A Bayesian Approach to Reconstruction from Incomplete Projections of a Multiple Object 3D Domain

YORAM BRESLER, MEMBER, IEEE, JEFFREY A. FESSLER, AND ALBERT MACOVSKI, FELLOW, IEEE

Abstract—An estimation approach to three-dimensional reconstruction from line integral projections, with incomplete and very noisy data, is described. Generalized cylinders parametrized by stochastic dynamic models are used to represent prior knowledge about the properties of objects of interest in the probed domain. The object models, a statistical measurement model, and the maximum *a posteriori* probability performance criterion are combined to reformulate the reconstruction problem as a computationally challenging nonlinear estimation problem. For computational feasibility, we describe a suboptimal hierarchical algorithm whose individual steps are locally optimal, and are combined to satisfy a global optimality criterion. The formulation and algorithm of this paper are restricted to objects whose center axis is a single valued function of a fixed spatial coordinate. Simulation examples demonstrate accurate reconstructions with as few as four views in a 135° sector, at an average signal-to-noise ratio of 3.3.

Index Terms—Hierarchical algorithm, nonlinear estimation and detection, parametric inference, stochastic object modeling, three-dimension tomography with incomplete data.

I. INTRODUCTION

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,[1] 3D reconstruction is simply obtained by stacking thin reconstructed slices.

[1]The 3D (or higher dimensional) version of the problem can be similarly formulated and direct 3D inversion algorithms have been derived (e.g., [1]), but appear to be rarely used in practice, owing to their increased complexity, and perhaps also the fact that commercial computerized tomography (CT) machines are geared for slice by slice data acquisition.

The applications of reconstruction from projections arise in diverse disciplines, ranging from medicine to geophysical exploration, and from astronomy to electron microscopy (see [2] and, in particular, [3] for extensive lists of applications and hundreds of references). Of all the applications, the success of CT is largely responsible for much of the current interest in a variety of reconstruction methods. However, as the applications expand into new areas, the situation often arises where the number of views, their angular coverage, and the number of rays within a view is severely restricted, and the available data are corrupted by noise, due to, e.g., time, physical, geometrical, or economic constraints in the data acquisition [4], [5]. This is almost always the case with 3D 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, due to the ill posedness of the problem, and in extreme cases, due to the inherent nonuniqueness of the solution [6]. Consequently, although 3D reconstruction would be an ideal tool in a variety of medical [7] and other applications, it is rarely attempted in practice with limited data.

The 2D limited-data reconstruction problem has been extensively researched [4], [8], [9] resulting in various methods that have varying degree of success in the noiseless case, but essentially fail when noise of significant level contaminates the data, as is typical in many practical applications. The development of effective algorithms for the limited-data high-noise case is therefore still an open problem.

While the major emphasis of research in reconstruction from projections has been on producing accurate, high resolution cross-sectional images, it has been observed [5] that in many applications the ultimate goal is often far more modest. The reconstructed image is postprocessed, either visually, or by automated techniques, to extract specific information on *objects* in the probed domain. Examples include the detection and localization of organs and tumors in the diagnostic interpretation of medical CT scans [10], the detection and tracking of high contrast thermal regions in oceans by oceanographic acoustic tomography [11], and the detection and localization of interior cracks and flaws in materials, in the area of nonde-

structive testing. The required information is often quantitative as in the case of imaging the heart ventricles [12], and the results are judged by the accuracy at which the parameters of interest are estimated, rather than the accuracy of the reconstruction of the underlying distribution. Moreover, the need for postprocessing to outline object boundaries is particularly noteworthy in 3D reconstruction [13], where a display of an unsegmented 3D density distribution is not very instructive by itself.

This paper addresses the 3D reconstruction problem for which the ultimate goal is to extract object-related information about the probed volume, and when the number, overall view angle, signal-to-noise-ratio (SNR), and sampling density of the projection measurements are limited to the point where current limited-view reconstruction techniques essentially fail. We describe an approach [14]–[16] which incorporates a priori information in the form of stochastic models of three-dimensional "objects" in the probed domain and of the measurement process, in order to overcome the inherent underdeterminancy of the solution from the data. Stochastic modeling is used in order to account for the associated uncertainty and in order to encompass a large class, i.e., an ensemble, of objects rather than a single "nominal" version. In particular, each object is represented by a parametrized cross-sectional density function, with a stochastic dynamic model for the evolution of these parameters along the object. Such a model may adequately represent in medical applications a variety of "smoothly shaped" anatomical organs, (the original motivation for this work has been the 3D reconstruction of systems of blood vessels from X-ray images [17]) as well as quite general structures in other application areas, e.g., detection of bubbles in nondestructive testing of castings.

The formulation and algorithm 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. An extension of this approach to remove this restriction is possible [15] and will be described in a forthcoming paper.

Combining the stochastic object model with a model for the limited noisy projection measurements, the reconstruction problem is reformulated as a challenging nonlinear state estimation problem where the parametrized object representations are directed estimated from the projection data. With a view toward operation with low SNR data, we seek an optimum solution, subject to the *maximum a posteriori probability* (MAP) performance criterion. That is, we seek object estimates that are globally optimal, in the sense that their posterior probability conditioned on all the available projection data is maximized. Hence, in distinction from conventional slice by slice tomographic reconstruction where the projection data corresponding to one slice is used only by reconstructing that slice, the estimate of any one point of any one of the objects, will make optimum use of all available measurements and prior information. Since the solution of the exact estimation problem is computationally infeasible, we describe a suboptimal hierarchical algorithm, whose individual steps are locally optimal, and are combined to satisfy a global optimality criterion. In addition to the quantitative morphological information provided by the algorithm, the results can be used to generate a synthetic 3D display of the reconstructed object, for visual evaluation.

Previous work using a similar philosophy has been limited, for the most part, either conceptually or by computational feasibility, to a 2D case. Hanson and Wecksung [18] have considered the MAP estimation of an object drawn from a restricted ensemble of fuzzy annuli. Rossi and Willsky [5] proposed a maximum-likelihood (ML) detection and estimation approach to limited view 2D cross-sectional imaging of an object with a cross-section that is functionally known up to a parameter vector of small dimensionality. Shmueli et al. [19] use an object model similar to ours, although in the narrow context of estimating the boundaries of a blood vessel from single planar projection of the volume. However, their MAP estimation algorithm is restricted by computational feasibility to the circular object cross-section they assume, and to a single projection angle, and hence, in a sense, to a 2D problem. Finally, the approach of Bresler [15, ch. 4] and Bresler and Macovski [16] uses the same object model as here, but differs in the choice of optimality criterion, in the computational structure, and in domain of applicability: 1) it uses the minimum mean square error (MMSE) rather than the MAP criterion, 2) the algorithm is in a single step, approximately optimal and recursive, in contrast to the suboptimal hierarchical scheme here, and 3) the computational requirements and certain observability limitations restrict the algorithm to a domain containing few objects. In contrast, the present algorithm is designed expressly for the reconstruction of a multiobject, densely populated domain.

In the next section, we introduce the statistical models for the object and the projection measurement process. In Sections III and IV, we pose the associated estimation problem and describe the algorithm for its solution. In Section V, we discuss its computational requirements, and then in Section VI, we present simulation results illustrating the method. Section VII contains concluding remarks.

II. STATISTICAL MODELS

A. Object Model

The probed region in 3D space, is the cylinder

$$\mathfrak{D} = \left\{ (\xi, \eta, z): \xi^2 + \eta^2 \leq (T/2)^2, 0 \leq z \leq L \right\}$$

(2.1)

which is assumed to have a known background density $f_b(\xi, \eta, z)$. We assume that $\mathfrak{D} \subset \mathfrak{R}^3$ contains J distinct 3D objects $O_j, j = 1, 2, \cdots, J$, each represented by a real function $f_j(\xi, \eta, z), j = 1, 2, \cdots, J$ which is defined on $\mathfrak{D}$, and its value describes an additive component to the density in each point in space, so that the overall den-

sity, which we ultimately desire to reconstruct, is given by

$$f(r, z) = f_b(r, z) + \sum_{j=1}^{J} f_j(r, z),$$ (2.2)

where $r = (\xi, \eta)'$ represents the 2D 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.) $f_j(\xi, \eta, z)$, has finite support $\mathbb{S}_j = \{(\xi, \eta, z): f_j(\xi, \eta, z) \neq 0\}$, $\mathbb{S}_j \subset \mathfrak{D}$, corresponding to the part of $\mathfrak{R}^3$ inside the "object boundaries."

A subclass of *generalized cylinders*, (or *generalized cones*) [20] is adopted in this paper as a representation for the objects, providing an effective tradeoff between model complexity and richness. An object (Fig. 1) is represented by a collection of vertically stacked cylindrical sections, which we call *primitives*. All primitives (Fig. 2) have equal height, which is scaled to 1 without loss of generality, and are each characterized by their cross-sectional density in the plane perpendicular to the z axis

$$d \cdot f_0(r; \gamma)$$ (2.3)

where d is the object contrast (determined, e.g., by its X-ray linear attenuation coefficient) and f_0 is a *known* function of r and of a vector of parameters γ characterizing the shape of the object cross-section at the position of the specific primitive. 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 $\gamma = [a, \mu, \phi]'$: a is the *radius* of the ellipse, defined as the geometrical mean of its major and minor semi-axes, μ is its *axis ratio*, and ϕ is its *orientation*, measured by the angle between its major axis and the ξ coordinate axis. This example will be used throughout the paper, although the algorithm derived here applies to an arbitrary known f_0 (e.g., the one shown in Fig. 2). For notational compactness, we define the parameter vector $\Gamma \triangleq [\gamma', d]'$, and describe a primitive by its density function $\mathrm{CYL}(r, z; \Gamma)$,

$$\mathrm{CYL}(r, z; \Gamma) \triangleq d \cdot f_0(r; \gamma) \Pi(z)$$ (2.4)

where

$$\Pi(z) \triangleq \begin{cases} 1 & \text{if } |z| < 1 \\ 0 & \text{otherwise.} \end{cases}$$ (2.5)

It follows that CYL, which is uniquely defined by the cross-section function f_0, is also a known function. Each object primitive is centered at a point $c(z) \in \mathfrak{R}^2$ on a discrete trajectory $c(z)$, $z = z_1, \cdots, z_1 + L$. These points define the (piecewise linear) *center axis* of an object of vertical *starting position* z_1 and *length* L. (Fig. 1). The variation of the parameter vector Γ along the object is described by the function $\Gamma(z)$, which, being constant within any one primitive, is understood to depend on the integer $round(z)$ that is nearest to z. Thus, the density function of a primitive located at position z along the center axis of the object O_j is CYL $[r - c_j(z); \Gamma_j(z)]$ where

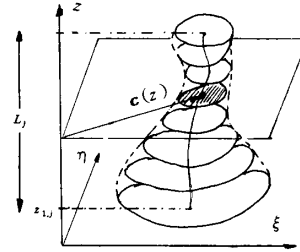
Fig. 1. Object formed from stacked primitives.

Fig. 2. Unit height cylindrical primitive.

the interpretation of $c(z)$ for noninteger values of z is analogous to that of $\Gamma(z)$. Finally, all the parameters of this primitive are lumped into the vector $\chi_j(z) = [c_j'(z), \Gamma_j'(z)]' = [c_{\xi_j}(z), c_{\eta_j}(z), a_j(z), \mu_j(z), \phi_j(z), d_j(z)]'$.

Note that since $c(z)$ and $\Gamma(z)$ are (single valued) functions of z, the foregoing representation is restricted to objects that do not "wind back" in the z direction, i.e., an object whose center axis intersects a horizontal plane at more than one point must be excluded. This requirement can often be circumvented in practice by a different choice of coordinate system, and/or by segmentation of the probed domain into a few components.

Common natural and biological structures are usually characterized by smooth surfaces of limited curvature, a fact which we attempt to employ. Since, in general, there will be significant uncertainty in our prior knowledge of the object's spatial properties, we choose the stochastic setting, in which our statements about the smoothness properties of the object are probabilistic. An object model in this context is a rule that assigns a probability measure to the different realizations in the given class defined by f_0. Objects with occasional regions of high curvature may thus be admitted, but a low probability assigned to them. Consequently, the class of objects modeled by this approach may be quite large, without compromising the use of prior information.

Prior knowledge about the object spatial properties is represented by a stochastic dynamic model for the primitive-to-primitive evolution of the cross-section parameters. Defining the augmented state vector $x_j(z) = [x_j^{(1)'}(z), \chi_j'(z)]'$ with $x_j^{(1)}$ augmenting the state to accommodate higher order models for $\chi_j(z)$, we assume the following discrete linear state space dynamic stochastic Markov model for the spatial evolution of $x_j(z)$

$$x_j(z + 1) = Ax_j(z) + Bw_j(z),$$

$$z = z_{1,j}, \cdots, z_{1,j} + L_j$$

$$E[x_j(z_{1,j})] = x_0 \quad \mathrm{COV}\{x(z_{1,j})\} = \Pi_0$$ (2.6)

where $w_j(z)$ is a white Gaussian noise vector-sequence of unit (identity matrix) covariance, uncorrelated with $x_j(z_{1,j})$. In addition, the object starting position $z_{1,j}$ is an unknown parameter, and a model for the object length L_j may also be available, for example, the exponential model

$$P(L_j) = \frac{1}{\overline{L}} \exp\left(-\frac{L_j}{\overline{L}} \right) \tag{2.7}$$

where $\overline{L}$ is the expected object length.

Quite general structures may be modeled by properly choosing $f_0(\cdot; \cdot)$, $\gamma(\cdot)$, and dynamic model parameters. The process noise $w_j(z)$ and the random initial state $x_j(z_{0,j})$ both represent the uncertainty as to the exact shape and position of the jth object: while a "nominal," deterministic object is given by the expected value $E[x_j(z)]$ which is obtained by initializing (2.6) with x_0 and setting $w_j(z) = 0$ for all z, the actual object is some random "variation" on this theme. The greater the noise covariance BB' and the initial state covariance Π_0, the higher the probability of objects with large deviations from the "nominal" object, and the richer the class of objects modeled.

The selection of model parameters A, B, Π_0, x_0, and $\overline{L}$ is discussed in [16]. In particular, they may be derived from a training set by a system identification procedure, or postulated to represent a reasonable set of assumptions on object boundary curvature. In Section VI, we describe one such choice, which we used in our simulation example.

The different objects will be assumed unrelated, so that given that they do not intersect in 3D space, the stochastic processes $\{x_j(z), z = z_{1,j}, \cdots, z_{1,j} + L_j\}$, will be assumed statistically mutually independent for different $j = 1, \cdots, J$, implying that

$$\mathrm{COV}\left\{ x_i(z_{1,i}), x_j(z_{1,j}) \mid \mathrm{object}_i \cap \mathrm{object}_j = \phi \right\}$$

$$= 0 \quad \forall i \neq j$$

$$E\left[w_i(l)\, w_j'(m) \mid \mathrm{object}_i \cap \mathrm{object}_j = \phi \right]$$

$$= 0 \quad \forall l, m. \tag{2.8}$$

This will turn out to be a key assumption in deriving an efficient computational algorithm. The objects will be further assumed to have identical statistics, and will therefore be modeled by identical state space models.

Given the representation and model for the objects, the jth object can be identified with the set $O_j = \{z_{1,j}, L_j, X_j\}$ where $X_j \triangleq \{x_j(z), z = z_{1,j}, \cdots, z_{1,j} + L_j\}$ is its state sequence. It is determined by this data set completely, and its 3D density function (with the dependence on O_j made explicit) can be immediately recovered from this representation as

$$f_j(r, z; O_j) = \sum_{l_j = z_{1,j}}^{z_{1,j} + L_j} \mathrm{CYL}\left[r - c_j(z), z - l_j; \Gamma_j(z) \right].$$

$$\tag{2.9}$$

Similarly, the density function $f(r, z; O_{1:J})$ of the whole domain is uniquely determined by the entire object data set $O_{1:J} \triangleq O_j, j = 1, \cdots, J$, and is obtained by substituting (2.9) into (2.2). Alternatively, a 3D graphic display of the objects, including their relative positions in space, can be generated directly from $O_{1:J}$. We ultimately seek the optimum estimate of this set, which in turn determines via (2.9) the 3D shapes of the objects.

B. Measurement Model

In the two-dimensional cross-sectional imaging problem, let $f(r)$ represent the distribution we desire to reconstruct. A line inclined at an angle θ with respect to the ξ axis and passing at the radial distance t from the origin can be parametrized as $l(t, \theta) = \{r: \theta'r = t\}$ where θ is the corresponding unit direction-vector $\theta \triangleq (\cos \theta, \sin \theta)'$. For a given projection angle θ, the projection of f evaluated at the radial distance t from the origin is the integral

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

$$\tag{2.10}$$

along the line $l(t, \theta)$. The projection $g(t, \theta)$ at any value of θ is a 1D function of t (Fig. 3). The mapping of f into g via (2.10) corresponds to the 2D Radon transformation [3] whereas the reconstruction problem of determining the distribution f from its projections g involves inverting the integral (2.10), or finding the inverse Radon transform.

Owing to the linearity of the Radon transform, the projection of the domain $\mathfrak{D}$ will be the sum of the individual projections of the background and the object superimposed on it, so that the known background contribution can be always subtracted off. Consequently, the background density $f_b(\xi, \eta, z)$ may be set to zero without loss of generality. The 3D density function of the domain is thus $f(\xi, \eta, z)$, the same as that of the objects in it. As in 3D axial tomography data acquisition, we assume that we are given projection data sets of 2D vertically stacked thin slices of the domain $\mathfrak{D}$, for $z \in [0, L]$. Considering parallel ray geometry, the 2D projection (2.10) of a slice at a given height z through a single primitive located at the origin is, according to (2.4),

$$\psi(t, z, \theta; \Gamma) \triangleq \int_{\theta'r = t} \mathrm{CYL}(r, z; \Gamma)\, dr$$

$$= \int_{\theta'r = t} df_0(r; \gamma)\, \Pi(z)\, dr$$

$$= dg_0(t, \theta; \gamma)\, \Pi(z) \tag{2.11}$$

where g_0 is a known function of t, θ, and γ; for each z, it is the projection of a unit density object slice with density $f_0(r; \gamma)$, located at the origin. For example, when f_0 is the indicator function on an ellipse with parameters $\gamma = [a, \mu, \phi]'$, $g_0(t, \theta; \gamma)$ can be shown [3] to be the half
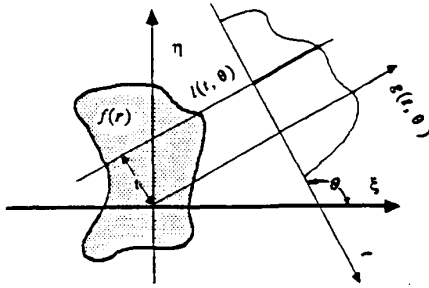
Fig. 3. 2D projection geometry.

ellipse,

$$g_0(t, \theta; \gamma) = l_e \sqrt{1 - (t/a_e)^2} \, \Pi\left(\frac{t}{2a_e}\right) \quad (2.12)$$

where $\Pi(\cdot)$ is defined in (2.5) and where a_e and l_e are determined by the parameters γ of the ellipse and by the view angle θ according to

$$a_e = a_e(\theta) = a\sqrt{A_1}$$

$$l_e = l_e(\theta) = 2a\sqrt{A_1},$$

$$A_1 = 0.5[(\mu + 1/\mu) + (\mu - 1/\mu)\cos 2(\phi - \theta)].$$

The parameters a_e and l_e are respectively the horizontal and vertical semiaxes of the half-ellipse function (2.12), and for a given ellipse they depend on the view angle.

Considered as a function of t, z and θ, the function $\psi(t, z, \theta; \Gamma)$ defines the Radon transform of an entire primitive. The projection of the same primitive located at $z = l$ and horizontal position c is given by

$$\int_{\theta'r=t} \text{CYL}(r - c, z - l; \Gamma) \, dr$$

$$= \psi(t - \theta'c, z - l, \theta; \Gamma). \quad (2.13)$$

The projection of $f(r, z; O_{1:J})$ is obtained by combining (2.4), (2.9), (2.2), and (2.11). By the linearity of the Radon transform, it follows that the Radon transform of all the J objects, with the jth object comprised of L_j primitives, is given by

$$\Psi(t, z, \theta; O_{1:J}) \triangleq \int_{\theta'r=t} f(r, z; O_{1:J}) \, dr = \sum_{j=1}^{J} \sum_{l_j=z_{1,j}}^{z_{1,j}+L_j}$$

$$\cdot \psi[t - \theta'c_j(z), z - l_j, \theta; \Gamma_j(z)]. \quad (2.14)$$

Note that at any z position, $\Psi(\cdot)$ consists of the superposition of at most J primitives. Neglecting for simplicity the effect of a nonzero imaging aperture, which may be easily taken into account [16], the actual noisy projection measurements are given by

$$y(t, z, \theta) = \Psi(t, z, \theta; O_{1:J}) + v(t, z, \theta) \quad (2.15)$$

where $v(t, z, \theta)$ is a white noise zero-mean Gaussian random field of intensity N_0. The modeling of the noise as additive and Gaussian is accurate in a variety of applications; in medical imaging it applies to magnetic resonance images where the thermal noise associated with the resistive loading of the pickup coil dominates [21], [22]. Even when the true statistics of the noise are Poisson, as in X-ray imaging, the Gaussian approximation is usually adequate. Finally, the algorithm is readily adapted to other noise statistics by modifying the likelihood function in Step 1, below.

To represent the limited data situation, the projection data is assumed to be available only at a discrete set $\theta \in \{\theta_m\}_{m=1}^{M}$ of M (possibly nonuniformly spaced) projection angles covering a sector $\theta_M - \theta_1 \le \pi$, 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. Consequently, the support set of the available projections of the probed region $\mathfrak{D}$ is

$$S_y = \left\{ (t, \theta, z) : t = -\frac{T}{2} + n\Delta, n = 0, \cdots, N - 1, \right.$$

$$\left. \theta \in \{\theta_m\}_{m=1}^{M}, z = 1, \cdots, L \right\}. \quad (2.16)$$

The interpretation of each projection $y(t, z, \theta_m)$ as a classical "X-ray" (radiograph) taken using parallel X-rays (Fig. 4), may be useful in visualizing the measurement geometry. Each line in a radiograph corresponds to a projection in the direction θ_m of a slice of unit thickness.

C. Summary of Modeling Assumptions

For ease of reference we summarize below the key modeling assumptions.

1) Known background density.
2) Objects described by generalized cylinders, with *functionally* known cross-section.
3) Center axis of objects is a single valued function of a fixed spatial coordinate denoted z.
4) Evolution of center axis and cross-section parameters described by linear Gaussian–Markov state-space model (2.6) with known parameters.
5) Model for object length possibly available as in (2.7), with known parameter $\bar{L}$.
6) Unknown total number J of objects and their z starting positions and lengths.
7) Stochastic processes describing distinct objects are independent, given that they are nonintersecting.
8) Measurement data consists of parallel-beam line-integral projections in planes perpendicular to the z axis.
9) The measurements are corrupted by additive white Gaussian noise.

Assumptions 8) and 9) are nonessential: the algorithm described in the sequel can be readily adapted to fan-beam data and to non-Gaussian noise statistics. Assumption 1) can be relaxed by assuming that the background is a smooth function of the coordinates, and can be modeled, e.g., by polynomials. The algorithm can then be extended to estimate on-line the unknown parameters of the background. Even if Assumption 2) does not hold, a hypoth-
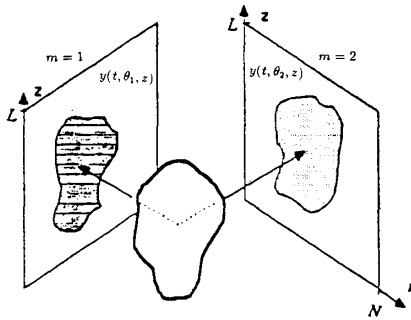
Fig. 4. Classical radiograph interpretation of $y(t, \theta_m, z)$.

esized elliptical cross-section may give an adequate fit to a variety of actual cross-sections of natural and man-made objects. Finally, simulation experiments with the algorithm, including those reported in Section IV, suggest that it is rather robust to violation of Assumption 4). Furthermore, as long as the parameters B and Π_0 are chosen to adequately represent the uncertainty about the exact evolution of the objects, the algorithm tends to degrade gracefully with deviation from modeling assumptions. The other assumptions are discussed at length in the text.

III. THE ESTIMATION PROBLEM

The estimation problem can be stated as

given the noisy set of projections $\mathcal{Y} \triangleq y(t, z, \theta)$, $(t, z, \theta) \in S_y$, the state space models (2.6)–(2.7), and possibly a model like (2.8) for L_j, determine optimal estimates $\hat{J}$ and $\hat{O}_{1:J}$ of the number of objects and their representations, respectively. The latter includes the objects' starting positions and lengths $\{\hat{z}_{1,j}, \hat{L}_j, j = 1, \cdots, \hat{J}\}$, and the set $X_{1:\hat{J}} \triangleq \{X_j, j = 1, \cdots, \hat{J}\}$ of their state sequences.

We follow a Bayesian approach, by choosing the *maximum a posteriori probability* (MAP) criterion [23] as the criterion for optimality. That is, we define the estimates $\hat{J}$ and $\hat{O}_{1:J}$ as the solution to

$$\max_{J, O_{1:J}} p(J, O_{1:J} \mid \mathcal{Y}) \qquad (3.1)$$

where $p(J, O_{1:J} \mid \mathcal{Y})$ is the posterior probability of J and $O_{1:J}$ given that the data $\mathcal{Y}$ was observed. The problem is one of joint estimation and detection, since both the number and lengths of the objects, as well as their state sequences X_j, need to be estimated from the measurements.

An appreciation of the computational requirements involved is obtained by considering just the estimation of the state sequences $X_{1:J}$. "Brute force" solution of (3.1) is impractical, in general, owing to the the extremely high dimensionality of the search space, and since the criterion will be typically multimodal, requiring, in effect, an exhaustive search. Even with the relatively efficient dynamic programming (DP) approach to the problem [19] the computational requirements grow exponentially with the total dimension of the state of the estimated process.

The latter is J dim (x_j), since the problem has to be solved jointly for all J objects, whose projections may overlap in any one of the available views. Assuming that each state component has only q possible transitions (the continuous-state space must be discretized in order to implement the DP algorithm) from one value of z to the next, the computational requirements are $O(q^{2J\dim(x)})$. For a relatively small example with $q = 10$, $J = 10$, and dim $(x_j) = 8$, we obtain $O(10^{160})$ flops, indicating the infeasibility of this approach.

IV. A FOUR-STEP ALGORITHM

In view of the infeasibility of globally optimal detection/estimation of the objects, we propose a suboptimal algorithm, producing estimates of the objects' position and shape in four hierarchical steps. First, object primitives are detected in 3D space and local maximum likelihood estimates of their parameters are obtained. This step requires a nonlinear search only in $\mathfrak{R}^{\dim(\chi)}$ [where χ is the parameter vector of a cylinder-primitive and dim $(\chi) \leq$ dim (x)]. In this step, massive data reduction is achieved, and the problem is converted into a finite combinatorial search, to find the optimal way to combine the detected primitives into objects. Since this combinatorial problem is still too large to be solved directly, additional steps are necessary. First, a set of feasible objects is constructed by combining individual detected primitives and their parameter estimates via causal minimum variance filtering. The assignment of detections to each feasible object and rejection of infeasible ones, is directed by a sequential hypothesis test. Then, a multiple hypothesis test is performed to determine which particular combination of the many feasible objects is most likely to represent the data. The number of objects is also determined in this step. The last step is again a continuous estimation procedure (in contrast to combinatorial optimization): final object estimates are obtained by linear fixed-interval smoothing of the individual objects that were selected in the previous step.

A. Step 1: Primitive Detection and Estimation

1) Introduction: We first detect cylinders in each slice (normalized to thickness 1) and estimate their local shape and position parameters using only the measurements for that slice. Hence, the problem reduces to the 2D plane. The approach of Rossi and Willsky [5] to the maximum-likelihood (ML) estimation of a single elliptical object in the plane from its projections does not apply to the multiple object case, since the projections of the different primitives in the slice may overlap, creating a coupling between the different parameter estimates. In fact, estimation of the parameters of the multiple cylinders whose projections are superimposed to produce $y(t, z, \theta)$ in (2.14), (2.15) is an extension to a higher dimension of a well known problem of estimating the parameters of overlapping pulses. For our imaging problem, the latter restricted 1D version corresponds to the measurement of a single slice at a particular projection angle. Such mea-
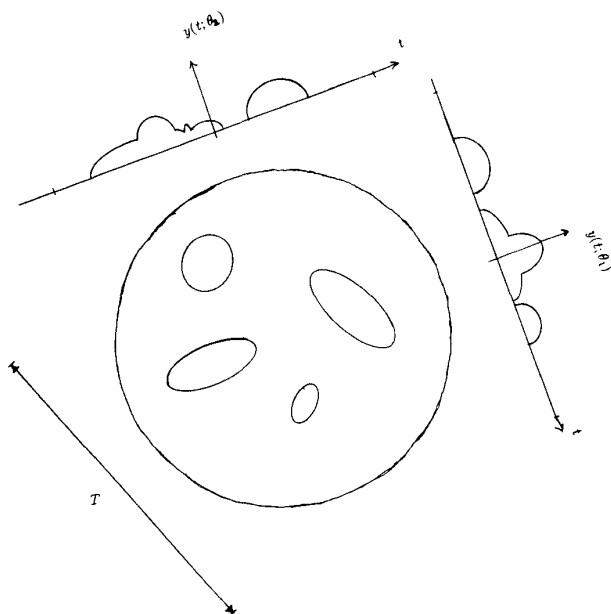
Fig. 5. Single slice projection geometry with overlapping projections.

surements are the superposition of several half-ellipse pulses of the form (2.12) of unknown position and shape parameters as shown in Fig. 5.

The latter problem of overlapping pulses, which arises, e.g., in radar and in geophysics, has been considered by several authors [24]–[29] and solved using suboptimal schemes of detection-deconvolution, inverse filtering, least squares, approximate maximum likelihood, eigen-structure analysis, and the iterative estimate-maximize approach. We have chosen to compute maximum-likelihood estimates using an approach we term alternating maximization (AM), which is derived from the "alternating variable" algorithm of numerical optimization [30]. A somewhat different version of the latter algorithm, termed alternating projection maximization (APM), has been applied to the problem of multiple source localization [31] demonstrating remarkable success. Though the global estimation problem is solved with the MAP criterion, for a given slice, there are no priors (except uniform distribution over the parameter space), so the maximum likelihood criterion is more appropriate.

2) Resolution of Superimposed Signals by the AM Algorithm: To see how the AM algorithm applies to primitive estimation, we first show in general how it can resolve J overlapping signals where J is known. The detection problem of estimating J is considered later. Consider a measurement of J superimposed signals in zero-mean white Gaussian noise:

$$y(t) = \sum_{j=1}^{J} \psi(t; \alpha_j) + \nu(t), \qquad (4.1)$$

$$E[\nu(t)\nu(s)] = N_0 \delta(t - s). \qquad (4.2)$$

The goal is to estimate the unknown set of parameters α

of the signals $\psi(t; \alpha_j)$ from Y where

$$\alpha = [\alpha_1, \cdots, \alpha_J], \, Y = \left\{ y(t): -\frac{T}{2} \le t \le \frac{T}{2} \right\}. \qquad (4.3)$$

The maximum likelihood solution is the set of parameters which has the highest probability, i.e.,

$$\alpha_{ML} = \arg \max_{\alpha} L(\alpha) = \arg \max_{\alpha} \log p_Y(Y; \alpha). \qquad (4.4)$$

This solution requires a search over a large parameter space, and is computationally infeasible when the number of signals and parameters is large. The AM algorithm reduces computation by maximizing over the parameters α_j of each signal individually while holding the other parameters fixed at their most recently estimated values according to the following iterative procedure

$n = 0;$
repeat
$\{$
    for $j = 1$ to J do
    $\{$
$$\hat{\alpha}_j^{(n+1)} = \arg \max_{\alpha_j} L_j(\alpha_j | \hat{\alpha}_{(j)}^n); \qquad (4.5)$$
    $\}$
    increment n by 1;
$\}$ until parameter change is insignificant.

where

$$\hat{\alpha}_{(j)}^n = [\hat{\alpha}_1^{(n+1)}, \cdots, \hat{\alpha}_{j-1}^{(n+1)}, \hat{\alpha}_{j+1}^{(n)}, \cdots, \hat{\alpha}_J^{(n)}] \qquad (4.6)$$

and

$$L_j(\alpha_j | \hat{\alpha}_{(j)}^n) \triangleq \log p_Y(Y; \alpha_j, \hat{\alpha}_{(j)}^n) \qquad (4.7)$$

is the conditional log-likelihood of the parameters α_j given all the other parameters. This algorithm, which searches over a subspace of dimension dim (α_j) at a time, is an intermediate form between full multidimensional search and the "alternating variable" algorithm [30] which reduces the search to a sequence of 1D searches parallel to the coordinate axes in parameter space. The advantage of searching over the parameters of one signal at a time is that the problem essentially decouples into several single-signal problems, as discussed below.

It is shown in Appendix A that the likelihood in (4.7) can be written as

$$L_j(\alpha_j | \hat{\alpha}_{(j)}^n) = \frac{1}{N_0} \langle z_j^n(t), \psi(t; \alpha_j) \rangle - \frac{1}{2N_0} \mathcal{E}(\alpha_j) \qquad (4.8)$$

where

$$\mathcal{E}(\alpha) = \| \psi(t; \alpha) \|^2 \qquad (4.9)$$

can be thought of as the energy of the signal (terms independent of α have been dropped), and the appropriate inner product and norm for continuous scalar time func-

tions are, respectively,

$$\langle f(t), g(t) \rangle = \int_{-T/2}^{T/2} f(t) \cdot g(t) \, dt \quad (4.10)$$

$$\| f(t) \|^2 = \langle f(t), f(t) \rangle. \quad (4.11)$$

For discrete measurements, the integral is replaced by a summation. The term

$$z_j^n(t) = y(t) - \sum_{i=1}^{j-1} \psi(t; \hat{\alpha}_i^{n+1}) - \sum_{i=j+1}^{J} \psi(t; \hat{\alpha}_i^n) \quad (4.12)$$

is a "cleaned" signal consisting of the original measurement with the most recently estimated contributions from the remaining $J - 1$ signals (which are held fixed) subtracted out. If these estimates are correct, the "cleaned" signal will be the noise plus the contribution due to the signal of interest, i.e., a noisy measurement of a *single* signal, and the estimate in (4.5) will correspond to the ML estimate. Not surprisingly, (4.8) coincides with the expression [32] for the log-likelihood of the single signal $\psi(t, \hat{\alpha}_j)$ in white Gaussian noise, when the measurement is the "cleaned" signal $z_j^n(t)$.

Since $\mathcal{E}(\alpha)$ can be precomputed for discretized α, the only computation in performing the maximization in (4.5) are the inner products in (4.8). Thus the AM algorithm is considerably more efficient than a full multidimensional search. We now apply the AM procedure to the projection imaging case and generalize it to handle multiple view measurements.

3) Primitive Estimation by the AM Algorithm: We still consider a single slice, so we can assume that $z = 0$ w.l.o.g., but now the measurement process consists of the M projections: $Y = \{ y(t, \theta_m): -T/2 \le t \le T/2, m = 1, \cdots, M \}$, so that the 1D index set $\{ t \}$ of the previous subsection is replaced by a 2D index set $\{ t, \theta \}$. Owing to the independence of the measurements, the results of the previous section are applied by simply replacing t with the pair (t, θ) and $\psi(t; \alpha_j)$ with $d_j \cdot g_0(t - \theta'c_j, \theta_m; \gamma_j)$ from (2.11)–(2.13). Thus the parameter set of the jth signal is $\alpha_j = \{ c_j, \gamma_j, d_j \}$. Since there are discrete projection angles, the inner product is redefined to be

$$\langle f(t, \theta), g(t, \theta) \rangle = \sum_{m=1}^{M} \int_{-T/2}^{T/2} f(t, \theta_m) \cdot g(t, \theta_m) \, dt. \quad (4.13)$$

It follows that the log-likelihood for the parameters of a single signal conditioned on all of the other signals at the $n + 1$ iteration is

$$L_j(c_j, \gamma_j, d_j) = \frac{1}{N_0} \langle z_j^n(t, \theta), d_j g_0(t - \theta'c, \theta; \gamma_j) \rangle$$

$$- \frac{1}{2N_0} \mathcal{E}(\gamma_j), \quad (4.14)$$

where now there is a "cleaned" signal for each projection

$$z_j^n(t, \theta_m) = y(t, \theta_m) - \sum_{i=1}^{j-1} \hat{d}_i^{(n+1)} \cdot g_0(t - \theta'\hat{c}_i^{(n+1)},$$

$$\theta_m; \hat{\gamma}_i^{(n+1)}) - \sum_{i=j+1}^{J} \hat{d}_i^{(n)} \cdot g_0(t - \theta'\hat{c}_i^{(n)},$$

$$\theta_m; \hat{\gamma}_i^{(n)}). \quad (4.15)$$

It is shown in Appendix A that the required maximization for (4.5) reduces to

$$[\hat{c}_j^{(n+1)}, \hat{\gamma}_j^{(n+1)}]$$

$$= \arg \max_{[c, \gamma]} \frac{\langle z_j^n(t, \theta), g_0(t - \theta'c, \theta; \gamma) \rangle^2}{\| g_0(t, \theta; \gamma_j) \|^2}, \quad (4.16)$$

and the maximizing d_j is given explicitly by

$$\hat{d}_j^{(n+1)} = \frac{\langle z_j^n(t, \theta), g_0(t - \theta'\hat{c}_j^{(n+1)}, \theta; \hat{\gamma}_j^{(n+1)}) \rangle}{\| g_0(t, \theta; \hat{\gamma}_j^{(n+1)}) \|^2}. \quad (4.17)$$

For a given shape parameter γ, the computation of the inner product in (4.16) can be interpreted [5] as a convolution-backprojection (CBP) [2] operation similar to the standard operation used in tomographic reconstruction, the only difference being that the standard convolution kernel is replaced by $g_0(-t, \theta; \gamma)$. Alternatively, the convolution step is interpreted as filtering the 'cleaned' measurement by a filter matched to the known *projection* $g_0(t, \theta; \gamma)$ of the signal. To see this, we decompose the inner product [defined in (4.13)] into the following two steps:

$$q_\theta(t) = \int_{-T/2}^{T/2} z_j^n(\tau, \theta) \cdot g_0(\tau - t, \theta; \gamma) \, d\tau$$

$$= z_j^n(t, \theta) * g_0(-t, \theta; \gamma) \quad (4.18)$$

$$\langle z_j^n(t, \theta), g_0(t - \theta'c, \theta; \gamma_j) \rangle^2 = \sum_{m=1}^{M} q_{\theta_m}(\theta'_m c). \quad (4.19)$$

The first, in (4.18), is the convolution, or matched filtering step, while the second, in (4.19), is the backprojection step, in which at each location c, the contributions from all available views are summed.

If the shape parameter space is discretized, then a bank of matched filters is used, one for each possible parameter set, and the maximization performed over all filters. For efficient implementation, the shape parameter space is first discretized coarsely, and rough estimates made by a search over this smaller set. Then the AM algorithm can be applied for refinement by iterating over the shape parameters.

Since at every iteration of the AM algorithm (4.5)–(4.7) we maximize the log-likelihood, it can only increase. Furthermore, the likelihood is bounded from above by its

global maximum corresponding to the ML estimate. Consequently, the algorithm converges. While the above properties are not a sufficient condition to guarantee convergence to a local maximum,[2] we have found that for discretized parameters, the AM algorithm converges to a local maximum within a few iterations. Convergence to the ML estimate (i.e., the global maximum) requires that the initial estimate be sufficiently close to the ML solution.

4) Initialization: Kwakernaak [29] proposes the following algorithm for estimating the parameters of superimposed signals: treat the measurements as if there were only a single signal, estimate its location and shape using a bank of matched filters. Then subtract from the measurements the estimated contribution due to that signal and repeat J times until the last signal is extracted. This procedure is identical to a single iteration of the AM algorithm if the first "estimates" are set to zero.

In Kwakernaak's application, the signals are of known shape (but unknown amplitude) and he successfully detects and estimates overlapping signals. However, as pointed out in [25] when the signal shapes are unknown: "The physical difference between the combination of two closely spaced signals and one signal located between them may be small, and the introduction of noise can obscure this difference. . . . The noise energy may be small when compared with either signal, but at the same time be large when compared with the difference we are trying to measure." We have found experimentally that the estimates can be improved significantly by a few iterations of the AM algorithm. Therefore we use Kwakernaak's algorithm for initialization and the AM algorithm for refinement. If other initialization procedures are also used, e.g., the ML estimates of the preceding slice, the likelihood of finding the true global maximum will be increased.

5) Detection of Primitives: If the number of primitives J is unknown, Kwakernaak's procedure could be repeated indefinitely and the likelihood would be monotone increasing, so Kwakernaak applies an information theoretic criterion based on the minimum description length principle [34] to estimate the number of signals. This approach has also been used [35] for estimating the number of sources sensed by an antennae array. This technique is applicable to our problem as well. For simplicity, in the sequel we use the following procedure: if J_{max} is an *a priori* upper bound on the number of primitives, then the estimation procedure is carried out assuming there are J_{max} primitives. If there are actually $J < J_{max}$ primitives in a slice, then $J_{max} - J$ "false alarms" (corresponding to hypothesized but nonexistent primitives) will be generated. The "cleaned" signal (4.15) for these primitives will just be background noise, so their ML estimates will be randomly distributed over the slice. Objects formed from

such primitives will have very low likelihood, and will easily be pruned by the following steps of the hierarchy.

6) Algorithm Output: At the end of the first step, the algorithm produces for each value of $z = 1, \cdots, L$ a set $\bar{\zeta}(z) = \{ \zeta_i(z) \}_{i=1}^{J_{max}}$ of J_{max} vectors (referred to in the sequel as "data points"), each describing the estimated parameters and location of a detected primitive (i.e., the vector χ for that primitive) at height z. The J_{max} points are assumed to be arbitrarily ordered and sequentially numbered. Thus, $\zeta_i(z)$ can be written as

$$\zeta_i(z) = \hat{\chi}_j(z) = \chi_j(z) + v_j(z) \qquad (4.20)$$

for some $1 \leq j \leq J$ where $\hat{\chi}$ denotes the suboptimal single-slice ML primitive estimate, and $v_j(z)$ is the associated estimation error. Let us define the measurement matrix H as the matrix that extracts the vector χ from the augmented state vector x, i.e.,

$$Hx = \chi \quad H = [0 | I]. \qquad (4.21)$$

Then, assuming the estimation error $v_j(z)$ to be Gaussian distributed[3] with covariance R and to be uncorrelated between data points,[4] each $\zeta_i(z)$ can be considered to provide a *linear* measurement of $x_j(z)$, corrupted by additive white Gaussian noise, for some object j, as in

$$\zeta_i(z) = Hx_j(z) + v_j(z). \qquad (4.22)$$

$$Ev_j(z) = 0 \quad E[v_j(m) v_k'(n)] = 0,$$

$$\forall k \neq j \; or \; m \neq n. \qquad (4.23)$$

The system and measurement models described by (2.6) and (4.22)–(4.23) seem to be in a classical form to which Kalman filtering [23] may be applied to obtain a minimum variance estimate of $\{x_j(z)\}$, thus reconstructing the object. Note, however, that contrary to the classical case, the uncertainty in (2.6) and (4.22)–(4.23) is not confined to the noises w_j and v_j; in fact, the association between a primitive measurement $\zeta_i(z)$ and an object O_j, is not known. Moreover, due to detection errors, "false alarm" primitives will be generated that do not correspond to any object. The situation is analogous to having a collection of beads (detected primitives) in 3-space, some of which need to be strung on threads in order to obtain necklaces (estimated objects) whose number is not prespecified, satisfying some global optimality criterion.

B. Unsupervised Pattern Recognition Problem

With the data from the first step available, our reconstruction problem can be reformulated as an unsupervised pattern recognition problem [37] in which it is necessary to estimate from the unclassified set of detected primitives

---

[2]Ziskind and Wax make this erroneous claim for their APM algorithm [31]. However, examples have been constructed [33] where the alternating variable algorithm converges to a point which is not an extremum.

[3]The ML estimator is known to be asymptotically Normal (in fact, Best Asymptotically Normal–BAN) [36].

[4]In general, the parameters of different primitives within a slice, which are estimated jointly, can be expected to be correlated, except when the projections of these primitives do not overlap in any of the given views. However, Monte-Carlo simulations show this correlation to be negligible even when the object density in the domain is so high that projections of each primitive overlap those of another one in two out of four available views.

both the number of clusters (objects) present, and parameters of individual objects. Although "trajectory" [in the broader sense of (2.6)] and measurement models are assumed known, it is not known which data point corresponds to which object.

A problem which is essentially equivalent to ours is that of multitarget tracking in a cluttered environment [38], which has been studied extensively. In the present case, however, the time axis has been replaced by a spatial axis z, and we are interested in obtaining an optimum "smoothing solution," i.e., one that uses the information at all "times" (z values) to construct the estimate at any given "time." In what follows, we present an adaptation to our problem of a batch type target-tracking technique, proposed by Morefield [39], which lends itself to a smoothing formulation.

We begin by introducing some notation and definitions.

Ordered Detection Set Ω is defined as the set of detections in the whole domain $\mathfrak{D}$, ordered in lexicographical order

$$\Omega \triangleq \left\{ \bar{\zeta}(1), \bar{\zeta}(2), \cdots, \bar{\zeta}(L) \right\}. \qquad (4.24)$$

The ith element ω_i of this set corresponds to a detected primitive $\zeta_k(z)$, for some k and z. The cardinality of this set is $\|\Omega\| = L \cdot J_{max}$.

Object Cluster λ^j is the set of data points ω_i corresponding to detected primitives belonging to a hypothesized object O_j^*, i.e.,

$$\lambda^j \triangleq \left\{ \omega_i : \omega_i \text{ belongs to } O_j^*, i \in \left\{ 1, \cdots, \|\Omega\| \right\} \right\}. \qquad (4.25)$$

Two additional constraints apply to the definition of λ^j: 1) it includes at most one data point at any value of z and 2) the data points in λ^j come from a contiguous range of z values. We will use $\lambda^j(z)$ to denote the element of λ^j corresponding to a primitive detected at position z. The number of primitives in λ^j is

$$L_j^* \triangleq \|\lambda^j\| \leq L \qquad (4.26)$$

corresponding to a hypothesis on the object length.

Association Hypothesis $\mathfrak{H}$ is a nonnegative integer J^* and a complete partitioning of Ω into J^* disjoint object clusters $\{ \lambda^j, j = 1, \cdots, J^* \}$ and a false alarm cluster Φ consisting of data points rejected from all λ^j. Thus, we have

$$\Phi = \Omega - \bigcup_{j=1}^{J^*} \lambda^j \qquad (4.27\text{-a})$$

$$\lambda^i \cap \lambda^j = \phi \quad i \neq j \qquad (4.27\text{-b})$$

$$\mathfrak{H} = \left\{ J^*, \left\{ \lambda^j \right\}_{j=1}^{J^*} \right\} \qquad (4.27\text{-c})$$

The requirement (4.27-b) that the distinct clusters be disjoint, corresponds, of course, to the assumption that no two objects occupy the same point in space, or share a primitive. The set S of all such valid hypotheses is defined as

$$S = \left\{ \mathfrak{H} : \text{the constraints } (4.27) \text{ are satisfied} \right\}. \qquad (4.28)$$

(a) z (b) λ_1 (c) λ_2 (d) λ_3

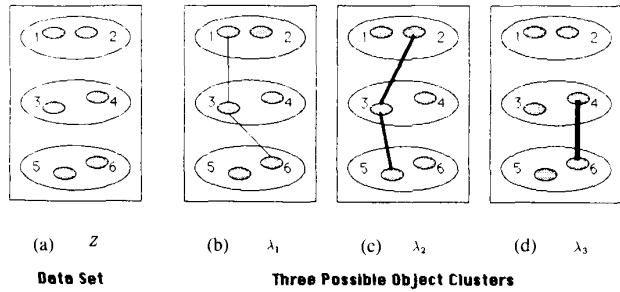**Data Set**   **Three Possible Object Clusters**

Fig. 6. Detected primitives and possible object clusters.

To illustrate the above definitions, consider the case [Fig. 6(a)] when the set of detections consists of six data points,

$$\Omega = \left\{ \omega_1, \omega_2, \cdots, \omega_6 \right\}.$$

Three possible object clusters [Figs. 6(b)–(d)] are

$$\lambda^1 = \left\{ \omega_1, \omega_3, \omega_6 \right\}, \lambda^2 = \left\{ \omega_2, \omega_3, \omega_5 \right\}, \lambda^3 = \left\{ \omega_4, \omega_6 \right\} \qquad (4.29)$$

Two valid association hypothesis and their corresponding false alarm clusters are $\mathfrak{H}_1 = \{ 2; \lambda^2, \lambda^3 \}$, $\mathfrak{H}_2 = \{ 1; \lambda^2 \}$, $\Phi_1 = \{ \omega_1 \}$, and $\Phi_2 = \{ \omega_1, \omega_4, \omega_6 \}$. The corresponding primitive-to-object associations are shown in Figs. 7(a) and (b), respectively. An instance of an *invalid* data partition [Fig. 7(c)], is $\mathfrak{H}_3 = \{ 2; \lambda^1, \lambda^2 \}$, with $\lambda^1 \cap \lambda^2 = \omega_3 \neq \phi$, in violation of (4.27-b).

A hypothesis $\mathfrak{H}$ thus completely specifies the data to object association. The choice of $\mathfrak{H} \in S$, which is an essential step in solving our problem, can now be formulated as a multihypothesis test. Here too, we employ the MAP criterion

$$\mathfrak{H} = \arg \max_{\mathfrak{H} \in S} P(\mathfrak{H} \mid \Omega) \qquad (4.30)$$

which is known [23] to minimize the associated Bayes risk, if errors are equally weighted and correct decisions not penalized. By Bayes' rule,

$$P(\mathfrak{H} \mid \Omega) = \frac{p(\Omega \mid \mathfrak{H}) \cdot P(\mathfrak{H})}{p(\Omega)}. \qquad (4.31)$$

The *a priori* association probability $P(\mathfrak{H})$, will be assumed constant for simplicity, although models, such as (2.8), of *a priori* information on object size (i.e., on the distribution of L_j) and of false alarm distribution could be incorporated. The independence of the objects conditioned on the constraint (4.27-b) [see also (2.8)], and of the estimation noise v_j together imply independence of the clusters λ^i conditioned on a specific association hypothesis $\mathfrak{H}$, so that, by (4.27-a) we obtain the following factored form, which is key to the efficient computation of the subsequent steps of the algorithm:

$$p(\Omega \mid \mathfrak{H}) = \prod_{j=1}^{J^*} p(\lambda^j \mid \mathfrak{H}, \lambda^j \in \mathfrak{H}) p(\Phi \mid \mathfrak{H}). \qquad (4.32)$$

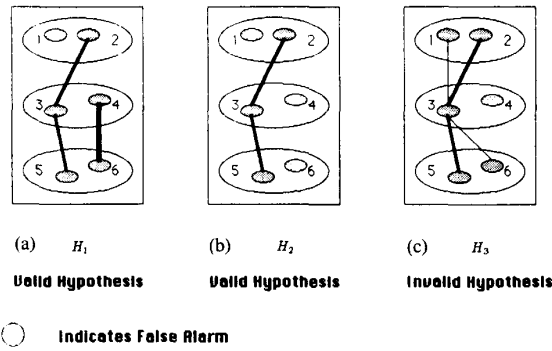| (a) H_1	(b) H_2	(c) H_3
Valid Hypothesis	Valid Hypothesis	Invalid Hypothesis

○   Indicates False Alarm

Fig. 7. Possible object hypotheses.

Assuming an equal likelihood of false alarms anywhere in the volume V of the probed domain, we obtain

$$p(\Phi \mid \mathcal{K}) = \left(\frac{1}{V}\right)^{n_\Phi} \qquad (4.33)$$

where the number of false alarms is $n_\Phi = L \cdot J_{max} - \Sigma_{j=1}^{J^*} L_j^*$ by (4.27-a), and V is in units of the volume of a typical primitive. Taking the logarithms of (4.31)–(4.33) and dropping terms independent of $\mathcal{K}$, we obtain the new problem

$$\max_{\mathcal{K} \in S} \sum_{j=1}^{J^*} \left[L(\lambda^j) + L_j^* \log V \right], \qquad (4.34)$$

where $L(\lambda^j)$ is the log-likelihood function of λ^j determined from the object and measurement models

$$L(\lambda^j) = \log p(\lambda^j \mid \mathcal{K}, \lambda^j \in \mathcal{K}). \qquad (4.35)$$

The computational difficulty in solving (4.34) by direct enumeration of the different $\mathcal{K}$ in S can be appreciated by evaluating the cardinality of S. Given that there are J_{max} detections in each slice, $\|S\|$ is bounded below by the number of association hypotheses with exactly $J^* = J_{max}$ objects of identical length $L_j^* = L$. This number is readily evaluated, yielding the following bound:

$$\|S\| \geq (J_{max}!)^{L-1}. \qquad (4.36)$$

Calculation of the object log-likelihood (4.35) (which in itself may be computationally expensive), for $\|S\|$ hypotheses, becomes infeasible for all but the smallest problems. For example, for a modest size problem (like the one simulated in Section VI) with $J_{max} = 5$ and $L = 100$ we have $\|S\| \geq 10^{272}$.

A key to reducing the computational complexity of (4.34) is the observation that different hypotheses $\mathcal{K} \in S$ are often not disjoint, with some objects appearing in more than one hypothesis. It is therefore inefficient to first form a hypothesis $\mathcal{K}$ and then check the feasibility of its objects, since the same object $O_j \in \mathcal{K}$ may appear in another hypothesis. Fortunately, the conditional independence of the data clusters allows us to decompose (4.34) into a step of feasible object construction followed by an integer optimization program, greatly reducing the computations.

C. Step 2: Feasible Object Construction

A clustering procedure is implemented in this step to detect clusters λ^j in Ω that are "reasonable" to incorporate in an assignment hypothesis $\mathcal{K}$, given the object and measurement models (2.6) and (4.22)–(4.23).

Owing to the mutual independence of the different λ^j conditioned on the hypothesis $\mathcal{K}$, relation (4.35) can be rewritten as

$$L(\lambda^j) = \log p[\lambda^j \mid model\ (2.6), (4.22)-(4.23)], \qquad (4.37)$$

so that $L(\lambda^j)$ can be evaluated independently of all the other entries in $\mathcal{K}$. The cost to be maximized in (4.34) is seen therefore to consist of a sum of terms, with each term depending only on its corresponding entry λ^j in $\mathcal{K}$. While complete decoupling of the problem into individual maximizations of each of these terms is impossible due to the coupling introduced by (4.27), the following scheme achieves partial decoupling.

If (4.34) is to be maximized, it makes sense to require that objects that are candidates for inclusion in $\mathcal{K}$ satisfy

$$L(\lambda^j) \geq \alpha_{L_j^*} \qquad (4.38)$$

for some threshold $\alpha_{L_j^*}$, which may be a function of the object's length L_j^*. Relation (4.38) is recognized as a likelihood test, rejecting object clusters whose likelihood is below some threshold. An object cluster O_j is incorporated into a feasible object set F, if it passes the hypothesis test (4.38). Thus, the feasible object set is defined by

$$F = \{ \lambda^j : \lambda^j \subset \Omega, L(\lambda^j) \geq \alpha_{L_j^*} \}. \qquad (4.39)$$

Only hypotheses $\mathcal{K} \subset F$ formed using the objects in F will be considered in the subsequent Bayesian decision process, effectively substituting a pruned feasible set F for the larger original set S. The cardinality $\|F\|$ of this set may be, however, significantly larger than the final number J^* of estimated objects. The threshold $\alpha_{L_j^*}$ is a critical design parameter in this procedure, determining a tradeoff between the accuracy of the algorithm and the computational requirements of the subsequent step: it determines the probability that a real object is mistakenly excluded from F and only parts of it recovered as smaller objects by an iterative post processing step (see below), versus the number of objects accepted into F that determine the size of the decision problem in the next step. The elimination of low probability (infeasible) hypotheses in this process contributes to great computational savings.

Note that conditioned on the hypothesized assignment (described by λ^j) of detections to the jth object, the measurement equation in 4.22 becomes

$$\lambda^j(z) = Hx_j(z) + v_j(z), \qquad (4.40)$$

where $\lambda^j(z)$ represents the measurement vector at position z associated with the jth object. To construct a feasible object corresponding to λ^j, we use the result that the

log-likelihood in (4.38) can be conveniently computed in terms of the white Gaussian innovation sequence [23]

$$\delta_j(z) = \lambda^j(z) - H\hat{x}_j(z|z - 1),$$

$$z = z^*_{1,j}, \cdots, z^*_{1,j} + L^*_j \quad (4.41)$$

where the "predicted estimate"

$$\hat{x}_j(z|z - 1) = E[x_j(z)|\lambda^j(z^*_{1,j}), \cdots, \lambda^j(z - 1)] \quad (4.42)$$

is a minimum variance causal estimate of the state sequence $x_j(z)$ of the object, based on "past" measurements in λ^j. Since the model (2.6), (4.40), (4.23) is in the form of a classical linear state-space model, the estimate in (4.42) is sequentially generated by a *linear* Kalman filter which is applied to the data λ^j in an increasing z order. For ease of reference, the equations of the Kalman filter [23] have been included in Appendix B. The resulting expression (Appendix B) for the log-likelihood function (4.33) for the object is given by

$$L(\lambda^j) = \tfrac{1}{2}L^*_j \dim \left[\delta_j(z)\right] \cdot \log 2\pi + \tfrac{1}{2} \sum_{z=z_{1,j}}^{z^*_{1,j}+L^*}$$

$$\cdot \log \left|R_\delta(z)\right| + \tfrac{1}{2} \sum_{z=z_{1,j}}^{z^*_{1,j}+L^*} \delta'_j(z) \, R_\delta^{-1}(z) \, \delta_j(z) \quad (4.43)$$

where $R_\delta(z)$ is the innovation covariance, $|R_\delta(z)|$ the determinant of this covariance, and $\dim [\zeta_i(z)]$ is the dimension of $\zeta_i(z)$. For models (2.6), (4.40), (4.23) with deterministic parameters (disregarding adaptive features which may be implemented to adjust model parameters in response to data collected), $R_\delta(z)$ is deterministic (and may be precomputed), and L^*_j is known for any given λ^j. Except for the first two terms in (4.43), which only shift the mean, the likelihood $L(\lambda^j)$ is a sum of quadratic forms in the white innovation vectors, and it is therefore χ^2 distributed with $\dim [\zeta_i(z)] \cdot L^*_j$ degrees of freedom [36]. It follows that the threshold $\alpha_{L^*_j}$ is readily set to fix the desired level of the test.

In practice, a feasible object is constructed sequentially by scanning the data with a depth first, backtracking search procedure. As each data point is added to a feasible object, the estimate (4.42) and the innovation (4.41) are recursively updated using the Kalman filter. The update of the cumulative object likelihood then involves only the addition of an extra term to the sums in (4.43). As many data points are added to an object as possible, before the test (4.38) is violated. When the test is failed, the algorithm backtracks, and examines another branch of the search tree. The final likelihood value associated with each feasible object is stored, along with the sequence λ^j of associated detections, for use in the next step of the algorithm.

For computational efficiency, the test (4.38) is pre-

ceded by coarser tests [39]. In particular, we test a distance measure d_{ij} between primitives $\zeta_i(z)$ and $\zeta_j(z + 1)$, to determine if they could possibly be assembled into a feasible object. d_{ij} can be a function of the the Euclidean distance between the locations of the corresponding primitives, and of their relative orientations. Primitives whose distance d_{ij} is below some threshold are declared *adjacent*. The threshold is set such that the probability of detection (i.e., correct decision on adjacency for neighboring primitives of an object) is close to unity, effectively guaranteeing that no real objects are excluded by this procedure. The adjacency relations determined by this coarse test, are used to guide the above depth-first search, thereby eliminating unnecessary evaluations of (4.43).

The main advantage of the sequential construction of F described above is its relatively low computational cost. Note however, that due to the depth-first search, it is biased towards long objects. Thus, although parts of a feasible object may be themselves feasible objects, they will not be produced by the algorithm. This may be justified by assumption that if a long object has been declared feasible by its likelihood exceeding a certain threshold, then the probability is low that its primitives constitute several distinct objects. In some cases, this bias must be overridden in favor of global optimality, as discussed below.

Consider the case where two feasible objects in F share one or more primitives. Eventually, only one of these objects will be included in the winning association hypothesis. If some of the primitives of the rejected object are not included in any other feasible object, these primitives will be declared as false alarms, although in fact they may together form a *shorter* object that would have been admitted as feasible, had it been considered in the first place. To address this difficulty, the processes of forming the feasible objects and choosing an association hypothesis are iterated on the data set, with the false alarm data cluster of one iteration, serving as the data set Ω of the next. Thus primitives that have been misclassified as false alarms can form an object in the subsequent iteration. The iteration is stopped when no further change occurs in Φ.

D. Step 3: Discrete Optimization

A subset $\{(\lambda^j)\}_{j=1}^{J^*}$ of F must now be chosen to maximize (4.34) subject to (4.27). This problem can be reformulated as a linear integer optimization problem, by adopting the following definitions. Let λ_j be a binary vector of length $\|\Omega\|$ describing an object cluster λ^j determined in the previous step, by using the following convention for its lth component:

$$\lambda_{i,j} = \begin{cases} 1 & \text{if } \omega_i \in \lambda^j \\ 0 & \text{otherwise.} \end{cases} \quad (4.44)$$

The whole set $F = \{\lambda^j, j = 1, \cdots, \|F\|\}$ containing $\|F\|$ feasible object clusters can be now alternatively described by the $\|\Omega\| \times \|F\|$ matrix $\Lambda = [\lambda_1, \cdots, \lambda_{\|F\|}]$, constructed from λ_j as columns. Let τ be a binary vector

of length $\|F\|$, defined by

$$\tau_j = \begin{cases} 1 & \text{if } \lambda^j \in \mathcal{H} \\ 0 & \text{otherwise} \end{cases} \qquad (4.45)$$

Then τ describes a hypothesis $\mathcal{H} \subset F$, which picks $J^* \leq \|F\|$ objects from F, and (4.27) can be expressed as

$$\Lambda\tau \leq [1, 1, \cdots 1]' \qquad (4.46)$$

where the vector inequality is considered satisfied if and only if all the components of the vector satisfy the inequality. Finally, define the cost associated with the choice of λ^j as

$$C_j = L(\lambda^j) + L_j^* \cdot \log V, \qquad (4.47)$$

and the cost vector $C = [C_1, \cdots C_{\|F\|}]'$. Note that $L(\lambda^j)$ has been computed in the previous step in (4.43), for each feasible object.

To illustrate this notation, consider Fig. 7 and the object clusters in (4.29). If F consists only of these three clusters, the matrix Λ is

$$\Lambda = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix},$$

while the τ vectors corresponding to the association hypotheses $\mathcal{H}_1 = \{2; \lambda^2, \lambda^3\}$ of Fig. 7(a), $\mathcal{H}_2 = [1; \lambda^2]$ of Fig. 7(b), and $\mathcal{H}_3 = \{2; \lambda^1, \lambda^2\}$ of Fig. 7(c) are, respectively,

$$\tau_1 = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} \qquad \tau_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \qquad \tau_3 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}.$$

We also have for these vectors

$$\Lambda\tau_1 = \begin{pmatrix} 0 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} \qquad \Lambda\tau_2 = \begin{pmatrix} 0 \\ 1 \\ 1 \\ 0 \\ 1 \\ 0 \end{pmatrix} \qquad \Lambda\tau_3 = \begin{pmatrix} 1 \\ 0 \\ 2 \\ 0 \\ 1 \\ 1 \end{pmatrix},$$

so that both τ_1 and τ_2 satisfy (4.46), as valid hypotheses should, and τ_3 violates (4.46), as expected of an invalid hypothesis.

Claim: With the above notation, the problem

$$\max_{\tau \in S^\tau} C'\tau \text{ subject to } \Lambda\tau \leq [1, 1, \cdots 1]' \qquad (4.48)$$

where $S^\tau = \{$ all binary vectors of length $\|F\| \}$, is equivalent to the original problem (4.34), (4.38), (4.27), provided that all the feasible objects have been included in F.

Proof: Simple substitution of the above definitions into (4.34).

The problem (4.48) is known as the set packing problem in 0-1 integer programming (cf. [40]), and specialized efficient algorithms for its solution have been devised (e.g., [41, 42]), taking advantage of the typical sparseness[5] of the matrix Λ. The performance of these algorithms may be significantly improved by preprocessing of Λ, to actually decouple (4.48) into a set of smaller problems [39]. Algorithms for the solution of the problem (4.48) are described in the above references and in the extensive general literature on integer programming (e.g., [43], [44]).

E. Step 4: Fixed Interval Linear Smoothing

At the completion of the previous step, an optimum association hypothesis $\mathcal{H} = \{\hat{J}, \{\hat{\lambda}^j\}_{j=1}^{\hat{J}}\}$ is available, specifying the estimated number of objects $\hat{J}$, their starting positions $\hat{z}_{1,j}$, lengths $\hat{L}_j, j = 1, \cdots, \hat{J}$, and which detected primitives belong to which object. We wish now to determine optimum estimates of the state sequences X_j, $j = 1, \cdots, \hat{J}$ of these objects.

Given the models (2.6), (4.40), (4.23) and the optimal association hypothesis $\mathcal{H}$, the problem decomposes (once again, owing to the conditional independence of the objects) into $\hat{J}$ independent linear Gaussian fixed interval smoothing problems [23], in which it is required to find the MAP estimate $\hat{X}_j(z \mid \hat{L}_j)$ of the state X_j given all the measurements in λ^j. Under the Gaussian assumption, the latter estimate coincides with the minimum mean square error estimate of the X_j, which is obtained by applying the two-pass formula of Rauch, Tung, and Striebel [45] (Appendix B), which we have chosen out of the several well-known fixed interval smoothing formulae [46], individually to the data sets $\lambda^j \in \hat{H}$.

A computed covariance of the estimation error is also available from the smoothing step, and may be used to assess the quality of the estimates. This covariance, however, reflects only the uncertainty associated with the measurement and process model noises, leaving out the uncertainty associated with the Bayesian decision process of Steps 2 and 3. It should be therefore regarded as a lower bound on the actual estimate variance, becoming tighter with higher measurement SNR and with better object spatial separation.

This step completes the solution of the estimation-detection/reconstruction problem, and the estimated density of the collection of objects in the probed domain is obtained by substituting $\hat{X}_{1:j}$ into (2.9) and (2.2).

---

[5]Λ will be typically less than $1/J_{max}$ full: while the length of a column of Λ is $\|\Omega\| = L \cdot J_{max}$ the average number of nonzero entries in it will be approximately equal to the average length $\bar{L}$ of an object; and since $\bar{L} \leq L$, it follows that the average density of the matrix is $\bar{L}/\|\Omega\| \leq 1/J_{max}$.

V. Computational Requirements

The set packing problem (4.48) is known to be NP-complete [47], and therefore, since it is conjectured that no polynomial-time algorithm exists for NP-complete problems, the best available upper bound on the computation will be exponential in the problem size. Using such a bound to predict the computational cost would invariably suggest that the feasibility of our whole algorithm is limited to relatively small problems.

Fortunately, efficient algorithms (e.g., those mentioned above) taking advantage of the typical sparseness of our constraint matrix, provide a solution to relatively large problems with a reasonable amount of computation, *on the average* (over a wide selection of test problems). Thus, we note that in the scheduling of airline crews, the problem (4.48) is solved routinely with several hundred constraints and thousands of variables. In the target tracking application, a problem with an 80×40, 2 percent full Λ matrix required 1.3 s on a CDC 7600 [39], whereas one with a 200×200, 10 percent full Λ matrix has been solved in less than one second on the average (virtual CPU seconds) on an IBM 3033 computer [48], using Toyoda's [49] algorithm to obtain "good" approximate solutions to randomly generated sparse test problems. Hence, although it appears extremely difficult to predict theoretically the computational requirements of any of these algorithms, the reported experience strongly suggests that the discrete optimization step can be easily handled by modern computers and currently available algorithms, even for relatively large problems.

A detailed evaluation [15] of the computational requirements of the other steps of the algorithm reveals that the computation is dominated by the first step, of primitive estimation. Consequently, we concentrate on this step.

The dominating computation for primitive estimation is the maximization in (4.16). For each choice of c and γ, the convolution in (4.18) must be performed, and then the maximum of the backprojected log-likelihood must be found. We first compute the requirements for these two procedures, assuming that floating point additions, multiplications, and comparisons take approximately the same amount of time. These expressions are for the sampled measurement case.

- *Convolution:* For convolving a small symmetric kernel with large vector, direction convolution is more efficient than using FET's. Owing to the symmetry of the half-ellipse kernels g_0, the flop count for a single convolution is $3 \cdot n_k N$ where n_k is the number of samples in an average kernel and N is from (2.16).

- *Backprojection:* The log-likelihood is maximized on a rectangular lattice within a single slice of the probed domain (2.1). At each lattice point, we sum the contributions from each view (4.19) and compare to the current maximum. Therefore, the operation count is $(M + 1) \cdot \pi (N/2n_s)^2$ where n_s is the sampling interval used for coarse sampling.

One search technique is to evaluate the likelihood for each possible radius, orientation, and axis ratio which are discretized to n_a, n_ϕ, and n_λ sample points, respectively. Although inefficient, this will provide an upper bound for the computational requirements. For each primitive, and for each possible parameter set, we must perform a convolution in each view, and then backproject. The operation count is therefore

$$J n_a n_\phi n_\lambda \left(M 3 n_k N + (M + 1) \cdot \pi \left(\frac{N}{2n_s} \right)^2 \right)$$

$$\approx J n_a n_\phi n_\lambda M N \left(3 n_k + \pi \frac{N}{(2n_s)^2} \right).$$

For example, if $N = 256$, $M = 4$, $J = 10$, $n_a = n_\phi = n_\lambda = 5$, $n_k = 7$, and $n_s = 2$, the operation count is $\approx 9 \cdot 10^7$.

The contribution to the operation count by subsequent iterations of the AM algorithm is negligible since these require only local searches in position, so the dominant $O(N^2)$ term is replaced by a much smaller term.

A more efficient procedure is to maximize over each parameter separately, i.e., first choose an "average" size primitive and find the best position, which requires convolution in each view and a backprojection. Then search over the orientation, axis ratio, and radius parameters individually, with only a local search in position, say in a $W \times W$ square about the initial estimate. This gives a more realistic figure for the operation count:

$$J \left(M 3 n_k N + (M + 1) \cdot \pi \frac{N^2}{2n_s} \right)$$

$$+ J (n_a + n_\phi + n_\lambda)(M 3 n_k W + (M + 1) \cdot W^2).$$

For the same example, with $W = 10$, the operation count is now $\approx 1 \cdot 10^6$. However, this latter procedure may require more iterations to achieve the same accuracy.

The above expressions are counts per slice, so the total operation count would then be L times larger. Since each slice is processed independently, parallel processing could be utilized to increase speed.

Although the computational requirements of the hierarchical algorithm are nontrivial, they are put in perspective by considering the alternative dynamic programming approach considered in Section III. We conclude that the hierarchical algorithm has a clear computational advantage (of hundreds of orders of magnitude) in solving multiple object problems.

VI. Implementation and Simulation Results

A simplified version of the above algorithm has been implemented and tested on simulated projection images. 3D objects are generated as explained below, their projections are synthesized, and white Gaussian noise is added. The current implementation is restricted to primitives with circular cross-sections of unknown radius and contrast. The shape parameter Γ is therefore just a and d.

A. Object Model

To demonstrate robustness of the algorithm to deviations from the Markov state model (2.6), objects were generated by using a few hand generated parameter values connected smoothly by splines. Two distinct sets of objects were generated—a training set and a test set. In the estimator, we chose to use first-order Markov models for the components of $\Gamma_j(z)$, and a second-order evolution model for $c_j(z)$, with the discrete second derivative $\Delta_c^2(z)$ $\triangleq c(z + 1) - 2c(z) + c(z - 1)$ driven by a white noise sequence, whose covariance corresponds to the root mean squared (rms) curvature of the center axis. The resulting definitions for the state vector and model matrices are

$$x_j(z) \triangleq \left[c_{\xi,j}(z - 1), \; c_{\eta,j}(z - 1), \; c_{\xi,j}(z), \; c_{\eta,j}(z), \right.$$

$$\left. a_j(z), \; d_j(z) \right] \tag{6.1}$$

$$A = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ -1 & 0 & 2 & 0 \\ 0 & -1 & 0 & 2 \\ & & & & 1 & 0 \\ & \mathbf{0} & & & 0 & 1 \end{bmatrix}$$

$$B = \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 0.2 & 0 \\ 0 & 0.2 \\ & & 0.2 & 0 \\ & \mathbf{0} & 0 & 0.1 \end{bmatrix}. \tag{6.2}$$

The parameters x_0 and Π of the object model (2.6) and the corresponding initial estimate x_0 and its covariance Π_0 were set to 0 and ∞, respectively, for each of the objects, representing a diffuse prior for the object parameters (i.e., corresponding to having no prior information about the state). The covariances B and R of (2.6) and (4.40) were then optimized using the algorithm on the training set of objects, and then applied to the test set. Obviously in practice it would be better to use a larger training set derived, e.g., from actual objects.

B. Measurement Model

The projections of the objects in a zero-density background[6] were computed using (2.12), and corrupted with pseudorandom white Gaussian noise with variance $\sigma_n^2 = 4$. Fig. 8(a) and (b) are two of four noisy projections at angles $0°$, $45°$, $90°$, and $135°$, discretized to 128 by

[6]As would result after subtracting off the contribution of a known background density (see Section II-B), or following a temporal subtraction procedure. The latter procedure, which is common practice in medical applications, relies on subtracting pairs of projections taken before and after administration of a contrast agent into the imaged object. The background which is invariant in the two projections is thus cancelled.
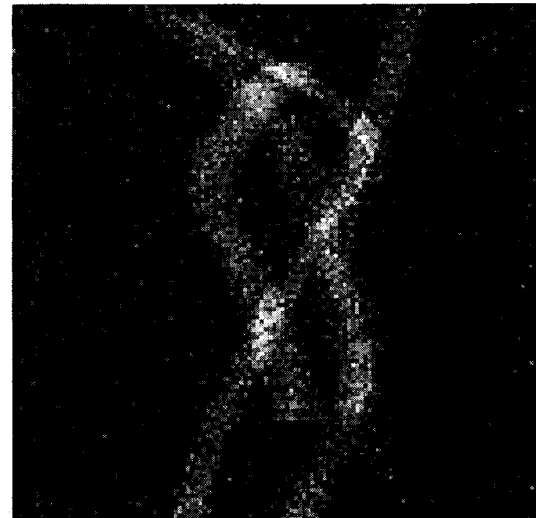
(a)
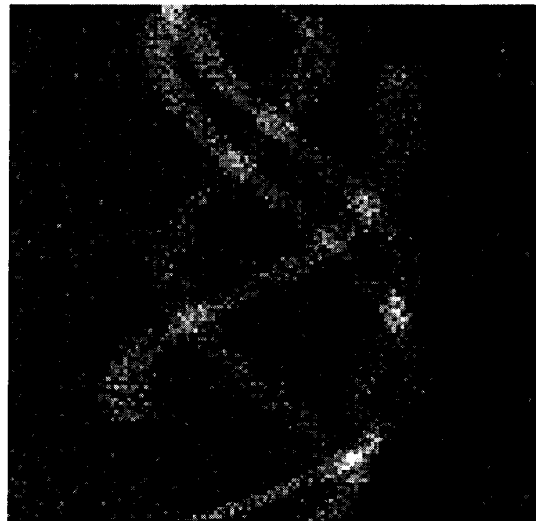
(b)

Fig. 8. Noisy projection measurement data. (a) $\theta = 0°$. (b) $\theta = 90°$.

128 pixels. In each projection, the signal-to-noise ratio for a given primitive, defined as the ratio of the projection height to noise standard deviation, is SNR $= 2ad/\sigma_n$ where a and d are, respectively, the radius and contrast of the primitive. For the test objects, the average SNR is 3.3. Although the outline of the objects can be visually detected in this example, quantitative assessment of the cross-section of the objects is very difficult.

C. Implementation

The training set of objects was used to determine the error standard deviations of the primitive parameter estimates obtained in the first step. These are used for the measurement model (4.40), (4.23), and were set to 0.7, 0.5, and 0.3 for c, a, and d, respectively.

The number of primitives per slice is assumed to be bounded above by $J_{max} = 4$, so as mentioned previously,
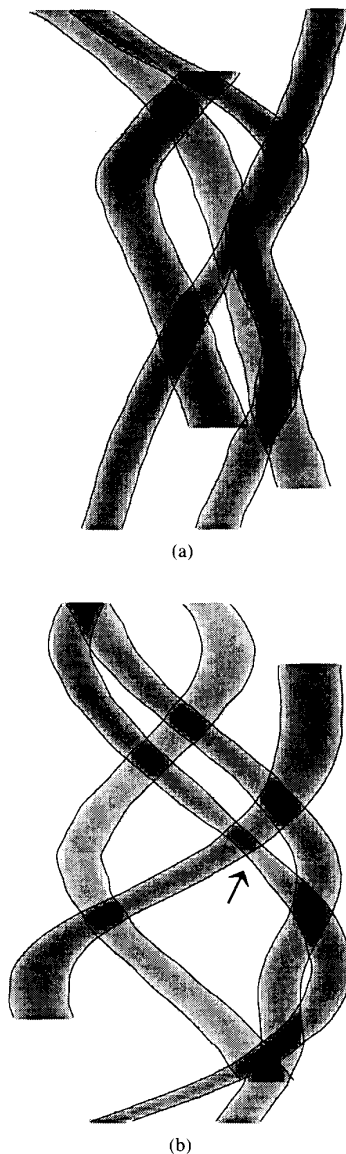
(a)

(b)

Fig. 9. Outline of reprojection of estimated objects superimposed on projections of true objects (shaded). (a) $\theta = 0°$. (b) $\theta = 90°$.

Fig. 10. True (solid) versus estimated (dashed) object parameters.

we estimate 4 primitives per slice, some of which will be false alarms. The threshold for sequential hypothesis testing in the construction of feasible objects was set at the 1 percent confidence level for the chi-squared distribution, i.e., α in (4.39) was chosen so that the probability of rejecting a correct primitive from an object is less than 1 percent. Nevertheless, for the examples considered, the number of feasible objects in F was small enough that the 0–1 integer programming minimization could be performed by hand.

The algorithm was implemented without regard to computational efficiency in C on a SUN 3/50, and took approximately 5 min to process the data. The radii were coarsely sampled to 1 pixel for initialization, and refined to 0.5 pixels by iteration.
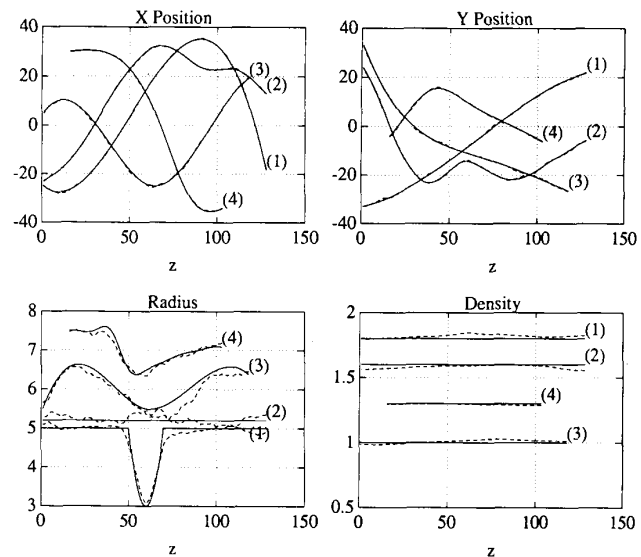
D. Results

Fig. 9(a) and (b) show two projections of the estimated objects superimposed on the (noiseless) projections of the actual objects. With a suitable graphics package, a 3D display of the estimated objects could be synthesized from their complete 3D representation, which is provided by the algorithm. The algorithm has tracked the 3-D evolution of the objects despite the low SNR and limited number of views. Note that the algorithm has successfully tracked the narrowing (indicated by the arrow) in the longest object, which was difficult to see in the projections. Fig. 10 allows a quantitative comparison of the actual and estimated object parameters. The smoothed position errors are less than a pixel, and the radius errors are less than half a pixel. This subpixel accuracy is a desirable feature of the parametric approach.

VII. Conclusions

We have presented an estimation framework for 3D reconstruction from limited-view noisy projections of a domain containing multiple objects, based on dynamic stochastic object models. Rather than attempt to reconstruct an arbitrary 3D distribution from the insufficient data, specific, quantitative information about objects in the probed domain is extracted, allowing, if desired, the production of a synthetic 3D display of the objects. The reduction of dimensionality achieved through this formulation is the key to the effectiveness of the approach with limited and noisy projection data: the state sequences that need to be determined to completely specify the objects have only a small number of degrees of freedom (typically in the hundreds), as compared to e.g., 10^6 for an arbitrary distribution in a $100 \times 100 \times 100$ voxel volume.

Since the globally optimal solution of the resulting joint detection-estimation problem is computationally infeasible for all but the smallest problems, a hierarchical struc-

ture is used in the algorithm to reduce the computation to reasonable amount for moderately sized problems. A divide and conquer strategy at several levels, extensive data reduction at the first step, successive hypothesis pruning, and a mix of continuous and combinatorial optimization, 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, and by a scheme to combine the results of these steps subject to a global optimality criterion. In general, such a decomposition of the problem does not preserve global optimality. However, significant reduction in computation with little loss in performance compared to the globally optimal solution can be obtained by judiciously choosing decision thresholds in the various steps of the algorithm, as described in the paper. This observation is supported by simulation results reported elsewhere [50], comparing this algorithm with the single-step approximately globally optimal algorithm of [16].

The hierarchical divide and conquer structure of our algorithm is reminiscent of approaches used in artificial intelligence (AI) in general [51] and computer vision in particular [52]. However, in distinction from most of those approaches, which often use ad-hoc criteria and procedures, our algorithm is characterized by well-defined quantitative optimality criteria, and by procedures that are either optimal, or represent a well-defined tradeoff between computation and optimality. These properties are based on a systematic probabilistic framework, in which both object and measurement processes are modeled by stochastic processes. 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 visual observer, the purpose of our algorithm is to perform a task of which neither the unaided visual observer nor current algorithms are capable—3D reconstruction from few projections at low SNR.

While the theoretical analysis of the performance of this algorithm is yet to be addressed, the simulation results are very encouraging, demonstrating "reconstructions" of a domain containing several objects from only 4 views in a 135° sector, at signal-to-noise ratios as low as 1.5.

The formulation and algorithm 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. An extension of this approach to remove this restriction is possible [15], and will be described in a forthcoming paper.

Appendix

A. Likelihoods for the AM Algorithm

When $\nu(t)$ in (4.1) is Gaussian white noise, the likelihood in (4.4) can be written, to within terms independent of $\hat{\alpha}$, as [32]

$$L_j(\alpha_j) = -\frac{1}{2N_0} \left\| y(t) - \sum_{j=1}^{J} \psi(t; \alpha_j) \right\|^2 \quad (A.1)$$

where the norm and inner product for scalar time functions are defined in (4.10) and (4.11), respectively. Fixing $\hat{\alpha}_{(j)}^n$ as in (4.6) we obtain

$$L_j(\alpha_j \mid \hat{\alpha}_{(j)}^n) = -\frac{1}{2N_0} \left\| z_j^n(t) - \psi(t; \alpha_j) \right\|^2 \quad (A.2)$$

where $z_j^n(t)$ is defined in (4.12). Expanding now the above norm and dropping $\| z_j^n(t) \|^2$ yields (4.8) and (4.9).

We now derive relations (4.16) and (4.17). Note first that by (4.9) we have

$$\mathcal{E}(\gamma_j) = d_j^2 \cdot \| g_0(t, \theta; \gamma_j) \|^2. \quad (A.3)$$

Substituting into (4.14) we obtain

$$L_j(c_j, d_j) = \frac{1}{N_0} d_j \langle z_j^n(t, \theta), g_0(t - \theta'c, \theta; \gamma_j) \rangle$$

$$- \frac{1}{2N_0} d_j^2 \cdot \| g_0(t, \theta; \gamma_j) \|^2. \quad (A.4)$$

Since the log-likelihood in (A.4) is quadratic in the contrast parameter d_j, its minimization with respect to d_j can be performed analytically for any values of the other parameters. The resulting estimate is

$$\hat{d}_j = \frac{\langle z_j^n(t, \theta), g_0(t - \theta'c, \theta; \gamma_j) \rangle}{\| g_0(t, \theta; \gamma_j) \|^2} \quad (A.5)$$

which, at the $n + 1$ iteration, coincides with (4.17).

Substituting the last expression into (A.4) yields

$$\hat{L}_j(c_j, \gamma_j) \triangleq \max_{d_j} L_j(c_j, \gamma_j, d_j)$$

$$= \frac{1}{2N_0} \frac{\langle z_j^n(t, \theta), g_0(t - \theta'c, \theta; \gamma) \rangle^2}{\| g_0(t, \theta; \gamma) \|^2} \quad (A.6)$$

which is the criterion maximized in (4.16).

B. Optimum Filtering and Smoothing

Since filtering, likelihood computation, and smoothing are applied individually and in identical manner to each of the feasible object clusters, we drop in this appendix the index j, for notational convenience.

1) Kalman Filter: Given the state-space model (2.6) and measurement (4.40), one version of the Kalman filter (KF) algorithm is given by [23] by the equations listed in Table I, which are applied recursively for $z = z_1^*, \cdots, z_1^* + L^*$.

2) Log-Likelihood from Innovations: The innovations are obtained from the measurement sequence $\{ \lambda(z) \}_{z_1^*}^{z_1^* + l^*}$ via (4.41) and the Kalman filter equations. It is readily verified that the converse is also true, i.e., the measurement sequence reconstructed via the same for-

TABLE I

Filtered Estimate	$\hat{x}(z\|z) = \hat{x}(z\|z - 1) + K(z)[\lambda(z) - H\hat{x}(z\|z - 1)]$
	$\hat{x}(z_1^*\|z_1^* - 1) = x_0$
Predicted Estimate	$\hat{x}(z + 1\|z) = A\hat{x}(z\|z)$
Kalman Gain	$K(z) = P(z\|z - 1)H'R_\delta^{-1}(z)$
Predicted Error Covariance	$P(z + 1\|z) = AP(z\|z)A' + BB'$
Filtered Error Covariance	$P(z\|z) = [I - K(z)H]P(z\|z - 1) \qquad P(z_1^*\|z_1^* - 1) = \Pi_0$
Innovation Covariance	$R_\delta(z) = HP(z\|z - 1)H' + R$

TABLE II

Fixed Interval Smoothed Estimate	$\hat{x}(z \mid z_1^* + L^*) = \hat{x}(z\|z) + G(z)[\hat{x}(z + 1 \mid z_1^* + L^*) - \hat{x}(z + 1\|z)]$
	$G(z) = P(z\|z)A'P(z + 1\|z)$
Smoothing Error Covariance	$P(z \mid z_1^* + L^*) = P(z\|z) + G(z)[P(z + 1 \mid z_1^* + L^*) - P(z + 1\|z)]A'$

mulae from the innovation sequence. The two are therefore statistically equivalent, and the likelihood of the measurement sequence can be written in terms of that of the innovation sequence as

$$p\left(\left\{\lambda(z)\right\}_{z_1^*}^{z_1^* + l^*}\right) = p\left(\left\{\delta(z)\right\}_{z_1^*}^{z_1^* + l^*}\right) = \prod_{z = z_1^*}^{z_1^* + L^*}$$

$$\cdot \, p\left(\delta(z)\middle| \text{model } (2.6), (4.40), (4.23)\right)$$

$$(\text{B.1})$$

where the factoring of the likelihood of the innovations sequence follows from the whiteness of the innovations. Now, since $\delta(z) \sim N(0, R_\delta(z))$, we have

$$p\left(\delta(z)\middle| \text{model } (2.6), (4.40), (4.23)\right)$$

$$= (2\pi)^{\dim\{\delta(z)\}/2}\left|R_\delta(z)\right|^{1/2}$$

$$\cdot \, \exp\left\{-\tfrac{1}{2}\delta'(z) \, R_\delta^{-1}(z) \, \delta(z)\right\}. \quad (\text{B.2})$$

Combining (B.1) and (B.2), taking logarithms, and reinserting the dependence on the index j, yields (4.43).

3) Smoothing Formulae: Given measurements as in (4.40), (4.23) over the interval $z = z_1^* \cdots z_1^* + L^*$, the fixed interval smoothed estimate of the state is

$$\hat{x}(z\middle|z_1^* + L^*) \triangleq E\left[x(z)\middle| \lambda^j(z), z = z_1^* \cdots z_1^* + L^*\right],$$

$$(\text{B.3})$$

i.e., the estimate at any point is based on data from the entire interval. The algorithm we use is based on the equations [23], [45] listed in Table II, which are propagated backwards (decreasing z) from the final conditions $\hat{x}(z_1^* + L^*|z_1^* + L^*)$ and $P(z_1^* + L^*|z_1^* + L^*)$. The latter quantities, as well as the filtered and smoothed estimates and either covariance, which are used in the smoothing algorithm, are obtained from the (forward pass of the) Kalman filter.

The computation of the smoothed error covariance, which is not used in the computation of the smoothed estimate, is optional. It may be used to assess the quality of the object estimates produced by the algorithm.

REFERENCES

[1] J. B. Ra and Z. H. Cho, "Generalized true three dimensional reconstruction algorithm," *Proc. IEEE*, vol. 69, pp. 668-670, May 1981.
[2] G. T. Herman, *Image Reconstruction from Projections.* New York: Academic, 1980.
[3] S. R. Deans, *The Radon Transform and Some of Its Applications.* New York: Wiley, 1984.
[4] R. M. Rangayyan, R. Gordon, and A. P. Dhawan, "Algorithms for limited-view computed tomography: An annotated bibliography and a challenge," in *Digest Topic. Meet. Indust. Appl. Comput. Tomog. NMR Imaging*, Opt. Soc. America, Hecla Island, Manitoba, Canada, Aug. 1984. (See also *Proc. IEEE Int. Conf. Comput., Syst., Signal Processing*, Bangalore, India, Dec. 1984.)
[5] D. J. Rossi and A. S. Willsky, "Reconstruction from projections based on detection and estimation of objects," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-32, pp. 886-906, 1984.
[6] A. K. Louis and F. Natterer, "Mathematical problems of computerized tomography," *Proc. IEEE*, vol. 71, pp. 379-390, Mar. 1983.
[7] P. Bloch and J. K. Udupa, "Application of computerized tomography to radiation therapy and surgical planning," *Proc. IEEE*, vol. 71, pp. 351-355, Mar. 1983.
[8] R. Gordon and G. T. Herman, "Three dimensional reconstruction from projections: A review of algorithms," *Int. Rev. Cytol.*, vol. 38, pp. 111-151, 1974.
[9] A. C. Kak, "Computerized tomography with x-ray, emission, and ultrasound sources," *Proc. IEEE*, vol. 67, pp. 1245-1272, 1979.
[10] P. G. Selfridge and J. M. S. Prewitt, "Organ detection in abdominal computerized tomography scans: Application to the kidney," *Comput. Graph. Image Processing*, vol. 15, pp. 265-278, 1981.
[11] W. Munk and C. Wunsch, "Ocean acoustic tomography: A scheme for large scale monitoring," *Deep Sea Res.*, vol. 26A, pp. 123-161, 1979.
[12] C. H. Slump and J. J. Gebrands, "A network flow approach to reconstruction of the left ventricle from two projections," *Comput. Graph. Image Processing*, vol. 18, pp. 18-36, 1982.
[13] J. K. Udupa, "Display of 3D information in discrete 3D scenes produced by computerized tomography," *Proc. IEEE*, vol. 71, pp. 420-431, Mar. 1983.
[14] Y. Bresler and A. Macovski, "A hierarchical Bayesian approach to reconstruction from projections of a multiple object 3-D scene," in *Proc. 7th Int. Conf. Pattern Recog.*, Montreal, P.Q., Canada, Aug. 1984, pp. 455-457.
[15] Y. Bresler, "Model based estimation techniques for 3-D reconstruction from projections," Ph.D. dissertation, Dep. Elec. Eng., Stanford Univ., 1985.
[16] Y. Bresler and A. Macovski, "3-D Reconstruction from projections with incomplete and noisy data by object estimation," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-35, pp. 1139-1152, Aug. 1987.
[17] Y. Bresler, A. Macovski, Y. Bresler, and A. Macovski, "Estimation of 3-D shape of blood vessels from x-ray images," *Proc. IEEE Int. Symp. Medical Images Icons, ISMII'84*, Arlington, VA, July 1984, pp. 251-258.
[18] K. M. Hanson and G. W. Wecksung, "Bayesian approach to limited-angle reconstruction in computed tomography," *J. Opt. Soc. Amer.*, vol. 73, pp. 1501-1509, Nov. 1983.
[19] K. Shmueli, W. R. Brody, and A. Macovski, "Estimation of blood vessels boundaries in x-ray images," *SPIE Conf. Digital Radiography*, Sept. 14-16, 1981.
[20] T. O. Binford, "Visual perception by computer," presented at IEEE Conf. Syst. Contr., Miami, FL, Dec. 1971.
[21] D. I. Hoult and R. E. Richards, "Signal-to-noise ratio of nuclear magnetic resonance experiment," *J. Magn. Reson.*, vol. 24, pp. 71-85, 1976.
[22] W. A. Edelstein et al., "Signal, noise, and contrast in NMR imaging," *J. Comput. Assisted Tomography*, vol. 7, pp. 391-401, 1983.
[23] A. P. Sage and J. L. Melsa, *Estimation Theory with Applications to Communications and Control.* New York: McGraw Hill, 1971.
[24] N. J. Nillson, "On the optimum resolution of radar signals in noise," *IRE Trans. Inform. Theory*, vol. 11, pp. 245-253, 1961.
[25] M. G. Lichtenstein and T. Y. Young, "The resolution of closely spaced signals," *IEEE Trans. Inform. Theory*, vol. IT-14, pp. 288-293, 1968.
[26] R. J. P. Figueiredo and A. Gerber, "Separation of superimposed signals by a cross-correlation method," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-31, pp. 1084-1089, Oct. 1983.

[27] A. M. Bruckstein, T. J. Shan, and T. Kailath, "The resolution of overlapping echoes," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-33, pp. 1357-1367, 1985.

[28] M. Feder and E. Weinstein, "Parameter estimation of superimposed signals using the EM Algorithm," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-36, pp. 477-489, Apr. 1988.

[29] H. Kwakernaak, "Estimation of pulse heights and arrival times," *Automatica*, vol. 16, pp. 367-377, 1980.

[30] R. Fletcher, *Practical Methods of Optimization*. New York: Wiley, 1987, 2nd ed., p. 18.

[31] I. Ziskind and M. Wax, "Maximum likelihood localization of multiple sources by alternating maximization," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Processing*, Dallas, TX, pp. 53.6.1-53.6.4, Apr. 1987.

[32] H. L. Van Trees, *Detection Estimation, and Modulation Theory, Part I*. New York: Wiley, 1968.

[33] M. J. D. Powell, "On search directions for minimization algorithms," *Math. Prog.*, vol. 4, pp. 193-201, 1973.

[34] J. Rissanen, "Modeling by shortest data description," *Automatica*, vol. 14, pp. 465-471, 1978.

[35] M. Wax and T. Kailath, "Estimating the number of signals by information theoretic criteria," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-33, pp. 387-392, Apr. 1985.

[36] T. S. Ferguson, *Mathematical Statistics: A Decision Theoretic Approach*. New York: Academic, 1967, p. 208.

[37] R. O. Duda and P. E. Hart, *Pattern Classification and Scene Analysis*. New York: Wiley, 1973.

[38] Y. Bar-Shalom, "Tracking methods in a multitarget environment," *IEEE Trans. Automat. Contr.*, vol. AC-23, pp. 618-626, 1978.

[39] C. L. Morefield, "Applications of 0-1 integer programming to multitarget tracking problems," *IEEE Trans. Automat. Contr.*, vol. AC-22, pp. 302-312, 1977.

[40] M. W. Padberg, "Covering, packing, and knapsack problems," in *Discrete Optimization Proceedings 1977*, Hammer, Johnson, and Korte Eds. Amsterdam, The Netherlands: North-Holland, 1979, pp. 265-287.

[41] J. F. Pierce, "Application of combinatorial programming to a class of all zero-one integer programming problems," *Manage. Sci.*, vol. 15, pp. 191-209, Nov. 1968.

[42] J. F. Pierce and J. S. Lasky, "Improved combinatorial programming for a class of all zero-one integer programming problems," *Manage. Sci.*, vol. 19, pp. 528-543, Jan. 1973.

[43] H. A. Taha, *Integer Programming: Theory, Applications, and Computations*. New York: Academic, 1975.

[44] A. Kaufman and A. Henry-Laborde're, *Integer and Mixed Programming: Theory and Applications*. New York: Academic, 1977.

[45] H. E. Rauch, F. Tung, and C. T. Striebel, "Maximum likelihood estimates of dynamic linear systems," *AIAA J.*, vol. 3, pp. 1445-1450, Aug. 1965.

[46] J. E. Wall, Jr., A. S. Willsky, and N. R. Sandell, Jr., "On the fixed interval smoothing problem," *Stochastics*, vol. 5, no. 1, 2, pp. 1-42, 1981.

[47] J. K. Lenstra and A. H. G. Rinnooy Kan, "Computational complexity of discrete optimization problems," in *Discrete Optimization Proceedings 1977*, Hammer, Johnson, and Korte, Eds. Amsterdam, The Netherlands: North-Holland, 1979, pp. 121-140.

[48] U. Suhl, "Implementing an algorithm: Performance considerations and a case study," in *Evaluating Mathematical Programming Techniques, Proceedings 1981*, J. M. Mulvey, Ed. Berlin, West Germany: Springer-Verlag, 1982.

[49] Y. Toyoda, "A simplified algorithm for obtaining approximate solutions to zero-one programming problems," *Manage Sci.*, vol. 21, no. 12, pp. 1417-1420, 1975.

[50] Y. Bresler, J. A. Fessler, and A. Macovski, "Model based estimation techniques for 3-D reconstruction from projections," *Machine Vision Appl.*, vol. 1, no. 2, pp. 115-126, 1988.

[51] N. J. Nillson, *Principles of Artificial Intelligence*. Palo Alto, CA: Tioga, 1980.

[52] D. H. Ballard and C. M. Brown, *Computer Vision*. Englewood Cliffs, NJ: Prentice Hall, 1982.

Yoram Bresler (S'82-M'85) received the B.Sc. and M.Sc. degrees from the Technion—Israel Institute of Technology, Haifa, Israel, in 1974 and 1981, respectively, and the Ph.D. degree from Stanford University, in 1985, all in electrical engineering.

From 1974 to 1979, he served as an electronics engineer in the Israeli Defense Force. From 1979 to 1981, he was a Consultant for the Flight Control Lab at the Department of Aeronautical Engineering, Technion, Israel, where he was involved in developing algorithms for autonomous TV aircraft guidance. From 1985 to 1987, he was a Research Associate at the Information Systems laboratory at Stanford University, where his research involved array signal processing and medical imaging. Since 1987 he has been an Assistant Professor at the Department of Electrical and Computer Engineering at the University of Illinois at Urbana-Champaign. His research interests include estimation theory and statistical signal processing and their application to imaging algorithms and sensor array signal processing.

Dr. Bresler is on the Editorial Board of Machine Vision and Applications. In 1988 and 1989 he received the Senior (best paper) awards from the IEEE Acoustics, Speech, and Signal Processing Society.
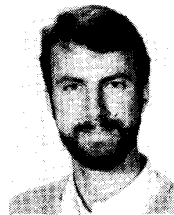
Jeffrey A. Fessler received the B.S. degree from Purdue University, West Lafayette, IN, in 1985, and the M.S. degree from Stanford University, Stanford, CA, in 1986 both in electrical engineering. Since 1986 he has been a National Science Foundation Graduate Fellow at Stanford, working toward the Ph.D. degree in electrical engineering and an M.S. in statistics.

His research interests include estimation theory and pattern recognition and their applications in medical imaging and bioengineering.
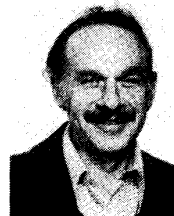
Albert Macovski (A'51-SM'56-F'76) received the B.E.E. degree from City College of New York in 1950, the M.E.E. degree from the Polytechnic Institute of Brooklyn in 1953, and the Ph.D. degree in electrical engineering from Stanford University in 1968.

From 1950 to 1957, he was a member of the Technical Staff at R.C.A. Laboratories. From 1957 to 1960, he was an Assistant and then Associate Professor at the Polytechnic Institute of Brooklyn, NY. From 1960 to 1971, he was a Staff Scientist at the Stanford Research Institute. Following one year as a special NIH fellow at the University of California Medical Center in San Francisco, he joined the Stanford Faculty as an Adjunct Professor and then full Professor of Electrical Engineering and Radiology, his present position. His research has been in a variety of imaging systems including television, facsimile, holography, and interferometry. During the past 15 years, he has been particularly involved with diagnostic imaging techniques in ultrasound, radiography and magnetic resonance.

In 1988, Dr. Macovski was elected to the National Academy of Engineering. He is a Fellow of the Optical Society of America, and he is a member of the American Association of Physicists in Medicine, Eta Kappa Nu and Sigma Xi. In 1958, he received the award from the IRE professional group on broadcast and television receivers. In 1973 he received the IEEE Zworykin Award. He has approximately 150 issued U.S. Patents and 150 publications in the fields of optics, electronics, television, imaging, radiography, ultrasonics, and magnetic resonance. In 1988, he was honored as *Inventor of the Year* by the Peninsula Patent Attorney Association of Northern California.