

Model-Based Estimation Techniques for 3-D Reconstruction from Projections

Yoram Bresler

Coordinated Science Laboratory, University of Illinois at Urbana-Champaign, Urbana, IL 61821 USA

Jeffrey A. Fessler* and Albert Macovski

Information Systems Laboratory, Department of Electrical Engineering, Stanford University, Stanford, CA 94305 USA

Abstract: A parametric estimation approach to reconstruction from projections with incomplete and very noisy data is described. Embedding prior knowledge about “objects” in the probed domain and about the data acquisition process into stochastic dynamic models, we transform the reconstruction problem into a computationally challenging nonlinear state-estimation problem, where the objects’ parametrized descriptions are to be directly estimated from the projection data. This paper is a review in a common framework and a comparative study of two distinct algorithms which were developed recently for the solution of this problem. The first, is an approximate, globally optimal minimum-mean-square-error recursive algorithm. The second is a hierarchical suboptimal Bayesian algorithm. Simulation examples demonstrate accurate reconstructions with as few as four views in a 135° sector, at an average signal to noise ratio of 0.6.

Key Words: 3D tomography, incomplete projections, Bayesian estimation, stochastic modeling

I. Introduction

Tomography, or the reconstruction of a multidimensional function from its line-integral projections, is a well-studied problem, typically arising in the context of determining the internal structure of an object from the results of external probing by electromagnetic or sound waves, or by subatomic particles. The problem is usually posed and solved

in two dimensions, where a cross section, or a thin slice, is reconstructed from its projections. Most often, 3-D reconstruction is simply obtained by stacking thin reconstructed slices. The applications of reconstruction from projections are seen in diverse disciplines, ranging from medicine and non-destructive testing to geophysical exploration, and from astronomy to electron microscopy (Deans 1984, Herman 1980).

Owing to various temporal, physical, geometrical, or economic constraints in the data acquisition (Bresler and Macovski 1987, Rossi and Willsky 1984) it is often impossible to acquire projection data at all angles (views), and the number of views and/or rays within a view is severely restricted. This is almost always the case with 3-D reconstruction, where a complete data-set is exceedingly large. An attempt to reconstruct the original distribution in this so-called *incomplete data* case results in images that suffer from artifacts such as streaking and geometric distortion, poor resolution, and high noise level, because the problem is ill-posed, and in extreme cases, because of the inherent nonuniqueness of the solution (Louis and Natterer 1983). Consequently, although 3-D reconstruction would be an ideal tool in a variety of medical (Bloch and Udupa 1983) and other applications, it is rarely attempted in practice with limited data.

In this paper we consider the incomplete data case in the extreme situation when the data are restricted and heavily corrupted by noise to the point where all current limited data reconstruction methods (Gordon and Herman 1974; Kak 1979; Rangayan et al. 1984; Stark 1987) produce unacceptable results. (For example, we are unaware of a current method producing diagnostically useful reconstructions from four views at a signal-to-noise ratio

This work was supported by the National Institute of Health contract NO1-38045, by the National Science Foundation contract ECS-8213959, and General Electric contract 22-84.

* Supported by a National Science Foundation Graduate Fellowship.

(SNR) of 0 dB.) The original motivation for this work has been the 3-D reconstruction of systems of blood vessels (e.g., coronary arteries and the vasculature in the brain) from X-ray images (Bresler and Macovski 1984a), using minimally invasive imaging techniques. In particular, we seek quantitative information about the cross-section parameters, which are required for assessment of vessel disease and about their location, which is essential for the planning of surgical intervention. The problem then corresponds to 3-D reconstruction from several entire-volume projection images. Two constraints, both of which are dictated by the requirement for a minimally invasive procedure, contribute to the low signal-to-noise ratio (SNR) condition. First, the radiopaque dye that is used to produce contrast between the blood and surrounding tissue is injected intravenously (rather than intra-arterially, as in direct catheter angiography), resulting in low contrast. Second, the minimization of the radiation dose results in higher photon count noise. A similar situation prevails in the "ultimate noninvasive technique," Magnetic Resonance Angiography, where the various subtraction flow-imaging techniques (Nishimura et al. 1986, 1987) that are used produce noisy projection images of blood vessels.

Our approach is based on the observation that in many practical cases abundant a priori information on the imaged domain is available, and the ultimate goal is not to obtain an image of an arbitrary distribution, but rather to extract specific, quantitative information about objects in the probed volume. Examples (Rossi and Willsky 1984) include the detection and localization of organs and tumors in medical diagnostic CT (Selfridge and Prewitt 1981) the detection and tracking of high contrast thermal regions in oceans by oceanographic acoustic tomography (Munk and Wunsch 1979), and the detection and localization of interior cracks and flaws in materials, in the area of nondestructive testing. We describe parametric techniques overcoming the inherent underdeterminacy of the solution from the data by incorporating a priori information into the reconstruction. Stochastic modeling for both the measurement process and the three-dimensional objects in the probed domain is used to account for the associated uncertainty and in order to encompass a large class, an ensemble of objects rather than a single nominal version. The reconstruction problem is thus converted from an algebraic problem into a computationally challenging estimation problem, with operations to be performed directly on the projection data.

With a view toward operation with low SNR

data, we have recently developed two distinct algorithms, which take advantage of the structure of this estimation problem for its approximate optimal solution. The first is a special approximate minimum mean square error (MMSE) smoothing algorithm providing an object estimate that is globally optimal, in the sense that at each point in space, the reconstruction error is minimized by optimum use of all available measurements and prior information. The second, tailored for the case when multiple objects may be simultaneously present, is a globally suboptimal maximum a posteriori (MAP) algorithm, which relies on a hierarchical structure to achieve computational feasibility. In addition to the quantitative morphological information provided by the algorithms, the results can be used to synthesize a 3-D display of the reconstructed objects.

Previous work using a similar philosophy has been limited, for the most part, either conceptually (Hanson and Wecksung 1983, Rossi and Willsky 1984) or by computational feasibility (Shmueli et al. 1981), to a 2-D case [see also the discussion in (Bresler and Macovski 1987).]

The purpose of this paper is twofold. First, it reviews in a common framework the algorithms, whose separate detailed derivations can be found in Bresler (1985) and Bresler and Macovski (1987). Second, the paper compares the performance of these two algorithms and a third, less sophisticated algorithm, on the same simulation example to determine their relative merits, and to verify the actual need for sophisticated near-optimal processing.

II. Statistical Models

Object Model

The probed region in 3-D space is the cylinder $D = \{(\xi, \eta, z) : \xi^2 + \eta^2 \leq (T/2)^2, 0 \leq z \leq L\}$, which is assumed to have a known background density which we set, without loss of generality (Bresler and Macovski 1987) to zero. We assume that $D \subset \mathcal{R}^3$ contains J distinct 3-D objects $O_j, j = 1, 2, \dots, J$, each represented by a real function $f_j(\xi, \eta, z), j = 1, 2, \dots, J$, which is defined on D , and its value describes an additive component to the density in each point in space. Thus, the overall density is given by

$$f(\mathbf{r}, z) = \sum_{j=1}^J f_j(\mathbf{r}, z), \quad (1)$$

where $\mathbf{r} = (\xi, \eta)'$ represents the 2-D location vector in the plane. (We use lower and upper case bold-

face letters to denote vectors and matrices, respectively, and $(\cdot)'$ to denote transposition.)

We adopt a subclass of *generalized cylinders*, (Binford 1971) as a representation for the objects, providing an effective tradeoff between model complexity and richness. An object (Figure 1) is represented by a collection of vertically stacked unit height cylindrical sections, which we call *primitives* (Figure 2). The cross-sectional density of each primitive perpendicular to the (vertical) z -axis is $f_0(\mathbf{r}; \boldsymbol{\gamma})$, where f_0 is a *known* function of \mathbf{r} and of a vector of shape parameters $\boldsymbol{\gamma}$. For example, f_0 can be the indicator function on an elliptical support (assuming a value of 1 on its support) whose shape is specified by the parameter vector $\boldsymbol{\gamma} = [\alpha, \lambda, \phi]'$: α is the *radius* of the ellipse, defined as the geometrical mean of its semi-axes; λ is its *axis ratio*; and its *orientation* ϕ is the angle between its major axis and the ξ coordinate axis. Each object primitive is centered at a point $(\mathbf{c}_j(z), z)$ on a discrete trajectory $\mathbf{c}_j(z) \in \mathcal{R}^2$, $z = z_{1,j}, \dots, z_{1,j} + L_j$, which defines the *center axis* of an object O_j of vertical *starting position* $Z_{1,j}$ and *length* L_j (Figure 1). Thus, the cross-sectional density function of such a primitive is $f_0[\mathbf{r} - \mathbf{c}_j(z); \boldsymbol{\gamma}_j(z)]$.

The object model is completed by a rule assigning a probability measure to the different object realizations in the given class defined by f_0 . In particular, we represent prior knowledge about the smoothness and curvature properties of an object O_j by a discrete linear state-space dynamic stochastic Markov model for the primitive-to-primitive evolution of the position and cross-sectional parameters. Defining the state vector $\mathbf{x}_j(z) = [\mathbf{c}_j'(z), \boldsymbol{\gamma}_j'(z)]'$ ($\mathbf{x}_j(z)$

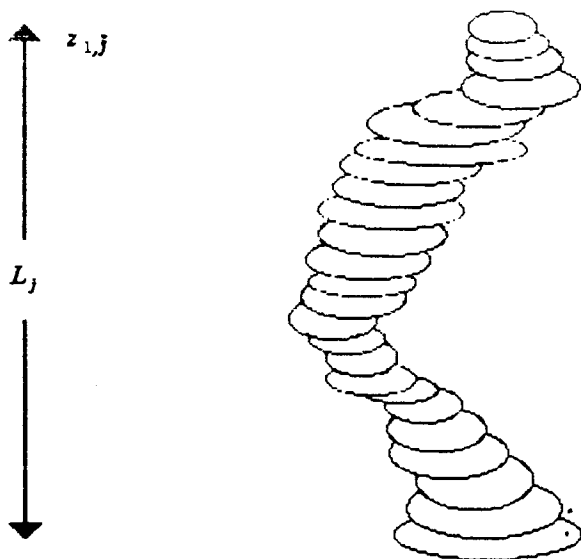


Figure 1. Object formed from stacked primitives.

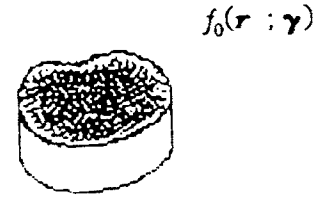


Figure 2. Unit height cylindrical primitive.

may be augmented to allow higher order evolution models) we have

$$\begin{aligned} \mathbf{x}_j(z + 1) &= \mathbf{A}\mathbf{x}_j(z) + \mathbf{B}\mathbf{w}_j(z), \quad z = z_{1,j}, \dots, z_{1,j} + L_j \\ E\mathbf{x}_j(z_{0,j}) &= \mathbf{x}_0 \quad \text{COV}\{\mathbf{x}(z_{0,j})\} = \boldsymbol{\Pi}_0 \end{aligned} \quad (2)$$

where $\mathbf{w}_j(z)$ is a white Gaussian noise vector-sequence of identity covariance, uncorrelated with $\mathbf{x}_j(z_{0,j})$. The different objects and their associated stochastic processes $\{\mathbf{x}_j(z), z = z_{1,j}, \dots, z_{1,j} + L_j\}$ are assumed statistically mutually independent implying that for $i \neq j$, $j = 1, \dots, J$, $\text{COV}\{\mathbf{x}_i(z_{0,i}), \mathbf{x}_j(z_{0,j})\} = \mathbf{0}$ and $E[\mathbf{w}_i(l)\mathbf{w}_j'(m)] = \mathbf{0} \forall l, m$. The objects are further assumed to have identical statistics, and are therefore modeled by identical state-space models. Diverse structures may be modeled by properly choosing $f_0(\cdot; \cdot)$, $\boldsymbol{\gamma}(\cdot)$ and dynamic model parameters (Bresler and Macovski 1987). The process noise $\mathbf{w}_j(z)$ and the random initial state $\mathbf{x}_j(z_{0,j})$ both represent the uncertainty about the exact shape and position of the j th object: While a nominal, deterministic object is given by the expected value $E \mathbf{x}_j(z)$, which is obtained by initializing equation (2) with \mathbf{x}_0 and setting $\mathbf{w}_j(z) = \mathbf{0}$ for all z , the actual object is some random variation on this theme.

The j th object is completely determined by the data set $O_j = \{z_{1,j}, L_j, \mathbf{X}_j\}$, where $\mathbf{X}_j \triangleq \{\mathbf{x}_j(z), z = z_{1,j}, \dots, z_{1,j} + L_j\}$ is its state sequence, and its 3-D density function (with the dependence on O_j made explicit) can be immediately recovered from this representation as

$$f_j(\mathbf{r}, z; O_j) = \sum_{l=z_{1,j}}^{z_{1,j}+L} f_0[\mathbf{r} - \mathbf{c}_j(z); \boldsymbol{\gamma}_j(z)]\delta_{l,z} \quad (3)$$

where $\delta_{k,j} = 1$ iff $j = k$ is the Kronecker delta function. Hence, by equations (1) and (3), the density function $f(\mathbf{r}, z; O_{1,j})$ of the entire domain is uniquely determined by the data set $O_{1,j} \triangleq \{O_j, j = 1, 2, \dots, J\}$, which we ultimately seek to estimate.

Measurement Model

The Radon Transform. The 2-D projection of the cross-sectional distribution $f(\mathbf{r})$ at given projection

angle θ and radial distance t from the origin is the integral

$$g(t, \theta) = \int_{-\infty}^{\infty} f(\mathbf{r})\delta(t - \theta' \mathbf{r})d\xi d\eta = \int_{\theta' \mathbf{r} = t} f(\mathbf{r})d\mathbf{r} \quad (4)$$

along the line $\mathbf{l}(t, \theta) = \{\mathbf{r}: \theta' \mathbf{r} = t\}$, where θ is the corresponding unit direction-vector $\theta \triangleq (\cos\theta, \sin\theta)'$. The projection $g(t, \theta)$ at any value of θ is a 1-D function of t (Figure 3). The mapping of f into g via equation (4) corresponds to the 2-D radon transformation (Deans 1984) whereas the reconstruction problem of determining the distribution f from its projections g involves inverting the integral equation (4), or finding the inverse radon transform.

As in 3-D axial tomography we assume that we are given sets of 2-D projection of vertically stacked thin slices of the domain D , for $z = 1, 2, \dots, L$. Denoting the radon transform of $f_0(\mathbf{r}; \gamma)$ by $\psi(t, \theta; \gamma)$, which is also a known function, it follows that the 2-D projection of a slice at a given height z containing a single primitive is

$$\int_{\theta' \mathbf{r} = t} f_0(\mathbf{r} - \mathbf{c}; \gamma)d\mathbf{r} = \psi(t - \theta' \mathbf{c}, \theta; \gamma). \quad (5)$$

By the linearity of the radon transform, it follows from equations (1), (3), and (5) that the radon transform of the J objects is given by

$$\begin{aligned} \Psi(t, \theta, z; O_{1..J}) &\triangleq \int_{\theta' \mathbf{r} = t} f(\mathbf{r}, z; O_{1..J})d\mathbf{r} \\ &= \sum_{j=1}^J \sum_{l_j=z_{1j}}^{z_{1j}+L} \psi[t - \theta' \mathbf{c}_j(z), \theta; \gamma_j(z)]\delta_{z,l_j} \end{aligned} \quad (6)$$

Note that at any z position, $\Psi(\cdot)$ consists of the superposition of the projections of at most J primitives. Neglecting for simplicity the effect of a finite imaging aperture, which may be easily taken into

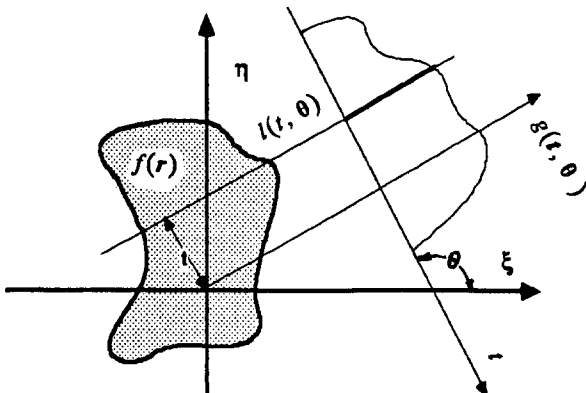


Figure 3. 2-D projection geometry.

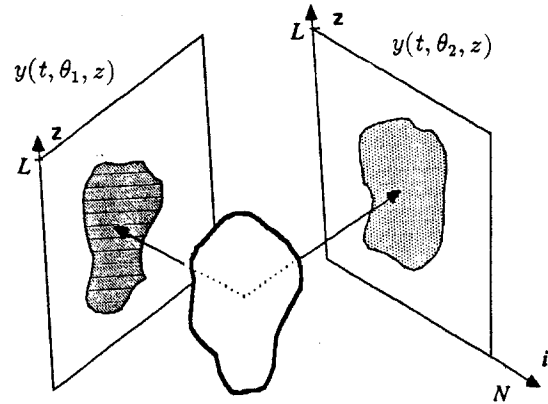


Figure 4. Classical radiograph interpretation of $y(t, \theta, z)$.

account (Bresler and Macovski 1987), the actual noisy projection measurements are modeled by

$$y(t, \theta, z) = \Psi(t, \theta, z; O_{1..J}) + v(t, \theta, z), \quad (7)$$

where $v(t, \theta, z)$ is a white noise zero-mean Gaussian¹ random field, of intensity N_0 .

In the limited data case, the projection data is available only at a discrete set $\theta \in \{\theta_m\}_{m=1}^M$ of M (possibly nonuniformly spaced) projection angles covering a sector $\theta_M - \theta_1 \leq \pi$, at N uniformly spaced values of t covering the range $[-T/2, T/2]$, and at L values of z spaced at the primitive height interval of 1. These values of (t, θ, z) define the support set S_y of the available measurements. For a fixed $\theta = \theta_m$, $y(t, \theta_m, z)$ may be regarded as a classical X-ray (radiograph) (Figure 4). Alternatively, using the indices n and m to index the discrete t and θ values, respectively, the measurements can be represented for each value of z by an $N \times M$ data matrix $Y(z) = [y_{nm}(z)]$ with elements given by the samples of $y(\cdot)$:

$$\begin{aligned} y_{nm}(z) &\triangleq y\left(-\frac{T}{2} + n\frac{T}{N}, \theta_m, z\right) = \Psi_{nm}(O_{1..J}) + \mu_{nm}(z), \\ &n = 0, \dots, N - 1 \\ &m = 1, \dots, M \end{aligned} \quad (8)$$

$$E[v_{n_1, m_1}(z_1) \cdot v_{n_2, m_2}(z_2)] = \sigma^2 \delta_{n_1, n_2} \delta_{m_1, m_2} \delta_{z_1, z_2}$$

¹ The modeling of the noise as additive and Gaussian is accurate in a variety of applications including MRI (Edelstein et al. 1983), and even when the true statistics of the noise are Poisson, as in X-ray imaging, the Gaussian approximation is often adequate. Furthermore, the Gaussian assumption can be dispensed with altogether in the hierarchical algorithm which is readily adapted to any known noise probability distribution, whereas the MMSE algorithm only requires that the Fourier transform of the noise field be Gaussian. The latter property applies asymptotically to any noise distribution with finite moments (Brillinger 1980).

The m th column of $Y(z)$ corresponds to a single scan line at height z in the radiograph at view angle θ_m .

III. The Estimation Problem

The object reconstruction problem may now be stated in the following form:

Given the noisy set of projections $\mathbf{y} \triangleq \{y(t, \theta, z), (t, \theta, z) \in S_y\}$ and the state space models from equation (2), determine optimal estimates \hat{J} and $\hat{O}_{1:\hat{J}}$ of the number of objects and their representations, respectively.

The problem is one of joint estimation and detection, since both the number of objects J and the number of primitives of which they comprise (i.e., their lengths L_j), as well as their state sequences \mathbf{X}_j , need to be estimated.

The problem is reduced to estimation only, if we assume a single object ($J = 1$) extending from bottom to the top of the imaged domain (i.e., $z_1 = 1, L_1 = L$). $\Psi(\cdot)$ in equation (7) is then replaced by $\psi[t - \theta'c(z); \gamma(z)]$. In this simplified case, stacking up the elements of $Y(z)$, for example, in lexicographical order into a column vector $\boldsymbol{\mu}(z)$ of length $N \cdot M$ (where the $t - \theta$ samples of $\psi(\cdot)$ and $v(\cdot)$ are similarly stacked into $\mathbf{h}[x(z)]$ and $\mathbf{v}(z)$, respectively), produces a vector nonlinear measurement equation with white additive noise.

$$\boldsymbol{\mu}(z) = \mathbf{h}[x(z)] + \mathbf{v}(z) \quad (9a)$$

$$E[\mathbf{v}(z_1)\mathbf{v}'(z_2)] = \sigma^2 \cdot \mathbf{I} \delta_{z_1, z_2} \quad (9b)$$

Note that $\mathbf{h}[x(z)]$, whose elements are the $t - \theta$ samples of $\psi[t - \theta'c(z); \gamma(z)]$ is a known nonlinear vector function of $c(z)$ and $\gamma(z)$, and hence of the state $x(z)$. Combining equation (9) with equation (2) and with the minimum mean square error (MMSE) criterion, the estimation problem for $x(z)$ takes on a standard form of a nonlinear state estimation problem. It is a *fixed interval smoothing* problem (Jazwinski 1970) since the estimate at each point $1 \leq z \leq L$ is formed using the measurement data for the entire interval, which is available simultaneously. While the optimum MMSE estimator is known to be given, in principle, by the conditional mean of the state given the measurements, its exact computation is not a finite process. Even an approximate solution by the widely used Extended Kalman Filter (EKF) (Jazwinski 1970) can not be obtained, owing to the severe nonlinearity of the measurements, which leads to filter divergence (Bresler and Kailath 1987). A special algorithm that is based on the EKF but yields extended convergence range is briefly reviewed below.

The abovementioned algorithm is restricted to only a few objects, owing to certain observability limitations. Our approach to the general multiple object detection-estimation problem is therefore different. We consider the *maximum a posteriori probability* (MAP) (Jazwinski 1970) estimates \hat{J} and $\hat{O}_{1:\hat{J}}$ defined as the solution to

$$\max_{J, O_{1:J}} p(J, O_{1:J} | \mathbf{y}), \quad (10)$$

where $p(J, O_{1:J} | \mathbf{y})$ is the posterior probability of J and $O_{1:J}$ given that the data \mathbf{y} were observed. Once more, while the solution is well defined, its actual computation is unfeasible for all but the smallest problems, owing to the high dimensionality of the search space. Consequently, we propose a suboptimal hierarchical algorithm, consisting of an optimal combination of individually optimal steps.

IV. Single Object Estimation Algorithm

The essence of this method lies in a 1-D Fourier transformation (FT) of the projection data from the spatial domain (indexed by t) to a frequency domain, and then a reordering of these measurements, to yield a new set of measurements that is more amenable to linearization, permitting the application of the EKF. The 1-D FT of the continuous (in t) measurement function, (with the z dependence suppressed for notational compactness) is

$$\begin{aligned} \psi[j\Omega, x, \theta] &\triangleq \text{FT}\{\psi[t - c'\theta, \theta; \gamma]\} \\ &= \psi_0(j\Omega, \theta, \gamma) \cdot e^{-j\Omega c'\theta} \end{aligned} \quad (11)$$

where $\psi_0(j\Omega, \theta; \gamma)$ the FT of $\psi(t, \theta; \gamma)$, is a known function. Note that the shift dependence has been transformed into an exponential phase factor,

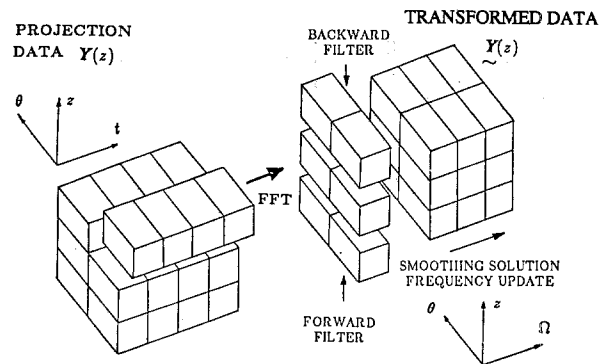


Figure 5. Data processing sequence for FSRS. Individual cuboids represent scalar samples. Slabs are processed as vectors.

transformed into an exponential phase factor, which can be effectively linearized in a neighborhood $\tilde{\mathbf{x}}$ of $\mathbf{x}(z)$, satisfying $|\Omega\tilde{\mathbf{x}}'\boldsymbol{\theta}_k| \ll 1$, where $\tilde{\mathbf{c}}$ is the corresponding neighborhood of the position \mathbf{c} . An EKF applied to the transformed measurements can be expected to converge for the range of estimate errors $\tilde{\mathbf{x}}$ satisfying the aforementioned requirement, if we assume that the “shift dependence,” (which has been transformed into the exponential), is indeed the major source of nonlinearity in the measurement. Hence, we term the $\tilde{\mathbf{x}}$ neighborhood satisfying $|\Omega\tilde{\mathbf{c}}'\boldsymbol{\theta}_k| \ll 1$ the *convergence region* of the filter at the measurement frequency Ω .

Clearly, low-frequency components (small Ω) imply larger convergence region than the higher components. This ordering is key to our scheme: Starting with the lowest frequency component of the measurement, (hence with a wide convergence region) use an EKF smoothing algorithm to obtain an estimate of the state for all $z = 1, \dots, L$. Next, linearizing about this estimate, run the EKF smoother to process the data of the next frequency component for all $z = 1, \dots, L$. As long as the estimation error is sufficiently reduced by the processing of one frequency component data-set so that it is within the convergence range of the next higher component, this process can be repeated without EKF divergence. The resulting estimate when all frequency components have been used, is, to within the approximation involved in the EKF linearization, the MMSE estimate. It is globally optimal, in the sense that at each point in space, the reconstruction error is minimized by optimum use

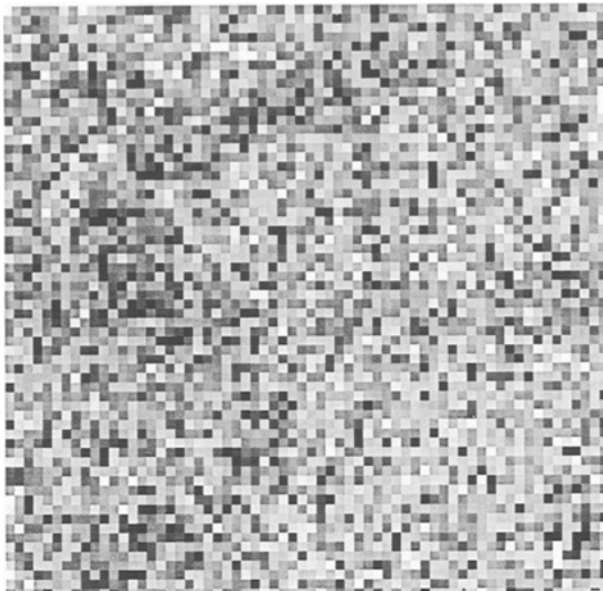


Figure 6. Noisy projection measurement at $\theta = 0^\circ$.

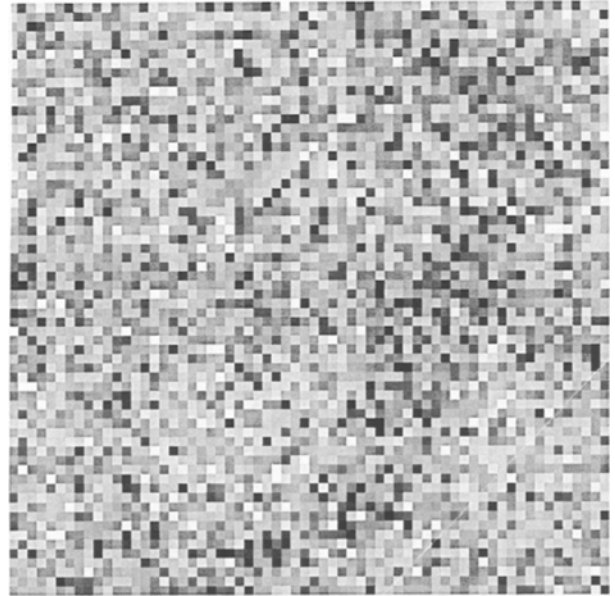


Figure 7. Noisy projection measurement at $\theta = 90^\circ$.

of all available measurements and prior information.

Such issues, as the replacement of the FT by a discrete FT (DFT) for the sampled measurements; the derivation of a recursive update scheme for a complete EKF smoothing solution to include an additional measurement set at the higher frequency; and convergence analysis of the algorithm, are addressed elsewhere (Bresler 1985, Bresler and Kailath 1987, Bresler and Macovski 1987). Here, we show in Figure 5 the overall structure of the resulting algorithm. In the preparatory step of the algorithm, a set of 1-D FFTs of the projection data is performed by columns of the measurement matrix $\mathbf{Y}(z)$, that is, separately for each view angle θ_m , transforming the radial index n to the frequency index u . The smoothing algorithm itself consists of a set of nested recursions. The inner recursion is in the spatial index z , where for a given frequency component u (a given vertical layer of transformed data in Figure 5) forward and backward filters are computed respectively for $z = 1, \dots, L$ and $z = L - 1, \dots, 0$, and then combined into the smoothing solution $\hat{\mathbf{x}}_{s,u}(z)$ for $L = 0, \dots, L$. The outside recursion is in u , where successive vertical layers of transformed data in Figure 5 are used to update the complete smoothing solution for $u = 0, \dots, N/2$. This recursive structure of the algorithm in a hybrid space-frequency domain (Fig. 5) has suggested its acronym FSRS (Frequency-Space Recursive Smoother).

Noting that higher-frequency components provide information about the finer structure of the data, the FSRS can also be interpreted as a multi-

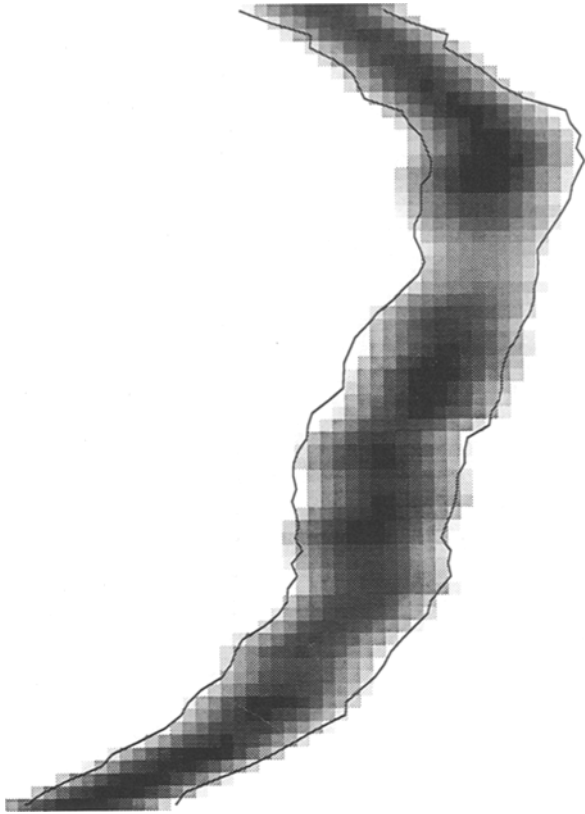


Figure 8. Noiseless object projection (shaded) with superimposed projection of FSRs-estimated object (solid line), $\theta = 0^\circ$.

resolution scheme, whereby the data are examined at increasing levels of resolution. It differs, however, from such schemes that have been proposed in various hierarchical computer vision algorithms in a fundamental way. First, it is recursive rather than hierarchical. Thus, no information is lost by proceeding in frequency order, the final result being an optimal combination of the data at all resolution levels. Second, rather than being postulated to reduce the computational load, the multiresolution structure here serves the purpose of increasing the domain of validity of the EKF-type linearization. Nonetheless, similarly to other multiresolution techniques, the FSRs also affords a trade-off between computation and estimate accuracy: The computation may be discontinued at $u < N/2$ short of using all the nonredundant frequency components, resulting in an object estimate that is less than optimum. In fact, since most of the energy of typical projection functions $\psi(\cdot)$ is concentrated in the low-frequency components (e.g., $|\psi_0(j\Omega, \theta, \gamma)|^2$ for an elliptical cross-section function falls off as Ω^{-3}), such truncated computation offers significant computational savings with little degradation in estimate accuracy.

V. Multiple Object Hierarchical Algorithm

The algorithm produces estimates of the objects' positions and shapes in four hierarchical steps.

Step 1: Primitive Detection and Estimation

First, we detect object primitives in each slice (normalized to thickness 1) and compute local maximum likelihood (ML) estimates of their parameters using only the measurements for that slice. Hence, the problem reduces to L decoupled 2-D problems, each requiring the maximization with respect to $\{\mathbf{x}_j\}$ of likelihood function

$$\begin{aligned} L(\{\mathbf{x}_j\}_{j=1}^K) &\triangleq \log p[\mathbf{Y}|\{\mathbf{x}_j\}_{j=1}^K] \\ &= C_1 - \frac{1}{2\sigma^2} \sum_{m=1}^M \sum_{n=1}^N (y_{nm} - \\ &\quad \sum_{j=1}^K \psi \left[-\frac{T}{2} + n \frac{T}{N} - \boldsymbol{\theta}_k' \mathbf{c}_j, \theta_k; \gamma_j \right])^2 \quad (12) \end{aligned}$$

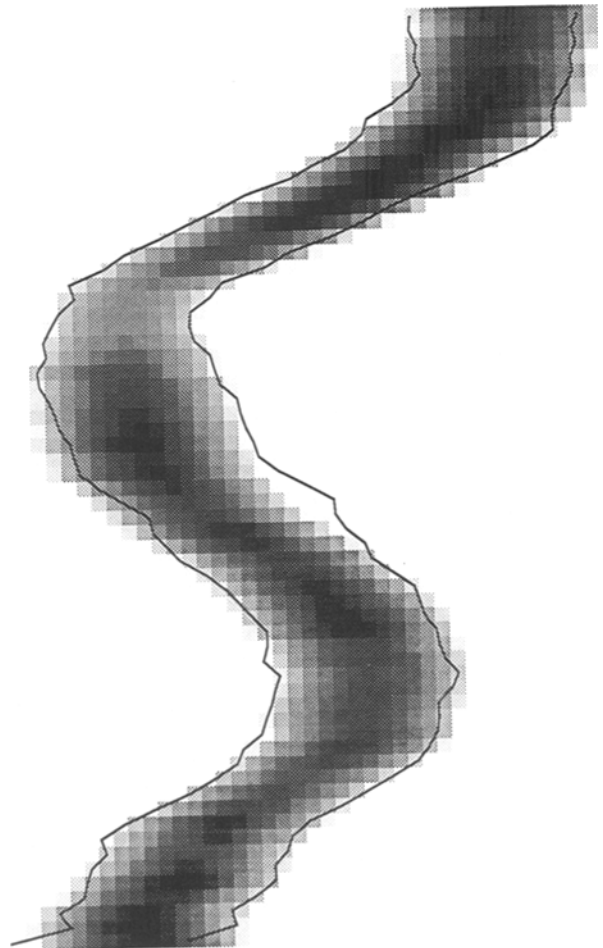


Figure 9. Noiseless object projection (shaded) with superimposed projection of FSRs-estimated object (solid line), $\theta = 90^\circ$.

where K is the number of primitives in that slice, and we suppress the z -dependence for convenience. The difficulty of the associated nonlinear search in $K \cdot \dim(\mathbf{x}_j)$ dimensional parameter space is alleviated (as will be described in detail in a forthcoming paper) by an algorithm based on Alternating Maximization (Ziskind and Wax 1987), where the likelihood is iteratively maximized over the parameters of one primitive, holding the others fixed, effectively requiring only K searches in $\dim(\mathbf{x}_j)$ dimensional space. Its structure consists of iterative matched filtering of a signal formed by subtracting out the contributions of all but one estimated primitive from the measurements, by a filter with a kernel matched to the *projection* of that primitive.

The detection problem of determining the number K of primitives in the slice can be addressed in several ways, including multiple hypothesis testing, and Information Theoretic criteria (Rissanen 1978). Alternatively, the estimation procedure can be carried out for a predetermined upper bound on the number of primitives in a slice, thus generating

false alarms that will have to be pruned by the following steps of the hierarchy.

In this step, the raw projection data are massively reduced, producing for each $z = 1, \dots, L$ a set $\tilde{\zeta}_j(z) = \{\zeta_{ij}\}_{i=1}^{K(z)}$ of $K(z)$ (arbitrarily ordered and sequentially numbered) ML parameter estimates of detected primitives. Equivalently, each $\zeta_j(z)$ provides a *linear* measurement of $\mathbf{x}_j(t)$ corrupted by the estimation error $\mathbf{v}_j(z)$ for some *unknown* object j , as in

$$\zeta_j(z) = \mathbf{x}_j(z) + \mathbf{v}_j(z). \quad (13)$$

Assuming the estimation error to be Gaussian distributed and uncorrelated between data points, the system and measurement models described by equations (2) and (13) seem to be in a classical form to which Kalman filtering (Jazwinski 1970) may be applied, to obtain a MMSE estimate of $\{\mathbf{x}_j(z)\}$ thus reconstructing the object. Note however, that contrary to the classical case, the uncertainty in equations (2) and (13) is not confined to the noises \mathbf{w} and

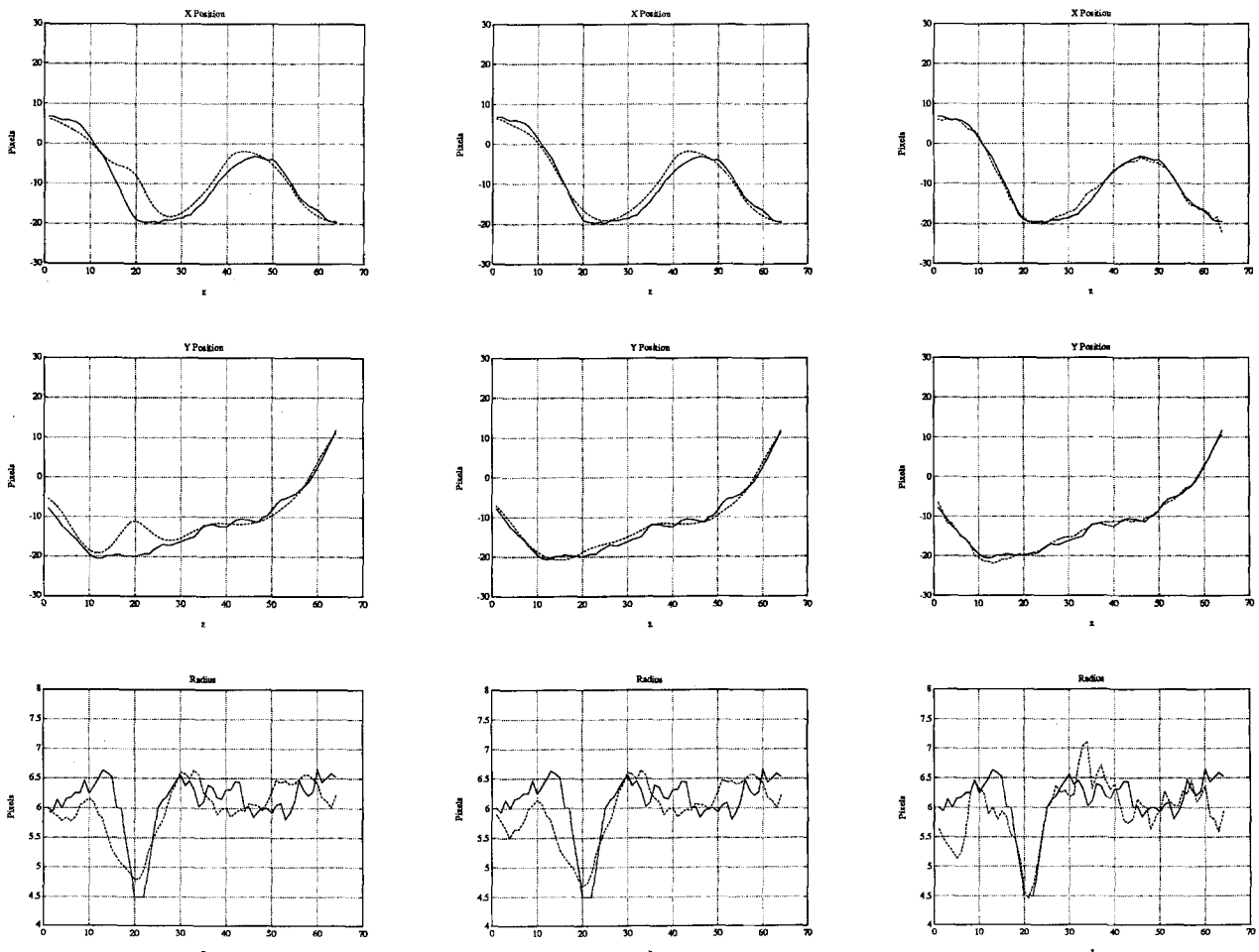


Figure 10. Comparison of true (solid) and estimated (dashed) parameters for the simple (left), hierarchical (center), and FSRS (right) algorithms.

v_j ; in fact, the association between a measurement $\zeta_j(z)$ and an object j is not known. Moreover, owing to detection error, “false alarm primitives” will be generated that do not correspond to any object.

Unsupervised Pattern Recognition Problem

The problem is first converted into a finite combinatorial search, to find the globally optimal way (in the MAP sense) to combine the detected primitives into objects. Equivalently, this is an unsupervised pattern recognition problem, in which it is necessary to estimate from the unclassified data ζ both the number of clusters (objects) present, and their individual parameters. Our solution (Bresler 1985, Bresler and Macovski 1984b) to this problem is inspired by a technique proposed by Morefield (1977) for multitarget tracking in a cluttered environment.

An *Association Hypothesis* H is a complete partitioning of the set Z of detections in all the slices into J^* disjoint *object clusters* $\{\lambda^j, j = 1, \dots, J^*\}$ and a *false alarm cluster* Φ consisting of data points rejected from all λ^j . Since no two objects occupy the same point in space, or share a primitive, the distinct clusters must be disjoint, and we have

$$\Phi = Z - \bigcup_{j=1}^{J^*} \lambda^j \quad (14a)$$

$$\lambda^i \cap \lambda^j = 0 \quad i \neq j \quad (14b)$$

Denoting the set of all such valid hypotheses by S , we formulate the problem as a multihypothesis test under the MAP criterion $H = \arg \max_{H \in S} P(H|Z)$, which is known (Jazwinski 1970) to minimize the associated Bayes risk, if errors are equally weighted and correct decisions not penalized.

Since this combinatorial problem is still too large to be solved directly² additional steps are necessary. Using the assumed independence of the objects, of the estimation noises v_j and the constraint (14b), which together imply independence of the clusters λ^j conditioned on a specific association hypothesis H , we decompose the problem into a step of feasible object construction followed by an integer optimization program, greatly alleviating the computational difficulty.

Step 2: Feasible Object Construction

A clustering procedure is implemented in this step to detect clusters λ^j in Z that are “reasonable” to

incorporate in an assignment hypothesis H , given the object and measurement models (2) and (7). Specifically, we construct a *feasible object set* containing object clusters whose likelihood exceeds a certain threshold by scanning the data with a depth first, backtracking search procedure; as many data points are added to an object as possible, before its likelihood test is failed, at which time the algorithm backtracks, and examines another branch of the search tree. The likelihood function for this sequential hypothesis test is computed from the white Gaussian innovation sequence (Jazwinski 1970) produced by a *linear* Kalman filter based on the model (2), (7), which is applied to the data λ^j as the search proceeds.

Only association hypotheses formed using the feasible objects will be considered in the subsequent Bayesian decision process, effectively substituting a pruned feasible set for the larger original set of all possible association hypotheses. The elimination of low probability (infeasible) hypotheses in this process contributes to great computational savings.

Step 3: Discrete Optimization

Note that while two or more (individually) feasible objects may share a primitive, only one of these objects can be included in the winning association hypothesis. Therefore, a multiple hypothesis test is performed next to determine the number of objects and which particular combination of the many feasible objects maximizes the posterior probability subject to the constraints (14). This final multihypothesis test can be structured as a linear integer optimization problem of the form

$$\min_{\tau} c^T \tau \quad \text{subject to } A\tau \leq [1, 1, \dots, 1] \quad (15)$$

where τ is a binary “selection” vector whose entries are zero or one according to whether or not the corresponding feasible object is included in the final selection. Each of the columns of the matrix A is used to similarly encode the selection of detected primitives in a corresponding feasible object obtained in the previous step, and the vector C contains the individual negative log-likelihoods associated with these feasible objects. Thus, the inequality in (15) represents the constraints (14), and the minimization of the linear cost function corresponds to a global MAP criterion. The problem (15) is known as the set packing problem in 0-1 integer programming (Taha 1975) and specialized efficient algorithms taking advantage of the typical sparseness of the matrix A are available for its solution

² With k detected primitives in each of L slices, the number of different hypotheses H in S is lower bounded by $\|S\| \geq k(k!)^{L-1}$, ruling out a solution by direct enumeration of these hypotheses in all but the smallest problems.

Table 1. Comparison of three algorithms

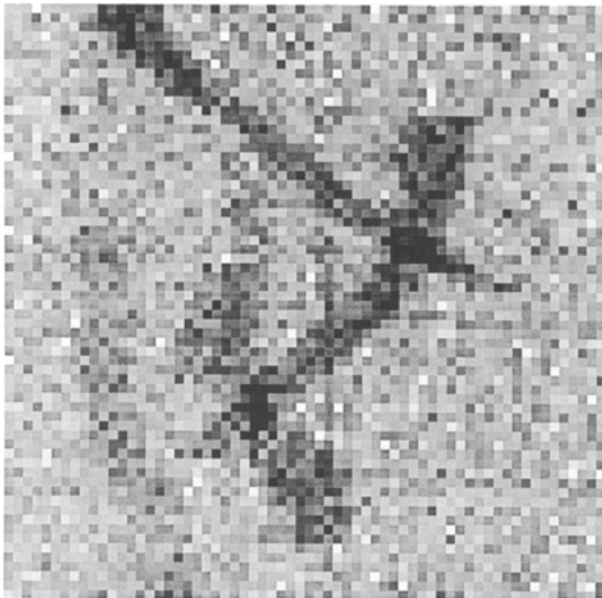
	Error Standard Deviations		
	Simple	Hierarchy	FSRS
X Position (actual)	3.06	1.49	0.94
X Position (predicted)	2.12	2.12	0.7
Y Position (actual)	2.53	0.99	0.76
Y Position (predicted)	2.12	2.12	0.7
Radius (actual)	0.39	0.39	0.40
Radius (predicted)	0.89	0.89	0.48

Step 4: Fixed Interval Linear Smoothing

At the completion of the previous step, an optimum association hypothesis \hat{H} is produced specifying the estimated number of objects \hat{J} , their starting positions $\hat{z}_{1,j}$, lengths \hat{L}_j , $j = 1, \dots, \hat{J}$, and which detected primitives belong to which object. Given this hypothesis and the models (2) and (13), the problem of determining optimum estimates of the state sequences X_j , $j = 1, \dots, \hat{J}$ of these objects decomposes (once again, owing to the conditional independence of the objects) into \hat{J} independent linear Gaussian fixed interval smoothing problems (Jazwinski 1970). A MMSE estimate of the X_j is obtained by applying standard Kalman filter smoothing formulae individually to the data sets $\lambda^j \in \hat{H}$.

Computational Requirements

The operation count is found (Bresler 1985) to be dominated by the first step, and in particular by the convolutions required to implement the matched filter operations. For example, depending on the specific implementation of the AM algorithm, the op-

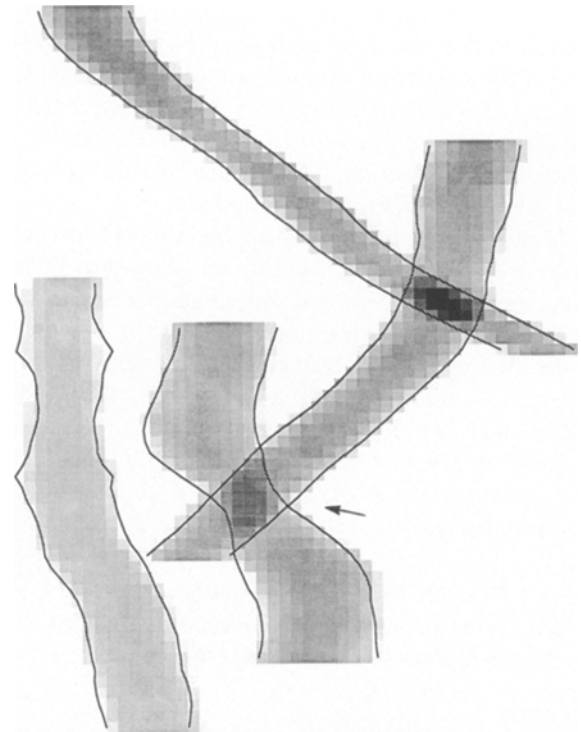
Figure 11. Noisy projection measurement at $\theta = 0^\circ$.

eration count for the simulations in section VI is between 10^6 and 10^8 .

VI. Simulation Results

Both of the foregoing algorithms have been implemented for the case of cylindrical objects with circular cross section. In addition, the following "simple" suboptimal algorithm has been implemented for comparison: for each slice, find the maximum likelihood estimate of the parameters of a cylindrical primitive. Smooth the resulting estimates using the object model.

For the simulation, we choose the object model to reflect the following assumptions: (a) low curvature of center axis; (b) independent evolution of the ξ and η coordinates; (c) continuity of the cross-section parameters; (d) uniform distribution of the initial horizontal position in D with $T = 64$ pixels, and of the radius in the range $[2,8]$ pixels. A random cylindrical object was generated by driving the object models with white noise. The resulting object has an average radius of 6 pixels. A 50% area reduction narrowing was inserted by hand. Simulated projection measurements of the object, (Figures 6 and 7) were generated at angles $0, 45, 90,$ and 135° with additive zero-mean white Gaussian noise re-

Figure 12. Noiseless object projection (shaded) with superimposed projection of FSRS estimated object (solid line), $\theta = 0^\circ$.

sulting in a projection SNR (the ratio of signal peak to noise-standard-deviation) of 0.5–0.7.

Figures 8 and 9 show synthesized projections of the estimated object (as estimated by the FSRS algorithm) superimposed on the projection of the actual (noiseless) object. With a suitable graphics package, a 3-D display of the estimated object could be generated. Figure 10 allows a quantitative comparison of estimated versus true object parameters for the three different approaches to the estimation problem. As can be seen from Figure 10a, the “simple” algorithm has substantial tracking error. This is due to the threshold effect for maximum likelihood position estimates (Zakai and Ziv 1969), which results in several of the primitives being poorly estimated. However, Figure 10b represents the result of applying the hierarchical algorithm with two primitive estimates per slice. The optimum estimated object formed from the estimated primitives is seen to be substantially more accurate. Finally, Figure 10c shows the FSRS MMSE algorithm estimates using 16 frequency components.

Table 1 compares the average performance of these algorithms. Note that the error performance predicted by the simple algorithm is optimistic, since it ignores the threshold effect. The hierarchical algorithm’s performance is comparable to that of the FSRS MMSE algorithm, indicating that the approximations used in its derivations hold at this (and higher) signal-to-noise ratios. The hierarchical algorithm has been demonstrated with simulations of multiple objects (to be reported in detail in a forthcoming paper); Figure 11 is one of four noisy projection measurements used there, and Figure 12 shows the superposition of the estimated objects on the projections of the actual objects. All of the estimation errors are in the subpixel range, including the tracking of a narrowing (indicated by the arrow) in the largest object which is difficult to see in the projections. Therefore, the hierarchical algorithm is the candidate for further investigation.

VII. Concluding Remarks

We have presented an estimation framework for 3-D reconstruction from limited-view noisy line-integral projections based on dynamic stochastic object models. Although the globally optimal solution of the resulting joint detection estimation problem is computationally infeasible, the two algorithms presented in this paper provide approximate and suboptimal solutions, respectively, drawing on a variety of techniques. The FSRS represents a synthesis of a signal processing technique, the DFT, with approximate recursive optimum nonlinear

smoothing, to produce a new algorithm for multidimensional signal processing. The hierarchical algorithm, on the other hand, employs a divide and conquer strategy at several levels, extensive data reduction at the first step, and a mix of continuous and combinatorial optimization, which all contribute to the computational efficiency of the algorithm. In the process, global optimality of a single step procedure is replaced by local optimality of the individual steps, whose results are combined to optimize a global performance criterion. However, judicious choice of decision thresholds in the various steps of the algorithm that produces a sequence of finer and finer tests, allows one to obtain significant reduction in computation, with little loss in performance.

The simulation results reported in this paper indicate the potential of the algorithms for reconstruction from extremely noisy and restricted data. The comparison of the two algorithms confirms that the hierarchical algorithm is capable of performance similar to the approximately optimal FSRS. In contrast, a simple algorithm consisting of single-slice maximum likelihood estimation followed by smoothing based on the object model performed much worse, suggesting that a sophisticated approach is indeed necessary for this problem.

Our hierarchical algorithm differs from other hierarchical algorithms that abound in artificial intelligence (AI) in general (Nilsson 1980) and computer vision in particular (Ballard and Brown 1982), in that rather than use ad hoc criteria and procedures, it is characterized by well-defined quantitative optimality criteria, and by procedures that are either optimal or represent a well-defined trade-off between computation and optimality. This pursuit of optimality is motivated by the adverse conditions under which the algorithm has to operate—limited data and poor SNR; while in most computer vision applications the aim is to match the performance of a human observer, the purpose of our the algorithm is to perform a task of which neither the unaided human observer nor current algorithms are capable, 3-D reconstruction from few projections at low SNR.

The formulation and algorithms as presented in this paper are restricted by the object representation and by the causal evolution model to objects whose center axis is a single valued function of a fixed spatial coordinate. While this limitation is inherent to the recursive structure of the FSRS, it can be overcome in a modified version of the hierarchical algorithm, as described in Bresler (1985). Another inherent limitation is to the class of generalized cylinders. However, as indicated before, by

appropriately choosing the cross-section function f_0 and dynamic model parameters, a good approximation to a variety of man-made and natural objects can be obtained.

References

- Ballard, DH, Brown CM (1982) Computer vision. Prentice Hall, Englewood Cliffs, NJ
- Binford TO (1971) Visual perception by computer. IEEE Conf. on Systems and Controls, Miami, FL
- Bloch P, Udupa JK (1983) Application of computerized tomography to radiation therapy and surgical planning. Proceedings of the IEEE 71: 351-355
- Bresler Y (1985) Model based estimation techniques for 3-D reconstruction from projections. Ph.D. Thesis, Department of EE, Stanford University
- Bresler Y, Kailath T (1987) Tracking of shift and shape parameters. Accepted for publication, IEEE Trans. Aut. Cont.
- Bresler Y, Macovski A (1984a) Estimation of 3-D shape of blood vessels from x-ray images. Proceedings of the IEEE Computer Society International Symposium on Medical Images and Icons, Arlington, VA: 251-258
- Bresler Y, Macovski A (1984b) A hierarchical Bayesian approach to reconstruction from projections of a multiple object 3-D scene. Proceedings of the 7th International Conference on Pattern Recognition, Montreal, Canada. 1: 455-457
- Bresler Y, Macovski A (1987) 3-D reconstruction from projections with incomplete and noisy data by object Estimation. IEEE Transactions Acoustic Speech Signal Processing ASSP-35: 1139-1152
- Brillinger DR (1980) Time series *Data analysis and theory*. Holden-Day, San Francisco
- Deans SR (1984) The radon transform and some of its applications. Wiley, New York
- Edelstein WA et al (1983) Signal, noise, and contrast in NMR imaging. Journal of Computer Assisted Tomography 7: 391-401
- Gordon R, Herman GT (1974) Three-dimensional reconstruction from projections: A review of algorithms. Int. Rev. Cytol. 38: 111-151
- Hanson KM, Wecksung GW (1983) Bayesian approach to limited-angle reconstruction in computed tomography. Journal of Optical Society of America 73: 1501-1509
- Herman GT (1980) Image reconstruction from projections. Academic Press, New York
- Jazwinski AH (1970) Stochastic processes and filtering theory. Academic Press, New York
- Kak AC (1979) Computerized tomography with X-ray, emission, and ultrasound sources. Proceedings of the IEEE 67: 1245-1272
- Morefield CL (1977) Applications of 0-1 Integer Programming to multitarget tracking problems. IEEE Trans Aut. Cont. AC-22: 302-312
- Munk W, Wunsch C (1979) Ocean acoustic tomography: A scheme for large scale monitoring. Deep Sea Research 26A: 123-161
- Louis AK, Natterer F (1983) Mathematical problems of computerized tomography. Proceedings of the IEEE 71: 379-390
- Nilsson NJ (1980) Principles of artificial intelligence. Tioga, Palo Alto, CA
- Nishimura DG, Macovski A, Pauly JM (1986) Magnetic resonance angiography. IEEE Trans. Med. Imag. 5(3): 140-151
- Nishimura DG, Macovski A, Pauly JM, Conoly SM (1987) MR Angiography by Selective Inversion Recovery. Magnetic Resonance in Medicine, 4: 193-202
- Rangayan RM, Gordon R, Dhawan AP (1984) Algorithms for limited-view computer tomography: An annotated bibliography and a challenge. Digest of Topical Meeting on Industrial Applications of Computer Tomography and NMR Imaging Optical Society of America, Hecla Island, Manitoba, Canada (see also Proc. IEEE Int. Conf. on Computers, Sys., and Signal. Proc. Bangalore, India)
- Rissanen J (1978) Modeling by shortest data description. Automatica 14: 465-471
- Rossi DJ, Willsky AS (1984) Reconstruction from projections based on detection and estimation of objects. IEEE Transactions on Acoustics, Speech & Signal Processing ASSP-32: 886-906
- Selfridge PG, Prewitt JMS (1981) Organ detection in abdominal computerized tomography scans: Application to the kidney. Computer Graphics and Image Processing 15: 265-278
- Shmueli K, Brody WR, Macovski A (1981) Estimation of blood vessels boundaries in X-ray images. SPIE Conference on Digital Radiography, Stanford, CA
- Stark H (Ed.) (1987) Image recovery: Theory and application. Academic Press, Orlando, FL
- Taha HA (1975) Integer programming: Theory, applications, and computations. Academic Press, New York
- Zakai M, Ziv J (1969) On the threshold effect in radar range estimation. IEEE Transactions on Information Theory, IT-15: 167-170
- Ziskind I, Wax M (1987) Maximum likelihood localization of multiple sources by alternating maximization. Proc. IEEE International Conference on Acoustics, Speech, Signal Processing, Dallas, TX: 53.6.1-4