International Journal of Computer Vision, 133–144 (1987)

Signal Matching Through Scale Space

ANDREW WITKIN, DEMETRI TERZOPOULOS, AND MICHAEL KASS
Schlumberger Palo Alto Research, 3340 Hillview Avenue Palo Alto, CA 94304, USA

Abstract

Given a collection of similar signals that have been deformed with respect to each other, the general signal-matching problem is to recover the deformation. We formulate the problem as the minimization of an energy measure that combines a smoothness term and a similarity term. The minimization reduces to a dynamic system governed by a set of coupled, first-order differential equations. The dynamic system finds an optimal solution at a coarse scale and then tracks it continuously to a fine scale. Among the major themes in recent work on visual signal matching have been the notions of matching as constrained optimization, of variational surface reconstruction, and of coarse-to-fine matching. Our solution captures these in a precise, succinct, and unified form. Results are presented for one-dimensional signals, a motion sequence, and a stereo pair.

1 Introduction

Given a collection of similar signals that have been deformed with respect to each other, the general signal-matching problem is to recover the deformation. Important matching problems include stereo vision, motion analysis, and a variety of registration problems such as template matching for speech and vision.

We cast the problem as the minimization of an energy functional $\mathscr{E}(\mathbf{V})$ where $\mathbf{V}$ is the deformation. The energy functional is the sum of two terms, one based on the correlation of the deformed signals, and the other based on the smoothness of the deformation.

In general, the energy functional $\mathscr{E}(\mathbf{V})$ can be highly nonconvex, so that ordinary optimization methods become trapped in local minima. Optimization by simulated annealing can be attempted, but at severe computational expense. Instead, we rely on *continuation methods* to solve the problem. By introducing a scale parameter σ, the minimization problem is embedded within a larger space. A suitable minimum can be achieved relatively easily for large σ because the signals and hence the energy landscape are very smooth. The solution of the original minimization problem is then obtained by continuously tracking the minimum as σ tends to zero. This is analogous to a coarse-to-fine tracking of extrema through scale space in the sense of Witkin [1].

The entire procedure consists of solving the first-order dynamic system

$$\dot{\sigma} = -c_1 \exp(-c_2 |\nabla \mathscr{E}(\mathbf{V}, \sigma)|), \dot{\mathbf{V}} = -\nabla \mathscr{E}(\mathbf{V}, \sigma)$$

where the dot denotes a time derivative, σ is the scale parameter and c_1 and c_2 are constants. Given an initial crude estimate for $\mathbf{V}$ at a coarse scale σ_0, the system minimizes $\mathscr{E}$ at σ_0 and follows a trajectory of minima through finer scales, thereby increasing the resolution of $\mathbf{V}$. Any of a number of well-known numerical techniques can be used to solve for the trajectory. Through a series of incremental deformations, correlations of deformed signals are optimized and balanced against the smoothness of the deformations while moving from coarse–to–fine scale. Thus, the first-order system compactly unifies a number of important yet seemingly disparate signal-matching notions.

In the remainder of this section, the relation of our technique to previous work on matching is discussed. Then in section 2, a minimization

framework for the matching problem is introduced. In section 3, the solution of the problem by continuation methods and the resulting single differential equation are developed. Section 4 describes the specific similarity term employed and, in section 5, the details of the smoothness term are discussed. Section 6 develops a discrete solution of the differential equation. Finally, section 7 presents several examples of matching results for one- and two-dimensional signals.

1.1 Background

An enormous amount of work has been done on signal matching, giving precedent for several components of our approach.

Optimization of constrained deformations guided by correlation or L_2 metrics can be found in prior work. In speech recognition, the problem of time warping speech segments to match input utterances with stored prototypes has been addressed in this context. Dynamic programming has been used to compute constrained warping functions (see Rabiner and Schafer [2] and Sankoff and Kruskal [3], part II). This particular optimization technique is readily applicable in matching situations involving sequentially ordered signals, such as speech, and unilateral continuity constraints. However, its stringent requirements on the energy functional appear incompatible with the unordered multidimensional signals and isotropic smoothness constraints that are of primary concern to us.

Smoothness constraints have been popular in computational vision. Consider the important problem of stereo matching. In the past, dense disparity maps have been computed through a two-step process of local matching followed by smooth [4], multiresolution [5], or piecewise continuous [6, 7] surface reconstruction from the sparse disparities. The approach in the present report unifies matching and piecewise smooth reconstruction into a single iterative optimization process.

Broit's work [8] in registering a deformed image to a model image resembles ours in that matching is explicitly formulated as a minimization problem involving a cost functional that combines both a deformation constraint and a similarity measure. His deformation model, which involves the strain energy of a globally smooth elastic body, is more elaborate than the deformation constraints inherent in the spring-loaded subtemplate matching technique of Fischler and Elschlager [9] or the iterative Gaussian-smoothed deformation models proposed by Burr [10]. Our controlled-continuity deformation model provides us with the additional capability to regulate the order of smoothness and to preserve discontinuities in the deformation.

Horn ([11] section 13.9.1) formulates an approach to stereo matching that involves the minimization of a cost functional that combines a global smoothness constraint and a squared difference image similarity measure. Poggio et al. ([12] table 1) consider a similar cost functional for stereo in the context of the regularization approach. Neither of these sources presents an implementation, but mention is made of low-pass or band-pass filtering of images as a means of simplifying the minimization. Stochastic optimization is the usually expensive alternative. Monte Carlo algorithms for signal matching are given by Marroquin ([13] chapter 6) and a simulated annealing algorithm for stereo matching is offered by Barnard [14]; both employ cost functionals very similar to Horn and Poggio.

Coarse-to-fine matching schemes have previously been treated as a multistage process in which a matching operation is performed at each successive level [15–19]. We have extended this idea into a matching process that evolves continuously toward finer spatial scale. The idea of progressing continuously through scale space derives from Witkin [1].

As our matching process computes the deformation iteratively, it is best to perform the similarity measurements by deforming the signals according to the current approximation of the deformation. This concern has also been addressed by the matching algorithms described by Mori et al. [15] Burr [10], Broit [8], and Quam [20].

2 Framework

Consider a vector of n similar signals $\mathbf{f}^*(\mathbf{x}) = [f_1^*(\mathbf{x}), \ldots, f_n^*(\mathbf{x})]$ defined in d dimensional space $\mathbf{x} = [x_1, \ldots, x_d] \in \mathcal{R}^d$, and a deformation mapping $\mathbf{V}: \mathcal{R}^d \rightarrow \mathcal{R}^{nd}$, such that $\mathbf{V}(\mathbf{x}) = [\mathbf{v}_1(\mathbf{x}), \ldots,$

$\mathbf{v}_n(\mathbf{x})]$, where each of the n disparity functions $\mathbf{v}_k$: $\mathscr{R}^d \rightarrow \mathscr{R}^d$ is a vector valued function $\mathbf{v}_k(\mathbf{x}) = [v_1(\mathbf{x}), \ldots, v_d(\mathbf{x})]^T$. Given a set of deformed signals $\mathbf{f}$ such that $\mathbf{f}^*(\mathbf{x}) = \mathbf{f}(\mathbf{V}^*(\mathbf{x}))$, the matching problem is to recover the deformation $\mathbf{V}^*(\mathbf{x})$.

Suppose that the similarity between the signals $\mathbf{f}$ for a given deformation $\mathbf{V}$ is measured by a functional $\mathscr{Q}(\mathbf{V})$: $\mathscr{R}^{nd} \rightarrow \mathscr{R}$ bounded from above by a value achieved by the best possible match. A reasonable objective is to find the deformation $\mathbf{U}$ which maximizes the quality of the match; i.e., to minimize $-\mathscr{Q}(\mathbf{V})$ over possible deformations $\mathbf{V}$. Thus, $\mathbf{U}$ represents an optimal approximation to $\mathbf{V}^*$.

This minimization problem is clearly ill-posed in the absence of constraints on admissible deformations, since, e.g., degenerate or chaotic deformations can always be contrived that achieve the minimum value. Such constraints may be encoded by a second functional $\mathscr{S}(\mathbf{V})$: $\mathscr{H}^{nd} \rightarrow \mathscr{R}$, where $\mathscr{H}^{nd} \subset \mathscr{R}^{nd}$ is the subset of admissible deformations.

Useful instances of similarity and constraint functionals will be formulated shortly. Their combination, however, leads to the following minimization problem: Find the deformation $\mathbf{U} \in \mathscr{H}^{nd}$ such that $\mathscr{E}(\mathbf{U}) = \inf_{\mathbf{V} \in \mathscr{H}^{nd}} \mathscr{E}(\mathbf{V})$, where the energy functional is given by

$$\mathscr{E} = -(1 - \lambda)\mathscr{Q} - \lambda \mathscr{S} \tag{1}$$

and where $\lambda \in (0, 1)$ is a weighting parameter.

Stabilization offers a general approach to a numerical solution through the construction of a discrete dynamic system whose fixed points include a discrete solution of the above optimization problem [21]. A simple dynamic system with this property is characterized by the differential equation

$$\dot{\mathbf{V}} + \nabla \mathscr{E} = 0 \tag{2}$$

where the dot denotes differentiation with respect to time t and $\nabla\varepsilon$ denotes the gradient of ε with respect to the free variables of the discrete deformation. Optimization occurs by dissipation of energy; energy cannot increase along the system's trajectory $\mathbf{V}(\mathbf{x}, t)$ in $\mathscr{H}^{nd}$, which follows the direction of the gradient of $\mathscr{E}$. Although the trajectory terminates at a local minimum of $\mathscr{E}$, there is no guarantee that the global minimum $\mathbf{U}$ will be attained by solving this initial value problem starting from an arbitrary initial condition $\mathbf{V}(\mathbf{x}, 0)$.

3 Continuation over Scale

The key remaining difficulty is that for obvious choices of $\mathscr{Q}$, such as linear correlation, $\mathscr{E}$ is likely to have many local minima, making the minimization problem highly nonconvex and therefore extremely difficult to solve. There are two options: solving this hard problem directly (for example by simulated annealing) or simplifying the problem by choosing $\mathscr{Q}$ to be convex or nearly so. We pursue the second option because annealing is expensive.

3.1 Continuation Methods

$\mathscr{Q}$ may be smoothed by subjecting $\mathbf{f}$ to a smoothing filter of characteristic width σ. We observe empirically that the best solution for $\mathbf{V}$ as σ increases tends to be an increasingly smoothed version of the correct solution. This means that slightly deblurring $\mathbf{V}$ by reducing σ produces a slightly better solution close to the one just obtained. To the extent this is so, we can solve the problem using equation (2) by means of *continuation methods* [22].

Continuation methods embed the problem to be solved,

$$\mathbf{g}(\mathbf{v}) = 0$$

in a family of problems

$$\mathbf{g}(\mathbf{v}, s) = 0$$

parameterized by s. Let $s_{i+1} = s_i + \Delta s$, $\mathbf{g}(\mathbf{v}, s_n)$ be the problem we wish to solve, presumably difficult, and $\mathbf{g}(\mathbf{v}, s_1)$ a readily solvable member of the family, and let

$$\mathbf{u}(s) = \mathbf{H}(\mathbf{g}, s, \mathbf{v}_0)$$

be the solution for $\mathbf{g}$ at s given $\mathbf{v}_0$ as an initial condition. Then $\mathbf{u}(s_n)$ is obtained from $\mathbf{u}(s_1)$ by the iteration

$$\mathbf{u}(s_{i+1}) = \mathbf{H}(\mathbf{g}, s_{i+1}, \mathbf{u}_i); \quad i = 1, \ldots, n - 1$$

that is, each solution is used as an initial condition to obtain the next one.

For the current problem, the continuation

parameter is σ, with $\Delta\sigma < 0$. We continue from an initial coarse scale σ_1 and an initial guess $\mathbf{V}_1$ by

$$\mathbf{V}_{i+1} = \mathbf{H}(\dot{\mathbf{V}} + \Delta\, \mathscr{E}, \sigma_{i+1}, \mathbf{V}_i)$$

to a fine scale σ_n and a final answer $\mathbf{V}_n$. To visualize this method, imagine the energy landscape at each value of σ as a contoured surface in 3-space. The surfaces are stacked one above the other, so that the topmost surface is very smooth, while the lower ones become increasingly bumpy. Imagine a hole drilled at each local minimum on each surface. A ball bearing dropped onto the topmost surface will roll down to the bottom of the hill. At this point, it falls through to the next level, rolls down again, falls through again, and so on to the bottom.

3.2 A Scale Space Equation

This iteration solves a separate initial value problem at each step. A more attractive alternative is to collapse the continuation over σ into a single differential equation. Ideally, the solution should follow a curve $\mathbf{V}(\sigma)$ satisfying $|\nabla\mathscr{E}(\mathbf{V}(\sigma))| \equiv 0$; i.e., a continuous curve of solutions over scale. A differential equation for this curve is

$$\mathbf{V}_\sigma = -(\nabla\mathscr{E})_\sigma(\nabla\nabla\mathscr{E})^{-1} \qquad (3)$$

The solution to this equation tracks a given coarse-scale solution continuously to fine-scale, in precise analogy with the coarse-to-fine tracking through scale space of Witkin [1]. Unfortunately, it is impractical to solve this equation for arbitrary $\mathscr{S}$ and $\mathscr{Q}$, since $\nabla\nabla\mathscr{E}$ is high dimensional.

To construct an approximate equation, we introduce the quantity $\Phi = -c_1 e^{-c_2|\nabla\mathscr{E}|}$, so that $\Phi = -c_1$ at a solution to equation (2), diminishing with distance from the solution at a rate determined by the space constant c_2. The equation

$$\dot{\sigma} = -c_1 e^{-c_2|\nabla\mathscr{E}(\mathbf{V},\sigma)|}, \quad \dot{\mathbf{V}} = -\nabla\mathscr{E}(\mathbf{V}, \sigma) \qquad (4)$$

approaches the desired behavior. Far from a solution, where Φ is small, question (4) approaches equation (2), changing $\mathbf{V}$ but not σ. Approaching a solution, σ begins to decrease. At a solution, $\dot{\mathbf{V}} = 0$ and $\dot{\sigma} = -c_1$. From an initial $\mathbf{V}(t_0)$, $\sigma(t_0)$, the solution $\mathbf{V}(t)$, $\sigma(t)$ moves through $\mathbf{V}$ at nearly constant scale until a minimum in $\mathscr{E}$ is approached, then it begins descending in scale staying close to a solution.[1]

Equation (4) finds a solution at the initial scale, then tracks it continuously to finer scales. To use equation (4), we choose a coarse-scale $\sigma(t_0)$, a crude initial guess $\mathbf{V}(t_0)$, and a terminal fine-scale σ_T. We then run the equation until $\sigma(t) = \sigma_T$, taking $\mathbf{V}(T)$ as the solution.

3.3 Ambiguous Solutions

From time to time, we expect to encounter instabilities in the solutions of equation (4), in the sense that a small perturbation of the data induces a large change in the solution curve's trajectory through scale space. These instabilities correspond to bifurcations of the solution curve, analogous to bifurcations that can be observed in Guassian scale space. We have considered two approaches to dealing with them. First, by adding a suitable noise term to $\mathscr{E}$, equation (4) becomes a hybrid of scale space continuation and simulated annealing. We believe that local ambiguities can be favorably resolved using low-amplitude noise, and hence with little additional computational cost. A second approach is to regard these instabilities as genuine ambiguities whose resolution falls outside the scope of the method. In that case, a set of alternative solutions can be explored by the addition of externally controlled bias terms to $\mathscr{E}$. These terms can reflect outside constraints of any kind, for example, those imposed by the operation of attentional processes.

In the following sections, we turn to specific choices for $\mathscr{S}$ and $\mathscr{Q}$.

[1] The solution to equation (4) oscillates around the exact solution (equation 3) with frequency and amplitude controlled by c_1 and c_2. This oscillation can be damped by the addition of second-order terms in t, but we have not found it necessary to do so in practice.

4 Similarity Functional

In general the similarity measure $\mathcal{Q}$ should capture what is known about the specific matching problem. In many cases, the undeformed signals are sufficiently similar that a simple correlation measure suffices. In this section we formulate a generic choice for this class of problems. Note that, by assumption, it is the undeformed signals $\mathbf{f}^*(\mathbf{x})$ that are similar, so the quality of a potential solution $\mathbf{V}$ should be measured by the similarity of the signals $\mathbf{f}(\mathbf{V}(\mathbf{x}))$.

Consider the case of two signals f_k and f_l. A general family of similarity measures is obtained by integrating local similarity measures over position. Let $Q_{k,l}(\mathbf{x})$ be a local measure of the similarity of $f_k(\mathbf{v}_k(\mathbf{x}))$ and $f_l(\mathbf{v}_l(\mathbf{x}))$ around $\mathbf{x}$. A number of possibilities exist for $Q_{k,l}(\mathbf{x})$. Normalized cross-correlation produces good results for several matching problems that we have examined. If $W\gamma(\mathbf{x})$ is a window function where γ denotes the width parameter, $\mu_k(\mathbf{x}) = \int f_k(\mathbf{v}_k(\mathbf{y} - \mathbf{x}))\, W\gamma(\mathbf{y})d\mathbf{y}$, and $\nu_k(\mathbf{x}) = \int [f_k(\mathbf{v}_k(\mathbf{x} - \mathbf{y})) - \mu_k(\mathbf{x} - \mathbf{y})]^2\, W\gamma(\mathbf{y})d\mathbf{y}$, then the normalized cross-correlation can be written

$$Q_{k,l}(\mathbf{x}) = [\nu_k(\mathbf{x})\nu_l(\mathbf{x})]^{-\frac{1}{2}} \int \{[f_k(\mathbf{v}_k(\mathbf{x} - \mathbf{y})) - \mu_k(\mathbf{x} - \mathbf{y})] \times [f_l(\mathbf{v}_l(\mathbf{x} - \mathbf{y})) - \mu_l(\mathbf{x} - \mathbf{y})] W\gamma(\mathbf{y})\} d\mathbf{y}$$

A global measure of similarity for f_k and f_l is given by

$$\mathcal{Q}_{k,l}(\mathbf{v}_k, \mathbf{v}_l) = \int Q_{k,l}(\mathbf{x})d\mathbf{x}$$

This functional generally has many local minima. In order to apply the continuation method, we compute $\mathcal{Q}_{k,l}(\mathbf{v}_k, \mathbf{v}_l)$ for signals f_k and f_l, which have been smoothed by Gaussians of standard deviation σ. The correlation window size W_γ should be large enough to provide an accurate local estimate of the mean and variance of the signals, but small enough that nonstationarities in the signals do not become a problem. A convenient way to set γ to a reasonable value is to make it a fixed multiple of the average autocorrelation widths of the smoothed signals. Then γ can be regarded as a function of σ. The resulting functional $\mathcal{Q}_{k,l}(\mathbf{v}_k, \mathbf{v}_l, \sigma)$ can then be made as smooth as is desired.

By simply adding up pairwise similarities, a global measure of similarity can be constructed for n signals:

$$\mathcal{Q}(\mathbf{V},\sigma) = \sum_{k=1}^{n} \sum_{l=k+1}^{n} \mathcal{Q}_{k,l}(\mathbf{v}_k, \mathbf{v}_l, \sigma)$$

Note that $\mathcal{Q}(\mathbf{V},\sigma)$ must be recomputed at each iteration, with the signals resampled to reflect the current choice of $\mathbf{V}$. If the deformation is very small, the distortion induced by failing to resample can be ignored, but the value of such resampling in stereo matching, for example, is well established [15, 20]. The Gaussian smoothing should also take place on the resampled signals $\mathbf{f}(\mathbf{V}(\mathbf{x}))$.

5 Smoothness Functionals

The functional $\mathcal{S}(\mathbf{V})$ places certain restrictions on admissible deformations in order to render the minimization problem better behaved. Perhaps the simplest possible restriction, and one that has been used often in the past, is to limit possible disparities between signals to prespecified ranges. A deformation can then be assigned within disparity bounds on a point-by-point basis according to maximal similarity criteria. Although simple, such limited searches are unfortunately error prone, since they are based on purely local information.

This problem can be resolved by imposing global constraints on the deformation that are more restrictive yet remain generic. Such constraints may be based on a priori expectations about deformations; for example, that they are coherent in some sense. In particular, admissible deformations may be characterized according to the *controlled-continuity constraints* defined by Terzopoulos [7]. These constraints, which are based on generalized splines, restrict the admissible space of deformations to a class of piecewise continuous functions. The deformation's order of continuity is controllable, and discontinuities of various orders (e.g., position, tangent, and curvature discontinuities) are permitted to occur, subject to an energy penalty.

A general controlled-continuity constraint is imposed on the deformation by the functional

$$\mathcal{S}(\mathbf{V}) = \sum_{i=1}^{n} S(\mathbf{v}_k)$$

where

$$S(\mathbf{v}) = -\sum_{m=0}^{p} \int_{\mathcal{R}^d} w_m(\mathbf{x}) \sum_{j_1+\ldots+j_d=m} \frac{m!}{j_1!\ldots j_d!} \left| \frac{\partial^m \mathbf{v}(\mathbf{x})}{\partial x_1^{j_1} \ldots \partial x_d^{j_d}} \right|^2 d\mathbf{x} \tag{5}$$

The positive integer p indicates the highest order generalized spline that occurs in the functional, and this determines the maximum order of continuity (C^{p-1}) of the admissible deformations. The nonnegative continuity control functions $\mathbf{w}(\mathbf{x}) = [w_0(\mathbf{x}), \ldots, w_p(\mathbf{x})]$ determine the placement of discontinuities. A discontinuity of order $q < p$ is permitted to occur at $\mathbf{x}_0$ by forcing $w_i(\mathbf{x}_0) = 0$ for $i \geq q$ (see Terzopoulos [7] for details).

The $p = 2$ order controlled-continuity constraint is employed in our implementations to date. If, for convenience, a "rigidity" function $\rho(\mathbf{x})$ and a "tension" function $[1 - \tau(\mathbf{x})]$ are introduced such that $w_0(\mathbf{x}) = 0$, $w_1(\mathbf{x}) = \rho(\mathbf{x})[1 - \tau(\mathbf{x})]$, and $w_2(\mathbf{x}) = \rho(\mathbf{x})\tau(\mathbf{x})$, then it is natural to view the functionals as characterizing "generalized piecewise continuous splines under tension." In particular, for the case of n signals in 1 dimension $\mathbf{x} = [x]$, the functional (5) reduces to

$$S(\mathbf{v}) = -\int_{\mathcal{R}} \rho(x)\{[1 - \tau(x)]|\mathbf{v}_x|^2 + \tau(x)|\mathbf{v}_{xx}|^2\} dx \tag{6}$$

while for the case of n signals in two dimensions $\mathbf{x} = [x, y]$, it becomes

$$S(\mathbf{v}) = -\int\int_{\mathcal{R}^2} \rho(x, y)\{[1 - \tau(x, y)](|\mathbf{v}_x|^2 + |\mathbf{v}_y|^2) + \tau(x, y)(|\mathbf{v}_{xx}|^2 + 2|\mathbf{v}_{xy}|^2 + |\mathbf{v}_{yy}|^2)\} dx\, dy \tag{7}$$

6 Discrete Solution

Equations (4) pose a standard first-order initial value problem for which solution methods abound. We have employed numerical methods of varying sophistication, each giving satisfactory results. In order of sophistication, these include Euler's method, a fourth-order Runge-Kutta method, and Adams-Moulton predictor-corrector methods. The latter offer the advantage that the step size can be automatically adapted, making them particularly robust [22].

Discretization of the continuous variational form of the n signal-matching problem can be carried out using standard methods. As an illustration of the iterative formulas involved, consider an explicit Euler step in which the time derivatives $\dot{\sigma}$ and $\dot{\mathbf{V}}$ are approximated by forward difference expressions with time step Δt. This leads to the following discrete time version of equations (4):

$$\sigma^{t+1} = \sigma^t - (\Delta t)c_1 e^{-c_2|\nabla \mathcal{E}^t|},$$
$$\mathbf{V}^{t+1} = \mathbf{V}^t - (\Delta t)\nabla \mathcal{E}^t$$

where $\nabla \mathcal{E}^t = [\nabla_1 \mathcal{E}^t, \ldots, \nabla_n \mathcal{E}^t]$, whose components, one for each signal, are given by

$$\nabla_k \mathcal{E}(\mathbf{V}, \sigma)^t = -(1 - \lambda) \sum_{\substack{l=1 \\ l \neq k}}^{n} \nabla_k \mathcal{Q}_{k,l}(\mathbf{v}_k, \mathbf{v}_l, \sigma)^t - \lambda \nabla S(\mathbf{v}_k)^t \tag{8}$$

Although finite element methods offer the greatest flexibility, for simplicity we employ standard central finite difference formulas for uniform meshes to approximate the spatial derivatives. This leads to locally computable expressions for equation (8).

The first term in equation (8) involves partial derivatives of the similarity functional with respect to each of the d disparity functions in $\mathbf{v}_k$; i.e., $\nabla_k \mathcal{Q}_{k,l}(\mathbf{v}_k, \mathbf{v}_l, \sigma) = [\partial \mathcal{Q}_{k,l}/\partial v_1, \ldots, \partial \mathcal{Q}_{k,l}/\partial v_d]$. These derivatives are computed numerically by taking first differences.

The components of $\nabla S(\mathbf{v})$, the second term in equation (8), are obtained by discretizing S using first and second differences. For the case of matching one-dimensional signals, discretization of equation (6) yields the components

$$\frac{\partial S(\mathbf{v})}{\partial \mathbf{v}_i} = \frac{1}{h^2}\{w_{1\,i}(\mathbf{v}_{i+1} - \mathbf{v}_i) - w_{1\,i-1}(\mathbf{v}_i - \mathbf{v}_{i-1})\} - \frac{1}{h^4}\{w_{2\,i-1}(\mathbf{v}_i - 2\mathbf{v}_{i-1} + \mathbf{v}_{i-2}) - 2w_{2\,i}(\mathbf{v}_{i+1} - 2\mathbf{v}_i + \mathbf{v}_{i-1}) + w_{2\,i+1}(\mathbf{v}_{i+2} - 2\mathbf{v}_{i+1} + \mathbf{v}_i)\} \tag{9}$$

where i indexes the nodes of the mesh defined on a compact domain of interest $\Omega \in \mathcal{R}$, $w_{1\,i} = \rho_i[1 - \tau_i]$, and $w_{2\,i} = \rho_i\tau_i$. Similarly, for the case of matching two-dimensional signals, the discretization of equation (7) yields the components

$$\frac{\partial S(\mathbf{v})}{\partial \mathbf{v}_{i,j}} = \frac{1}{h^2}\{w_{1\,i,j}(\mathbf{v}_{i+1,j} - \mathbf{v}_{i,j})$$
$$- w_{1\,i-1,j}(\mathbf{v}_{i,j} - \mathbf{v}_{i-1,j})$$
$$+ w_{1\,i,j}(\mathbf{v}_{i,j+1} - \mathbf{v}_{i,j}) - w_{1\,i,j-1}(\mathbf{v}_{i,j} - \mathbf{v}_{i,j-1})\}$$
$$- \frac{1}{h^4}\{w_{2\,i-1,j}(\mathbf{v}_{i,j} - 2\mathbf{v}_{i-1,j} + \mathbf{v}_{i-2,j})$$
$$- 2w_{2\,i,j}(\mathbf{v}_{i+1,j} - 2\mathbf{v}_{i,j} + \mathbf{v}_{i-1,j})$$
$$+ w_{2\,i+1,j}(\mathbf{v}_{i+2,j} - 2\mathbf{v}_{i+1,j} + \mathbf{v}_{i,j})$$
$$+ 2w_{2\,i-1,j-1}(\mathbf{v}_{i,j} - \mathbf{v}_{i-1,j} - \mathbf{v}_{i,j-1} + \mathbf{v}_{i-1,j-1})$$
$$- 2w_{2\,i,j-1}(\mathbf{v}_{i+1,j} - \mathbf{v}_{i,j} - \mathbf{v}_{i+1,j-1} + \mathbf{v}_{i,j-1})$$
$$- 2w_{2\,i-1,j}(\mathbf{v}_{i,j+1} - \mathbf{v}_{i-1,j+1} - \mathbf{v}_{i,j} + \mathbf{v}_{i-1,j})$$
$$+ 2w_{2\,i,j}(\mathbf{v}_{i+1,j+1} - \mathbf{v}_{i,j+1} - \mathbf{v}_{i+1,j} + \mathbf{v}_{i,j})$$
$$+ w_{2\,i,j-1}(\mathbf{v}_{i,j} - 2\mathbf{v}_{i,j-1} + \mathbf{v}_{i,j-2})$$
$$- 2w_{2\,i,j}(\mathbf{v}_{i,j+1} - 2\mathbf{v}_{i,j} + \mathbf{v}_{i,j-1})$$
$$+ w_{2\,i,j+1}(\mathbf{v}_{i,j+2} - 2\mathbf{v}_{i,j+1} + \mathbf{v}_{i,j})\} \qquad (10)$$

where i, j index the nodes of the mesh defined on the compact domain $\Omega \in \mathscr{R}^2$, $w_{1\,i,j} = \rho_{i,j}[1 - \tau_{i,j}]$, and $w_{2\,i,j} = \rho_{i,j}\tau_{i,j}$. In equations (9) and (10), h denotes the distance between neighboring nodes.

The above local computations are valid only in regions where the deformation has C^1 continuity. The computations must be modified at nodes that are within a distance of $2h$ from the position discontinuities on the boundary of Ω (the natural boundary conditions). The same is true within Ω regarding position and tangent discontinuities in the deformation, which have been prescribed or which may be introduced during the matching process. The iterative computations are automatically modified as required using the technique of summation and inhibition of computational molecules (see [6]). The automatic detection of discontinuities requires the estimation w_1 and w_2 during matching, and this is accomplished using techniques described by Terzopoulos [7].

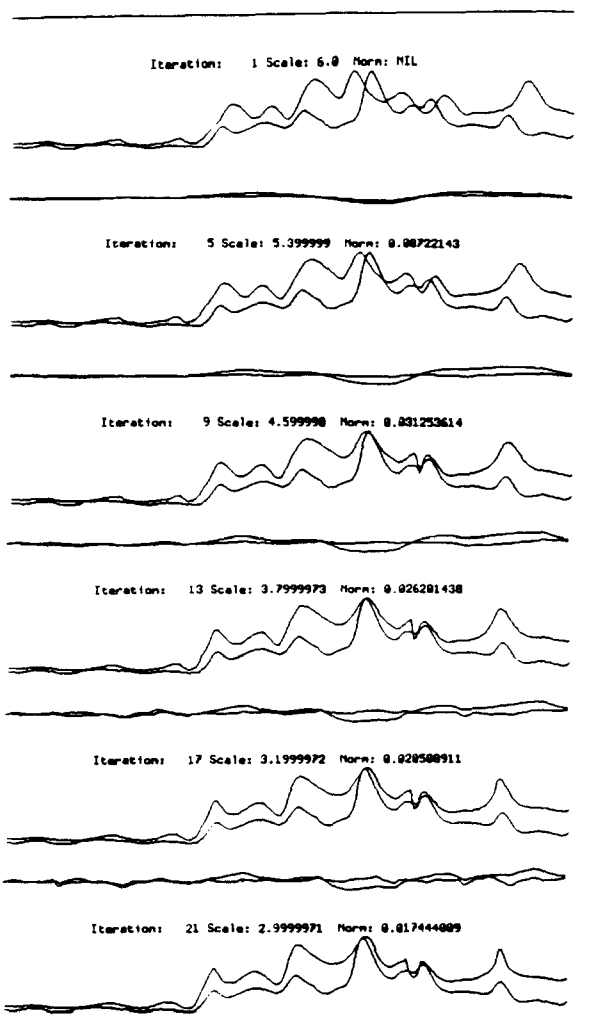


Fig. 1. Matching two one-dimensional signals. The deformation function **V** is shown above each signal trace. From *top to bottom*, the signals first appear in their original form, then partially deformed at intermediate stages of the matching process, and finally showing the end result.

7 Results

7.1 One-Dimensional Signals

The method is applicable to matching n signals, each of which is d-dimensional. Figure 1 shows the simplest case, that of matching two one-dimensional signals. The signals are measurements of the resistivity of a geologic structure as a function of depth at two different locations. From top to bottom, the signals first appear in their original form, then partially deformed at intermediate stages of the matching process, finally showing the end result. The deformation function **V** is shown above the signals.

Figure 2 shows a more challenging example in which four signals are matched simultaneously. The signals are intensity profiles from a complex natural image. On the left, the four signals are shown superimposed at several points in the matching process. As before, the original signals appear at the top and the final result is at the bottom. Note that coarse-scale features are aligned

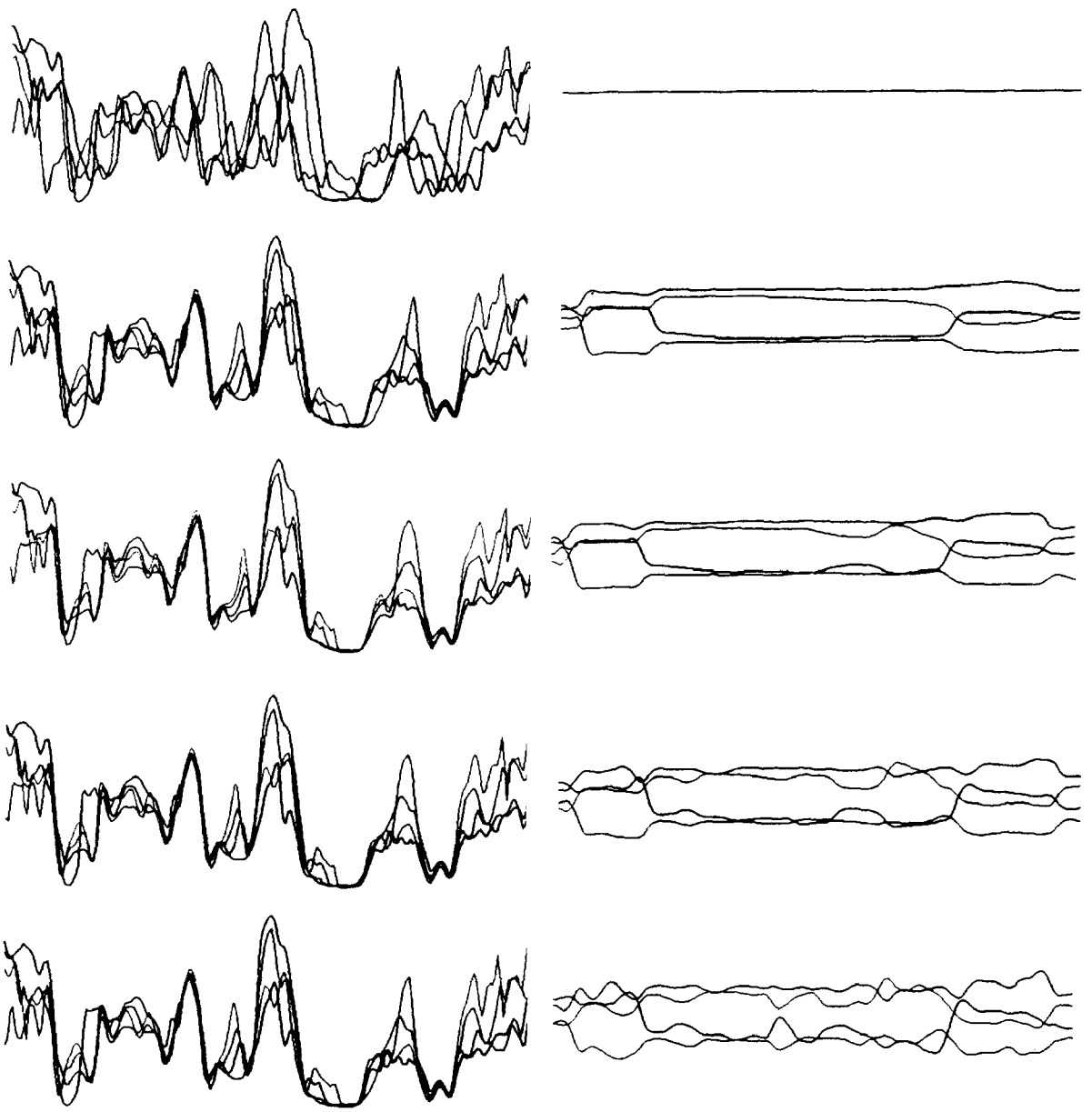


Fig. 2. Simultaneous matching of four signals. The signals are superimposed at several points in the matching process (*left*). The original signals appear at the *top* and the final result is at the *bottom*. The four corresponding deformation functions $v_i(x)$ are shown (*right*).

(a)

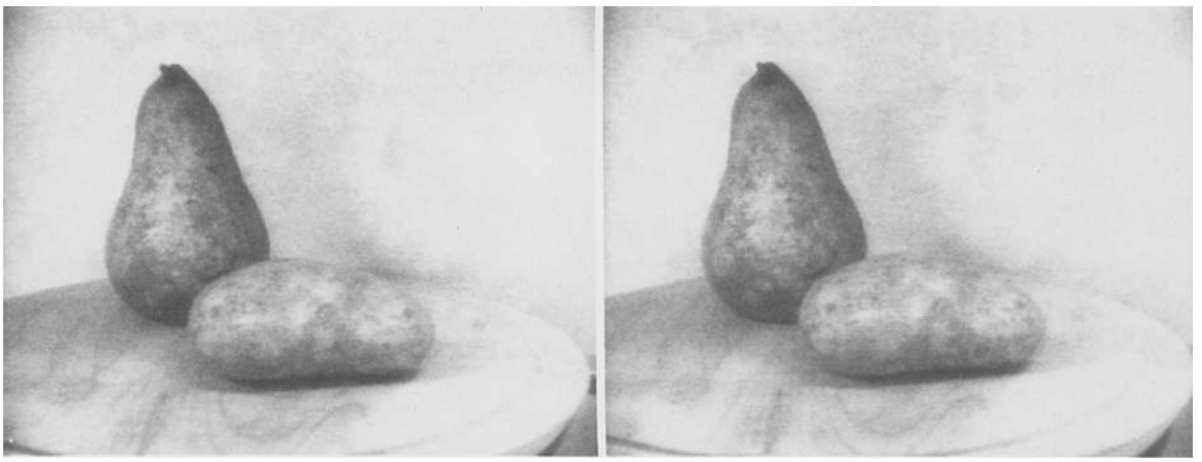
(b)

Fig. 3. Two frames from a motion sequence (**a**). The original image has been texture mapped onto a surface that encodes speed as elevation (**b**). The raised area indicates motion before a stationary background.

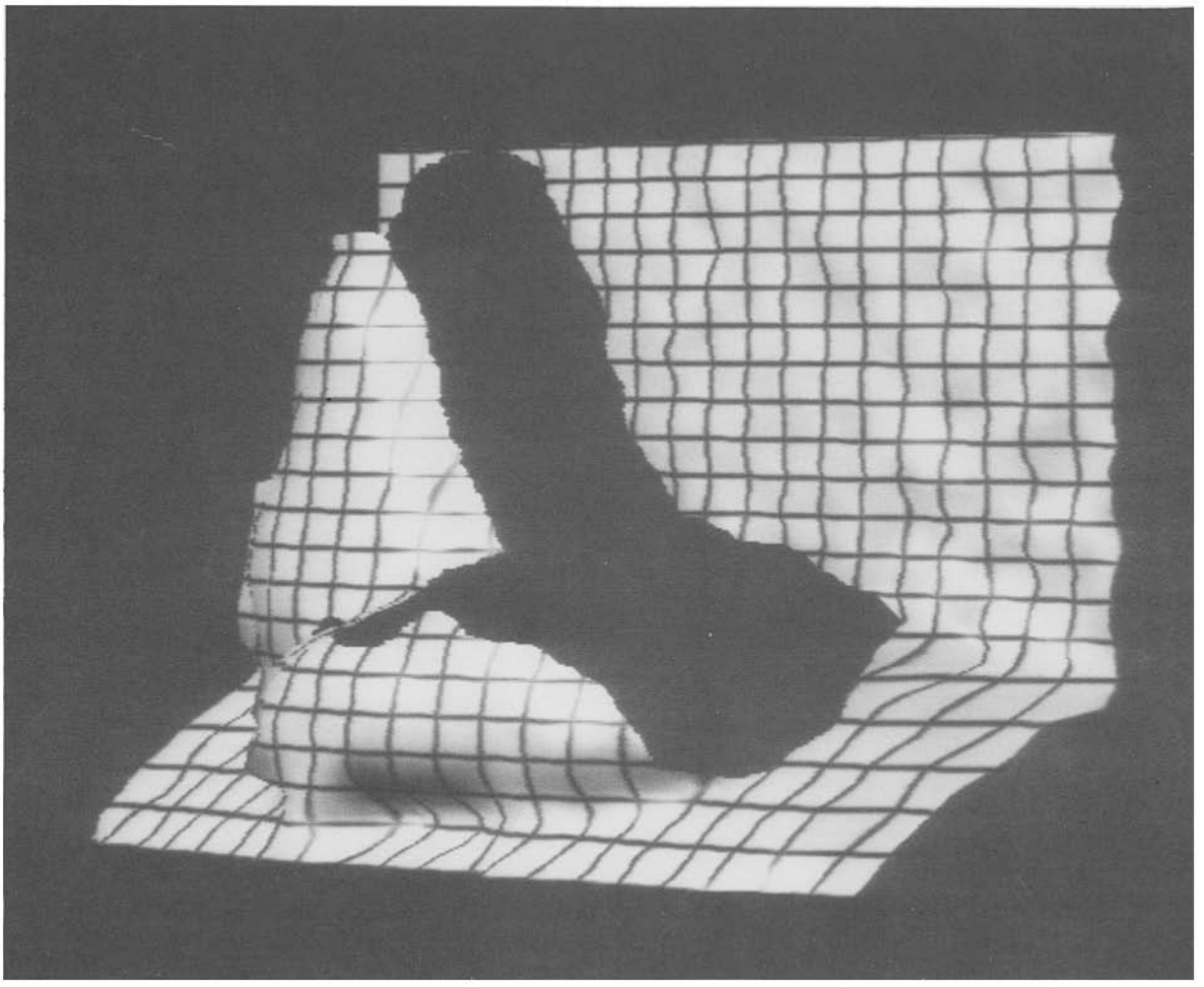
first in the matching process while fine-scale features are matched later. The four corresponding deformation functions $v_i(x)$ are shown to the right.

7.2 Motion Sequence

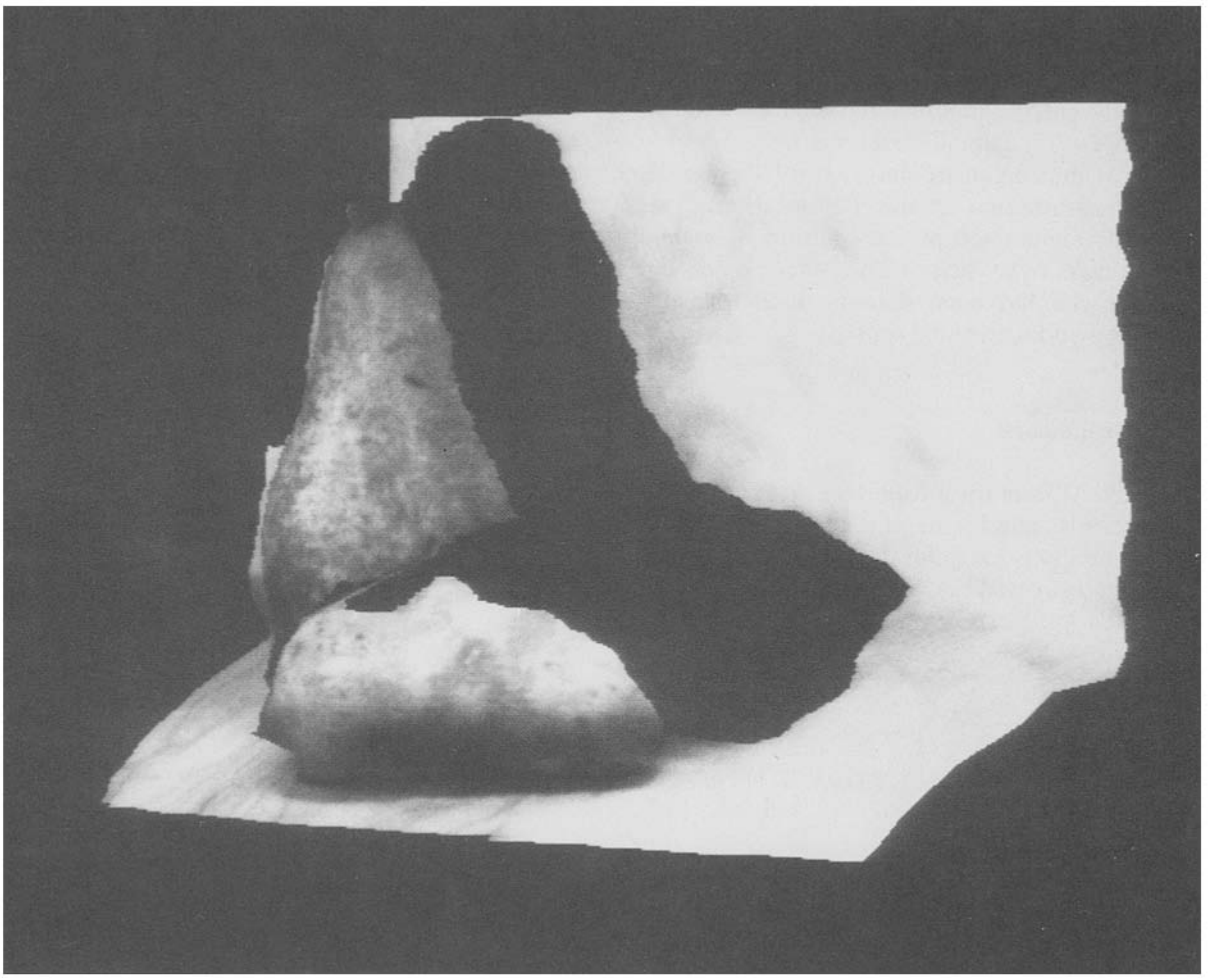
Figure 3 shows two frames from a motion sequence showing M. Kass moving against a stationary background. The frames are separated in time by about 1.5 s. Results of the matching process are shown as follows: The original image has been mapped onto a surface that encodes estimated speed as elevation. The raised area shows the region in which the algorithm detects motion. No attempt has been made in this example to detect motion discontinuities.



(a)



(b)



(c)

Fig. 4. A stereogram showing a potato partly occluding a pear (the images are reversed for free fusing) (**a**). The matching results are rendered as two shaded surfaces with the image coordinate grid mapped onto the first (**b**) and the left image mapped onto the second (**c**). The visible portions for the reconstructed surfaces are rendered from an oblique viewpoint showing the computed surface discontinuities.

7.3 Stereo Matching

Figure 4 contains a stereogram showing a potato partly occluding a pear. The matching results are rendered as two shaded surfaces with depth computed from the disparity. An image coordinate grid is mapped onto the first surface and the left image is mapped onto the second. The reconstructed surfaces are rendered from an oblique viewpoint showing the computed surface discontinuities. Only those portions of the scene visible in the original stereogram are shown.

8 Conclusion

A variational approach to signal matching has been developed. The approach unifies several established themes in signal matching including constrained optimization, piecewise continuous reconstruction, coarse-to-fine matching, and incremental deformation of signals. The main technical contribution of this paper is twofold. First, we introduced the notion of tracking the solution to the matching problem continuously over scale. This continuation method for finding good local

minima offers an efficient deterministic alternative to stochastic optimization techniques. Second, we developed a single system of first-order differential equations that characterize this process. The system is governed by an energy functional that balances similarity of the signals against smoothness of the deformation. This balance is maintained as the solution is tracked continuously over scale. The effectiveness of this approach has been demonstrated for both one- and two-dimensional signals.

Acknowledgments

We thank Al Barr for introducing us to continuation methods, and for helping us with numerical solution methods for differential equations. Keith Nishihara provided us with stereo correlation data.

References

1. A. Witkin, "Scale Space Filtering," in PROC. INT. JOINT CONF. ARTIF. INTELL., Karlsruhe, FRG, 1983, pp. 1019–1021.
2. L.R. Rabiner and R.W. Schafer, DIGITAL PROCESSING OF SPEECH SIGNALS. Prentice-Hall: Englewood Cliffs, NJ, 1978.
3. D. Sankoff and J.B. Kruskal (eds.), TIME WARPS, STRING EDITS AND MACROMOLECULES: THE THEORY AND PRACTICE OF SEQUENCE COMPARISON. Addison-Wesley: Reading, MA, 1983.
4. W.E.L. Grimson, "An implementation of a computational theory of visual surface interpolation," COMPUT. VISION, GRAPHICS, IMAGE PROCESSING vol. 22, pp. 39–69, 1983.
5. D. Terzopoulos, "Multilevel computational processes for visual surface reconstruction," COMPUT. VISION, GRAPHICS, IMAGE PROCESSING vol. 24, pp. 52–96, 1983.
6. D. Terzopoulos, "The role of constraints and discontinuities in visible-surface reconstruction," in PROC. 8TH INT. JOINT CONF. ARTIF. INTELL., Karlsruhe, FRG, 1983, pp. 1073–1077.
7. D. Terzopoulos, "Regularization of inverse visual problems involving discontinuities," IEEE TRANS. PAMI vol. PAMI-8, pp.413–424, 1986.
8. C. Broit, "Optimal registration of deformed images," PhD thesis, Computer and Information Science Department, University of Pennsylvania, Philadelphia, PA, 1981.
9. M.A. Fischler, and R.A. Elschlager, "The representation and matching of pictorial structures," IEEE TRANS. COMPUT. vol. C-22, pp. 67–92, 1973.
10. D.J. Burr, "A dynamic model for image registration," COMPUT. GRAPHICS IMAGE PROCESSING vol. 15, pp. 102–112, 1981.
11. B.K.P. Horn, ROBOT VISION. M.I.T. Press: Cambridge, MA, 1986.
12. T. Poggio, V. Torre, and C. Koch, "Computational vision and regularization theory," NATURE vol. 317, pp. 314–319, 1985.
13. J.L. Marroquin, "Probabilistic solution of inverse problems," M.I.T. Artif. Intell. Lab., Cambridge, MA, A.I.-TR 860, 1985.
14. S.T. Barnard, "A stochastic approach to stereo vision," in PROC. NATL. CONF. ARTIF. INTELL. AAAI-86, Philadelphia, PA, 1976, pp. 676–680.
15. K. Mori, M. Kidodi, and H. Asada, "An iterative prediction and correction method for automatic stereo comparison," COMPUT. GRAPHICS IMAGE PROCESSING vol. 2, pp. 393–401, 1973.
16. M.J. Hannah, "Computer matching of areas in stereo images," Stanford Artif. Intell. Lab., A.I. Memo. AIM-239, July 1974.
17. H.P. Moravec, "Towards automatic visual obstacle avoidance," in PROC. FIFTH INT. JOINT CONF. ARTIF. INTELL. Cambridge, MA, 1977, p. 584.
18. D. Marr and T. Poggio, "A theory of human stereo vision," PROC. R. SOC. LOND. [B] vol. 204, pp. 301–328, 1979.
19. D.B. Gennery, "Modeling the environment of an exploring vehicle by means of stereo vision," PhD thesis, Stanford Artif. Intell. Lab., A.I. Memo. 339, 1980.
20. L.H. Quam, "Hierarchical warp stereo," in PROC. IMAGE UNDERSTANDING WORKSHOP, New Orleans, LA, October 1984, pp. 149–155.
21. N.S. Bakhvalov, NUMERICAL METHODS. Mir: Moscow, 1977.
22. G. Dahlquist and A. Björck, NUMERICAL METHODS, N. Anderson (trans.). Prentice-Hall: Englewood Cliffs, NJ, 1974.