectioning microscopy.[16]

In one dimension, Good's continuous first-order roughness penalty may be written in several equivalent ways:

$$4\int\left[\frac{d}{d\tau}\sqrt{\sigma(\tau)}\right]^2 d\tau = \int\sigma(\tau)\left[\frac{d}{d\tau}\ln\sigma(\tau)\right]^2 d\tau = \int\frac{\sigma'(\tau)}{\sigma(\tau)}d\tau = -\int\sigma(\tau)\frac{d^2}{d\tau^2}\ln\sigma(\tau)d\tau \tag{14}$$

The last equality in (14), established in Ref. 58, follows from analogy with steps in the proof the Cramer-Rao bound. Consider the rightmost expression. The penalty is straightforwardly extended to two dimensions (see pp. 155-156 of 34 or Sec. 3 of 53):

$$-\int\int \sigma(\tau,f)\left(\frac{\partial^2}{\partial\tau^2}+\frac{\partial^2}{\partial f^2}\right)\ln\sigma(\tau,f)d\tau df. \tag{15}$$

Discretizing (15) yields

$$\Phi_G(\sigma) = -\sum_{\ell,k}\sigma_{\ell,k}[\ln\sigma_{\ell+1,k}+\ln\sigma_{\ell-1,k}+\ln\sigma_{\ell,k+1}+\ln\sigma_{\ell,k-1}-4\ln\sigma_{\ell,k}]. \tag{16}$$

As noted by O'Sullivan,[58] the discretized penalty (16) has an appealing information-theoretic interpretation in terms of I-divergences between neighboring pixel values.

At each iteration, the new $\mathbf{\Sigma}$ may be found by solving a set of nonlinear difference equations:

$$\begin{aligned}-\frac{1}{\sigma_{\ell,k}}+\frac{\sigma^{un}_{\ell,k}(\mathbf{\Sigma}^{old})}{\sigma^2_{\ell,k}} \quad &+ \quad \alpha[(\ln\sigma_{\ell+1,k}+\ln\sigma_{\ell-1,k}+\ln\sigma_{\ell,k+1}+\ln\sigma_{\ell,k-1}-4\ln\sigma_{\ell,k})\\ &\quad +\frac{1}{\sigma_{\ell,k}}(\sigma_{\ell+1,k}+\sigma_{\ell-1,k}+\sigma_{\ell,k+1}+\sigma_{\ell,k-1}-4\sigma_{\ell,k})]=0.\end{aligned} \tag{17}$$

5.2.3. Silverman's roughness

Inspired by the work of Good and Gaskins,[52] Silverman[59] suggested alternative penalties which employ differential operators of the logarithm[§] of the function. Like Good's roughness, Silverman proposed his penalty in the context of density estimation. We consider the special case

$$\int\int\left[\left(\frac{\partial}{\partial\tau}+\frac{\partial}{\partial f}\right)\ln\sigma(\tau,f)\right]^2 d\tau df. \tag{18}$$

Discretizing (18) yields

$$\Phi_S(\sigma) = \sum_{\tau,f}[(\ln\sigma_{\ell+1,k}-\ln\sigma_{\ell,k})^2+(\ln\sigma_{\ell,k+1}-\ln\sigma_{\ell,k})^2]. \tag{19}$$

Using this penalty requires solving the following set of nonlinear equations at each iteration of the EM algorithm:

$$-\frac{1}{\sigma_{\ell,k}}+\frac{\sigma^{un}_{\ell,k}(\mathbf{\Sigma}^{old})}{\sigma^2_{\ell,k}}+\alpha\frac{2}{\sigma_{\ell,k}}[\ln\sigma_{\ell+1,k}+\ln\sigma_{\ell-1,k}+\ln\sigma_{\ell,k+1}+\ln\sigma_{\ell,k-1}-4\ln\sigma_{\ell,k}]=0 \tag{20}$$

6. DIRECTIONS FOR FUTURE WORK

One of the main difficulties with EM algorithms in general is their slow convergence rate. A variety of EM variants have been proposed which boast faster convergence. For instance the Space-Alternating Generalized EM (SAGE) algorithm of Fessler and Hero[60] updates the parameters in groups instead of all at once; each group has its own associated *hidden data space*, which would be a complete data space if the remaining parameters were known. We are exploring applying SAGE to the radar imaging problem.

The most formidable obstacle to making the EM algorithm practical in this particular application is the computational complexity of inverting $\mathbf{K}$ in (7). This was the primary reason for the exploration of small 20×20 images in Sec. 4. Table 1 presents approximate computation times for our MATLAB implementation running on a Sun Enterprise 3500. The MATLAB code is highly vectorized, so it runs relatively efficiently; we do not expect substantial gains to be made by, for instance, recoding it in C.

[§]In the realm of astronomical imaging, Molina and Ripley[49] suggest smoothing the logarithm of the image, noticing that "astronomers tend to look at the raw data on a logarithmic scale." The same observation often seems true of radar engineers.

Image size	Total time	Time for inverse
20 x 20	15 seconds	6 seconds
32 x 32	8 minutes	4 minutes
40 x 40	32 minutes	13 minutes

Table 1. Computation time for one step of the EM algorithm, assuming that the collected data vector equals the number of pixels in the image.

As discussed in Ref. 61, $\mathbf{K}$ may have some structure which can aid in crafting faster inverse algorithms, although it would be preferable to avoid having to compute a direct, brute-force inverse altogether. One possible approach would be to exploit the fact that since $\mathbf{\Sigma}$ will not change drastically from iteration to iteration, $\mathbf{K}$ is unlikely to change drastically either, so the $\mathbf{K}$ from a previous EM iteration might be helpful as an initial value in an inner iteration which finds a new $\mathbf{K}$.

An appealing aspect of the statistical formulation is that it allows the prediction of estimator performance via Cramer-Rao bounds. This provides a method for selecting the transmitted waveform in various applications.[62] For large images, inverting the Fisher information can become cumbersome; Hero and Fessler[63] propose an iterative algorithm for computing Cramer-Rao bounds on parameter subsets which avoids explicit inversion of the Fisher information matrix via a complete/incomplete data formulation analogous with the EM algorithm.

ACKNOWLEDGMENTS

This work was supported by a grant from DARPA under Contract F49620-98-1-0498, administered by AFOSR.

REFERENCES

1. D. Snyder, J. O'Sullivan, and M. Miller, "The use of maximum-likelihood estimation for forming images of diffuse radar-targets from delay-doppler data," *IEEE Trans. on Information Theory* **35**, pp. 536–548, May 1989.
2. A. D. Dempster, N. M. Laird, and D. B. Rubin, "Maximum likelihood from incomplete data via the EM algorithm," *Journal of the Royal Statistical Society B* **39**, pp. 1–37, 1977.
3. G. McLachlan and T. Krishnan, *The EM Algorithm and Extensions*, John Wiley and Sons, New York, 1996.
4. M. Bernfeld, "Chirp doppler radar," *Proc. of the IEEE* **72**, pp. 540–541, April 1984.
5. D. Snyder, H. Whitehouse, J. T. Wohlschlaeger, and R. Lewis, "A new approach to radar/sonar imaging," in *Advanced Algorithms and Architectures*, vol. SPIE Proc. 696, (San Diego, CA.), August 1986.
6. E. Feig and F. Grünbaum, "Tomographic methods in range-Doppler radar," *Inverse Problems* **10**, pp. 793–808, 1994.
7. Y. Dai, E. Rothwell, K. Chen, and D. Nyquist, "Time-domain imaging of radar targets using algorithms for reconstruction from projections," *IEEE Trans. on Antennas and Propagation* **45**, pp. 1227–1235, August 1997.
8. M. V. Blaricum and R. Mittra, "A technique for extracting the poles and residues of a system directly from its transient response," *IEEE Trans. on Antennas and Propagation* **23**, pp. 777–781, November 1975.
9. J. Costas, "A study of a class of detection waveforms having nearly ideal range-doppler ambiguity properties," *Proc. of the IEEE* **72**, pp. 996–1530, August 1994.
10. H. L. Van Trees, *Detection, Estimation and Modulation Theory, Part III: Radar-Sonar Signal Processing and Gaussian Signals in Noise*, John Wiley and Sons, New York, 1971.
11. L. Guosui, G. Hong, and S. Weimin, "Development of random signal radars," *IEEE Trans. on Aerospace and Electronic Systems* **35**, pp. 770–776, July 1999.
12. A. Lanterman, "Tracking and recognition of airborne targets via commercial television and fm radio signals," in *Acquisition, Tracking, and Pointing XIII*, M. Masten and L. Stockum, eds., vol. SPIE Proc. 3692, pp. 189–198, (Orlando, FL), April 1999.
13. L. A. Shepp and Y. Vardi, "Maximum-likelihood reconstruction for emission tomography," *IEEE Trans. on Medical Imaging* **MI-1**, pp. 113–121, 1982.
14. M. I. Miller, D. L. Snyder, and T. R. Miller, "Maximum likelihood reconstruction for single photon emission computed tomography," *IEEE Trans. on Nuclear Science* **NS-32,No.1**, pp. 769–778, 1985.

15. D. L. Snyder, A. M. Hammoud, and R. L. White, "Image recovery from data acquired with a charge-coupled-device camera," *J. Optical Society of America A* **10**, pp. 1014–1023, 1993.
16. S. Joshi and M. Miller, "Maximum a Posteriori estimation with Good's roughness for optical sectioning microsopy," *Optical Society of America A* **10**, pp. 1078–1085, May 1993.
17. P. Moulin, J. O'Sullivan, and D. Snyder, "A method of sieves for multiresolution spectrum estimation and radar imaging," *IEEE Trans. on Information Theory* **38**, pp. 801–813, March 1992.
18. J. O'Sullivan, D. Snyder, D. Porter, and P. Moulin, "An application of splines to maximum likelihood radar imaging," *International Journal of Imaging Systems and Technology* **4**, pp. 256–264, 1992.
19. J. Shapiro, B. Capron, and R. Harney, "Imaging and target detection with a heterodyne-reception optical radar," *Applied Optics* **20**, pp. 3292–3313, October 1981.
20. J. Burg, D. Luenberger, and D. Wenger, "Estimation of structured covariance matrices," *Proc. of the IEEE* **70**, pp. 963–974, Sept. 1982.
21. T. Marzetta, "EM algorithm for estimating the parameters of a multivariate complex Rician density for polarimetric SAR," in *Proceedings of the International Conference on Acoustice, Speech, and Signal Processing*, vol. 5, pp. 3651–3654, 1995.
22. M. DeVore, A. Lanterman, and J. O'Sullivan, "ATR performance of a Rician model for SAR images," in *Automatic Target Recognition X*, F. Sadjadi, ed., vol. SPIE Proc. 4050, (Orlando, FL), April 2000.
23. A. Lanterman, "Maximum-likelihood estimation for hybrid specular/diffuse models of radar imaging and target recognition," *IEEE Trans. on Aerospace and Electronic Systems* , to be submitted.
24. C. Mallows, "Linear processes are nearly gaussian," *Journal of Applied Probability* **4**, pp. 313–329, 1967.
25. P. Moulin, *A Method of Sieves for Radar Imaging and Spectrum Estimation*, D.Sc. Dissertation, Dept. of Electrical Engineering, Sever Institute of Technology, Washington Univ., St. Louis, MO, May 1990.
26. J. D.C. Munson, J. O'Brien, and W. Jenkins, "A tomographic formulation of spotlight-mode synthetic aperture radar," *Proceedings of the IEEE* **71**, pp. 917–925, August 1983.
27. J. O'Sullivan, R. Blahut, and D. Snyder, "Information-theoretic image formation," *IEEE Trans. on Information Theory* **44**, pp. 2094–2123, October 1998.
28. L. Scharf, *Statistical Signal Processing: Detection, Estimation and Time Series Analysis*, Addison-Wesley, Reading, MA, 1991.
29. V. Poor, *An Introduction to Signal Detection and Estimation*, Spinger-Verlag, New York, NY, second ed., 1994.
30. P. Green, "Radar astronomy measurement techniques," Technical Report 282, Lincoln Laboratory, Lexington, MA, Dec. 1962.
31. R. A. Tapia and J. R. Thompson, *Nonparametric probability density estimation*, Johns Hopkins University Press, Baltimore,Md., 1978.
32. D. Snyder and M. Miller, "The use of sieves to stabilize images produced with the EM algorithm for emission tomography," *IEEE Trans. on Nuclear Science* **32**, pp. 3864–3872, 1985.
33. D. Snyder, M. Miller, J. L.J. Thomas, and D. Politte, "Noise and edge artifacts in maximum-likelihood reconstruction for emission tomography," *IEEE Trans. on Medical Imaging* **6**(3), pp. 228–237, 1987.
34. D. L. Snyder and M. I. Miller, *Random Point Processes in Time and Space*, Springer-Verlag, 2nd ed., 1991.
35. U. Grenander, *Abstract Inference*, John Wiley and Sons, New York, 1981.
36. D. Politte and D. Snyder, "The use of constraints to eliminate artifacts in maximum-likelihood image estimation for emission tomography," *IEEE Trans. on Nuclear Science* **35**, pp. 608–610, February 1988.
37. D. G. Politte and D. L. Snyder, "Corrections for accidental coincidences and attenuation in maximum-likelihood image reconstruction for positron-emission tomography," *IEEE Trans. on Medical Imaging* **10**, pp. 82–89, March 1991.
38. P. Moulin, "A wavelet regularization method for diffuse radar-target and speckle-noise reduction," *Journal of Mathematical Imaging and Vision* **3**(1), pp. 123–134, 1993.
39. B. R. Frieden, "Restoring with maximum likelihood and maximum entropy," *J. Opt. Soc. Am.* **62, No. 4**, pp. 511–518, 1972.
40. *Maximum-Entropy and Bayesian Methods in Science and Engineering*, Kluwer Academic, Dordrecht, 1988.
41. P. Eggermont, "Maximum entropy regularization for Fredholm integral equations of the first kind," *SIAM Journal of Mathematical Analysis* **24**, pp. 1557–1576.

42. U. Amato and W. Hughes, "Maximum entropy regularization of Fredholm integral equations of the first kind," *Inverse Problems* **7**, pp. 793–808, 1991.
43. L. Mead, "Approximate solution of fredholm integral equations by the maximum entropy method," *Journal of Mathematical Physics* **27**, pp. 2903–2907, 1986.
44. J. Skilling, A. W. Strong, and K. Bennett, "Maximum-entropy image processing in gamma-ray astronomy," *Mon. Not. R. Astr. Soc.* **187**, pp. 145–152, 1979.
45. J. Skilling and R. Bryan, "Maximum entropy image reconstruction: general algorithm," *Mon. Not. R. Astr. Soc.* **211**, pp. 111–124, 1984.
46. B. Frieden and D. C. Wells, "Restoring with maximum entropy. iii. poisson sources and backgrounds," *J. Opt. Soc. Am.* **68, No. 1**, pp. 93–103, 1978.
47. S. Gull and G. Daniell, "The maximum entropy algorithm applied to image enhancement," *Proc. of the IEEE* **5**, p. 170, 1980.
48. S. Wernecke and L. D'Addario, "Maximum entropy image reconstruction," *IEEE Trans. on Computers* **26**, pp. 351–364, 1977.
49. R. Molina and B. Ripley, "Using spatial models as priors in astronomical image analysis," *Journal of Applied Statistics* **16**(2), pp. 193–206, 1989.
50. D. Donoho, I. Johnstone, J. Hoch, and A. Stern, "Maximum entropy and the nearly black object," *Journal of the Royal Statistical Society B* **54**(1), pp. 41–81, 1992.
51. B. Borden, "Maximum entropy regularization in synthetic aperture radar imagery," *IEEE Trans. on Signal Processing* **40**, pp. 969–973, April 1992.
52. I. J. Good and R. A. Gaskins, "Nonparametric roughness penalties for probability densities," *Biometrika* **58,2**, pp. 255–277, 1971.
53. M. Miller and B. Roysam, "Bayesian image reconstruction for emission tomography incorporating Good's roughness prior on massively parallel processors," *Proc. of the Natl. Acad. of Sci.* **88**, pp. 3223–3227, April 1991.
54. A. McCarthy and M. Miller, "Maximum likelihood SPECT in clinical computation times using mesh-connected parallel computers," *IEEE Trans. on Medical Imaging* **10**, pp. 426–436, September 1991.
55. C. Butler and M. Miller, "Maximum a posteriori estimation for SPECT using regularization techniques on massively-parallel computers," *IEEE Trans. on Medical Imaging* **12**, pp. 84–89, March 1993.
56. M. Miller and C. Butler, "3-D maximum a posteriori estimation for SPECT on massively-parallel computers," *IEEE Trans. on Medical Imaging* **12**(3), pp. 560–565, September 1993.
57. C. S. Butler, M. Miller, T. R. Miller, and J. W. Wallis, "Massivel parallel computers for 3D single-photon-emission computed tomography," *Physics in Medicine and Biology* **39**, pp. 575–582, 1994.
58. J. O'Sullivan, "Roughness penalties on finite domains," *IEEE Trans. on Image Processing* **4**, pp. 1258–1268, September 1995.
59. B. W. Silverman, "On the estimation of a probability density function by the maximum penalized likelihood method," *The Annals of Statistics* **10**(3), pp. 795–810, 1982.
60. J. Fessler and A. Hero, "Space-alternating generalized expectation-maximization algorithm," *IEEE Trans. on Signal Processing* **42**, pp. 2664–2677, October 1994.
61. J. O'Sullivan, P. Moulin, D. Snyder, and S. Jacobs, "Computational considerations for maximum-likelihood radar imaging," in *Proc. of the Conf. on Information Sciences and Systems*, (Baltimore, MD), March 1990.
62. J. O'Sullivan, P. Moulin, and D. Snyder, "Cramer-rao bounds for constrained spectrum estimation with application to a problem in radar imaging," in *Proceedings of the 26th Allerton Conference on Communication, Control, and Computing*, Univ. of Illinois at Urbana-Champaign, (Urbana, IL), October 1988.
63. A. Hero and J. Fessler, "A recursive algorithm for computing Cramer-Rao-type bounds on estimator covariance," *IEEE Trans. on Information Theory* **40**, pp. 1205–1210, July 1994.