COMPUTER VISION, GRAPHICS, AND IMAGE PROCESSING 31, 139–155 (1985)

Early Vision: From Computational Structure to Algorithms and Parallel Hardware

TOMASO POGGIO

Artificial Intelligence Laboratory and Center for Biological Information Processing, Massachusetts Institute of Technology, 545 Technology Square, Cambridge, Massachusetts 02139

Received November 15, 1984

I review a new theoretical framework that from the computational nature of early vision leads to algorithms for solving them and suggests a specific class of appropriate hardware. The common computational structure of many early vision problems is that they are mathematically ill-posed in the sense of Hadamard. Standard regularization analysis can be used to solve them in terms of variational principles that enforce constraints derived from a physical analysis of the problem, see T. Poggio and V. Torre (Artificial Intelligence Lab. Memo No. 773, MIT, Cambridge, Mass., 1984). Studies of human perception may reveal whether some principles of a similar type are exploited by biological vision. It can also be shown that the corresponding variational principles are implemented in a natural way by analog networks, see T. Poggio and C. Koch (Artificial Intelligence Lab. Memo No. 783, MIT, Cambridge, Mass., 1984). Specific electrical and chemical networks for localizing edges and computing visual motion are derived. These results suggest that local circuits of neurons may exploit this unconventional model of computation. © 1985 Academic Press, Inc.

1. INTRODUCTION

One of the best definitions of early vision is that it is inverse optics—a set of computational problems that both machines and biological organisms have to solve. While in classical optics the problem is to determine the images of physical objects, vision is confronted with the inverse problem of recovering 3-dimensional shape from the light distribution in the image. Most processes of early vision such as stereomatching, computation of motion, and all the "structure from" processes can be regarded as solutions to inverse problems. This common characteristic of early vision can be formalized: *most early vision problems are "ill-posed problems" in the sense of Hadamard*. In this article we will first review a framework proposed by Poggio and Torre [54]. They suggested that the mathematical theory developed for regularizing ill-posed problems leads in a natural way to the solution of early vision problems in terms of variational principles of a certain class. They argued that this is a theoretical framework for some of the variational solutions already obtained in the analysis of early vision processes. They also showed how several other problems in early vision can be approached and solved. Thus the computational, ill-posed nature of early vision problems dictates a specific class of algorithms for solving them, based on variational principles of a certain class. It is natural to consider next which classes of parallel hardware may efficiently implement regularization algorithms. We are especially interested in implementations that are suggestive for biology. I will thus review a model of computation proposed by Poggio and Koch [53] that maps easily into biologically plausible mechanisms. They showed that a natural way of implementing variational principles of the regularization type is to use electrical, chemical, or neuronal networks. They also showed how to derive specific networks

0734-189X/85 $3.00

for solving several low-level vision problems, such as the computation of visual motion and edge detection.

1.1. *Variational Solutions to Vision Problems*

In recent years, the computational approach to vision has begun to shed some light on several specific problems. One of the recurring themes of this theoretical analysis is the identification of physical constraints that make a given computational problem determined and solvable. Some of the early and most successful examples are the analyses of stereo matching (Marr and Poggio [43, 44]; Grimson [15, 16]; Mayhew and Frisby [45]; Kass [32]; for a review, see Nishihara and Poggio [50]) and structure from motion (Ullman [68]). In these studies constraints such as continuity of 3-D surfaces in the case of stereo matching and rigidity of objects in the case of structure from motion play a critical role for obtaining a solution.

More recently, variational principles have been used to introduce specific physical constraints. A variational principle defines the solution to a problem as the function that minimizes an appropriate cost function. Many problems can be formulated in this way, including laws that are normally expressed in terms of differential equations. In physics, for instance, most of the basic laws have a compact formulation in terms of variational principles, that require the minimization of a suitable functional, such as the Lagrangian for classical mechanics. In vision, the problem of interpolating visual surfaces through sparse depth data can be solved by minimizing functionals that embed a constraint of smoothness [16, 17, 60, 61]. Thus, the surface that best interpolates the data minimizes a certain cost functional which measures how much the surface deviates from smoothness. Computational of the motion field in the image can be successfully performed by finding the smoothest velocity field consistent with the data [26, 20, 21]: in other words, among all possible velocity fields that are consistent with the data a solution can be found by choosing the velocity field that varies the least. In a similar way, shape can be recovered from shading information in terms of a similar variational method [28].[1]

We wish to show that these variational principles follow in a natural and rigorous way from the ill-posed nature of early vision problems. We will then propose a general framework for "solving" many of the processes of early vision.

1.2. *Ill-Posed Problems*

Hadamard (1923) defined a mathematical problem to be *well posed* when its solution

(a) exists

(b) is unique

(c) depends continuously on the initial data (this condition is essentially equivalent to saying that the solution is robust against noise, because it will change only a little for small perturbations of the input data).

[1] The computation of subjective contours [67, 4, 25], of lightness [24], and of shape from contours [1, 5] can also be formulated in terms of variational principles. Terzopoulos [60, 63] has recently reviewed the use of a certain class of variational principles in vision problems within a rigorous theoretical framework.

Most of the problems of classical physics are well posed, and Hadamard argued that physical problems have to be well posed. "Inverse" problems, however, are usually ill-posed. Inverse problems can usually be obtained from the direct problems by exchanging the role of solution and data. Consider, for instance,

$$y = Az \tag{1}$$

where A is a known operator. The direct problem is to determine y from z, the inverse problem is to obtain z when y ("the data") are given. Though the direct problem is usually well posed, the inverse problem is usually ill-posed, when z and y belong to a Hilbert space.

Typical ill-posed problems are analytic continuation, backsolving the heat equation, superresolution, computer tomography, image restoration, and the determination of the shape of a drum from its frequency of vibration, a problem which was made famous by Kac [30]. In early vision, most problems are ill-posed because the solution is not unique (but see later the case of edge detection), since the operator corresponding to A is usually not injective, as in the case of shape from shading, surface interpolation, and computation of motion (see Poggio and Torre [54]).

1.3. Regularization Methods

Rigorous regularization theories for "solving" ill-posed problems have been developed during the past years (see especially Tikhonov [64], Tikhonov and Arsenin [65], and Nashed [48, 49]. Most ill-posed problems are not sufficiently constrained. To regularize them and make them well posed, one has to introduce generic constraints on the problem. In this way, one attempts to force the solution to lie in a subspace of the solution space, where it is well defined. The basic idea of regularization techniques is to restrict the space of acceptable solutions by choosing the function that minimizes an appropriate functional. The regularization of the ill-posed problem of finding z from the data y such that $Az = y$ requires the choice of norms $\|\cdot\|$ (usually quadratic) and of a *stabilizing functional* $\|Pz\|$. The choice is dictated by mathematical considerations, and, most importantly, by a physical analysis of the generic constraints on the problem. Three methods that can be applied (see Bertero [3]) among the several standard regularization techniques are:

(I) Among z that satisfy $\|Pz\| \leq \mathbf{C}$, where C is a constant, find z that minimizes

$$\|Az - y\|, \tag{2}$$

(II) Among z that satisfy $\|Az - y\| \leq \mathbf{C}$, find z that minimizes

$$\|Pz\|, \tag{3}$$

(III) Find z that minimizes

$$\|Az - y\|^2 + \lambda\|Pz\|^2, \tag{4}$$

where λ is a regularization parameter.

The first method consists of finding the function z that satisfies the constraint $\|Pz\| \leq \mathbf{C}$ and best approximates the data. The second method computes the function z that is sufficiently close to the data (C depends on the estimated errors and is zero if the data are noiseless) and is most "regular." In the third method, the regularization parameter λ controls the compromise between the degree of regularization of the solution and its closeness to the data. Standard regularization theory provides techniques to determine the best λ [65, 69]. It also provides a large body of results about the form of the *stabilizing functional* P that ensure uniqueness of the result and convergence. For instance, it is usually possible to ensure uniqueness in the case of Tikhonov's stabilizing functionals (also called *stabilizers of pth order*) defined by

$$\|Pz\|^2 = \int \sum_{r=0}^{p} c_r(\xi)\left(\frac{d^r z}{d\xi^r}\right)^2 d\xi, \tag{5}$$

where $c(\xi)$ are positive weighting factors. Equation (5) can be extended in the natural way to several dimensions. If one seeks regularized solutions of Eq. (1) with P given by Eq. (5) in the Sobolev space W_2^p of functions that have square-integrable derivatives up to pth order, the solution can be shown to be unique (up to the null space of P), if A is linear and continuous. This is because for every p the space W_2^p is a Hilbert space and $\|Pz\|^2$ is a quadratic functional (see Theorem 1, [65, p. 63]). They all correspond to either interpolating or approximating splines (for method II and method III, respectively). In the following, I will refer to regularization methods based on Tikhonov stabilizers as *standard regularization theory*. It turns out that most stabilizing functionals used so far in early vision are of the Tikhonov type (see also Terzopoulos, [60, 62]). I will discuss later the limitations of standard regularization theory and the need to develop nonstandard regularization methods (possibly, but not necessarily, of the type of Eqs. (2), (3), and (4)) for solving satisfactorily basic problems in vision.

1.4. Example I. Motion

Our first claim is that variational principles introduced recently in early vision for the problem of computation of motion and surface interpolation and approximation are exactly equivalent to standard regularization techniques. The associated uniqueness results are directly provided by regularization theory. We briefly discuss the case of motion computation in its recent formulation by Hildreth [20, 21].

Consider the problem of determining the 2-dimensional velocity field along a contour in the image. Local motion measurements along contours provide only the component of velocity in the direction perpendicular to the contour. The component of velocity tangential to the contour is invisible to a local detector that examines a restricted region of the contour. Figure 1 shows how the local velocity vector $\mathbf{V}(s)$ is decomposed into a perpendicular and a tangential component to the curve

$$\mathbf{V}(s) = v^{\mathrm{T}}(s)\mathbf{T}(s) + v^{\perp}(s)\mathbf{N}(s). \tag{6}$$

The perpendicular component $v^{\perp}$ and direction vectors $\mathbf{T}(s)$ and $\mathbf{N}(s)$, are given directly by the initial measurements, the "data." The tangential component $v^{\mathrm{T}}(s)$ is not and must be recovered to compute the full 2-dimensional velocity field $\mathbf{V}(s)$. Thus the "inverse" problem of recovering $\mathbf{V}(s)$ from the data is ill-posed because the

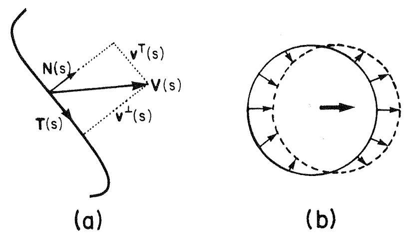

FIG. 1. Decomposition and ambiguity of the velocity field: (a) The local velocity vector $\mathbf{V}(s)$ in the image plane is decomposed according to Eq. (6) into components perpendicular and tangent to the curve. (b) Local measurements cannot measure the full velocity field; the circle undergoes pure translation; the arrows represent the perpendicular components of velocity that can be measured from the images [54].

solution is not unique. Mathematically, this arises because the operator K defined by

$$v^{\perp} = K\mathbf{V} \tag{7}$$

is not injective. Equation (7) describes the imaging process as applied to the physical velocity field $\mathbf{V}$ which consists of the x and y components of the velocity field on the image plane.

Intuitively, the set of measurements given by $v^{\perp}(s)$ over an extended contour should provide considerable constraint on the motion of the contour. An additional generic constraint, however, is needed to determine this motion uniquely. For instance, rigid motion on the plane is sufficient to determine $\mathbf{V}$ uniquely but is very restrictive, since it does not cover the case of motion of a rigid object in space. Hildreth suggested, following Horn and Schunck [26], that a more general constraint is to find the smoothest velocity field among the set of possible velocity fields consistent with the measurements. The choice of the specific form of this constraint was guided by physical considerations—the real world consists of solid objects with smooth surfaces whose projected velocity field is usually smooth—and by mathematical considerations—especially uniqueness of the solution. Hildreth proposed two algorithms: in the case of exact data the functional to be minimized is a measure of the smoothness of the velocity field

$$\|P\mathbf{V}\|^2 = \int \left(\frac{\partial \mathbf{V}}{\partial s}\right)^2 ds \tag{8}$$

subject to the measurements $v^{\perp}(s)$. Since in general there will be error in the measurements of $v^{\perp}$, the alternative method is to find $\mathbf{V}$ that minimizes

$$\|K\mathbf{V} - v^{\perp}\|^2 + \lambda \int \left(\frac{\partial \mathbf{V}}{\partial s}\right)^2 ds. \tag{9}$$

It is immediately seen that these schemes correspond to the second and third regularizing method, respectively. (The constraint of rigid translatory motion in the image plane corresponds to using the first regularizing method with the same P and $C = 0$.) Uniqueness of the solutions (proved by Hildreth for the case of Eq. (8)) is a direct consequence for both Eqs. (8) and (9) of standard theorems of regularization theories. In addition, other results can be used to characterize how the correct solution converges depending on the smoothing parameter λ (Geiger, in preparation).

1.5. Example II. Edge Detection

We have recently applied regularization techniques to another classical problem of early vision—edge detection. Edge detection, intended as the process that attempts to detect and localize changes of intensity in the image (this definition does not encompass all the meanings of edge detection) is a problem of numerical differentiation [66]. Notice that differentiation is a common operation in early vision and is not restricted to edge detection. The problem is ill-posed because the solution does not depend continuously on the data.

The intuitive reason for the ill-posed nature of the problem can be seen by considering a function $f(x)$ perturbed by a very small (in L_2 norm) "noise" term $\varepsilon \sin \Omega x$. $f(x)$ and $f(x) + \varepsilon \sin \Omega x$ can be arbitrarily close for very small ε, but their derivatives may be very different if Ω is large enough. This simply means that a derivative operation "amplifies" high-frequency noise.

In 1-D, numerical differentiation can be regularized in the following way. The "image" model is $y_i = f(x_i) + \varepsilon_i$, where y_i is the data and ε_i represent errors in the measurements. We want to estimate f'. We chose a regularizing functional $\|Pf\| = \int (f''(x))^2 \, dx$, where f'' is the second derivative of f. This choice corresponds to a constraint of smoothness on the intensity profile. Its physical justification is that the (noiseless) image is indeed very smooth because of the imaging process: the image is a bandlimited function and has therefore bounded derivatives. The second regularizing method (no noise in the data) is equivalent then to using interpolating cubic splines for differentiation. The third regularizing method, which is more natural since it takes into account errors in the measurements, leads to the variational problem of minimizing (see [57])

$$\sum (y_i - f(x_i))^2 + \lambda \int (f''(x))^2 \, dx. \tag{10}$$

Poggio *et al.* [55] have shown (a) that the solution f of this problem can be obtained by convolving the data y_i (assumed on a regular grid and satisfying appropriate boundary conditions) with a convolution filter R, and (b) that the filter R is a cubic spline with a shape very close to a Gaussian and a size controlled by the regularization parameter λ (see Fig. 2). Differentiation can then be accomplished by convolution of the data with the appropriate derivative of this filter. The optimal value of λ can be determined for instance by cross validation and other techniques. This corresponds to finding the optimal scale of the filter (see [54]).

These results can be directly extended to two dimensions to cover both edge detection and surface interpolation and approximation. The resulting filters are very similar to two of the edge detection filters derived and extensively used in recent years [42, 8, 66]. The Laplacian of the optimal filter found in this way seems, however, to have slightly better performance than the Laplacian of a Gaussian.

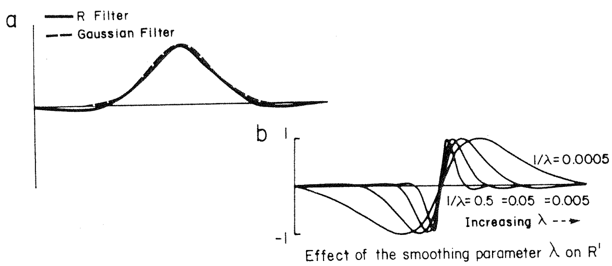

FIG. 2. The edge detection filter: (a) The convolution filter obtained (15) by regularizing the ill-posed problem of edge detection with method (III) (solid line). It is a cubic spline, very similar to a Gaussian (dotted line). (b) The first derivative of the filter for different values of the regularizing parameter λ, which effectively controls the *scale* of the filter. This 1-dimensional profile can be used for 2-dimensional edge detection by filtering the image with oriented filters with this transversal crossection and choosing the orientation with maximum response (see 16). The second derivative of the filter (not shown here) is quite similar to the second derivative of a Gaussian [66].

It is important to notice that the result about convolution holds true more in general: *if the data are given on a regular grid with appropriate boundary conditions, the solution of a standard regularization principle (Eq. 4, with Tikhonov stabilizers) is equivalent to convolving the data with a precomputed filter* [55].

Other problems in early vision such as shape from shading [28] and surface interpolation [16, 17, 60–62] in addition to the computation of velocity, have already been formulated and "solved" in similar ways using variational principles of the type suggested by regularization techniques (although this was not realized at the time—also, Ikeuchi and Horn's formulation is nonquadratic). It is also clear that other problems such as stereo and structure from motion can be approached in terms of, possibly nonstandard, regularization analysis (see [54]).

1.6. Physical Plausibility of the Solution

Uniqueness of the solution of the regularized problem—which is ensured by formulations such as Eqs. (2)–(4)—is not the only (or even the most relevant) concern of regularization analysis. Physical plausibility of the solution is the most important criterion. The decision regarding the choice of the appropriate stabilizing functional cannot be made judiciously from purely mathematical considerations. A physical analysis of the problem and of its generic constraints play the main role. Regularization theory provides a framework within which one has to seek constraints that are rooted in the physics of the visual world. This is, of course, the challenge of regularization analysis.

In our example of the computation of motion the constraint of smoothness is justified by the observation that the projection of 3-dimensional objects in motion onto the image plane tends, in a probabilistic sense, to yield smoother velocity fields (see [20, 21]). In the case of edge detection the constraint on the derivative of image intensity is justified by the bandlimiting properties of the optics. In the case of motion, however, as more dramatically in the case of surface reconstruction, the constraint of smoothness is not always correct. This suggests that more general stabilizing functionals are needed to deal with the general problem of discontinuities (see discussion).

A method for checking physical plausibility of a variational principle is, of course, computer simulation. A simple technique we suggest is to use the Euler–Lagrange equation associated with the variational problem. In the computation of motion, Yuille [70] has obtained the following *sufficient and necessary* condition for the solution of the variational principle equation (8), to be the correct physical solution

$$\mathbf{T} \cdot \frac{\partial^2 \mathbf{V}}{\partial s^2} = 0$$

where $\mathbf{T}$ is the tangent vector to the contour and $\mathbf{V}$ is the true velocity field. The equation is satisfied by uniform translation or expansion and by rotation only if the contour is polygonal. These results suggest that algorithms based on the smoothness principle will give correct results, and hence be useful for computer vision systems, when (a) motion can be approximated locally by pure translation, rotation, or expansion, or (b) objects have images consisting of connected straight lines.

In the case of edge detection (intended as numerical differentiation), the solution is correct *if and only if* the intensity profile is a polynomial spline of odd degree [55].

From a more biological point of view, a careful comparison of the various "regularization" solutions with human perception promises to be a very interesting area of research, as suggested by Hildreth's work on the computation of motion. For some classes of motions and contours, the solution of Eqs. (8) and (9) is not the physically correct velocity field. In these cases, however, the human visual system also appears to derive a similar, incorrect velocity field [20, 21].

2. ANALOG NETWORKS FOR SOLVING VARIATIONAL PROBLEMS

We consider now the question of which class of parallel hardware could efficiently solve variational principles of the regularization type. The specific architecture depends of course on the norm and the stabilizer P that are chosen. Standard regularization methods (with quadratic norms and Tikhonov's stabilizers) map into two main classes of algorithms: *convolution algorithms* (for data on a regular grid) and *multigrid algorithms*. These two classes of algorithms can be efficiently implemented by architectures of N simple processors with local interconnections (possibly with a multilevel structure) [60, 62]. Digital architecture of this type have only a limited interest for biology. Poggio and Koch [53] have suggested a more "exotic" type of hardware for implementing regularization solutions that suggests a new model for neural computations.

It is well known that analog networks—chemical, electrical, or mechanical—are a natural computational model for solving variational principles. The behavior of such systems can be described using variational principle. Electrical network representations have been constructed for practically all of the field equations of physics—many of them are equivalent to variational principles (for an electrical network implementation of Schrodinger's equation, see Kron [36]). A fundamental reason for the natural mapping between variational principles and electrical or chemical networks is Hamilton's least action principle (for more details see Koch and Poggio [35]).

The class of variational principles that can be computed by analog networks is dictated by Kirchhoff's current and voltage laws (KCL and KVL), which simply represent conservation and continuity restrictions satisfied by each network component (appropriate variables are usually voltage and current for electrical networks and affinity, i.e., chemical potential and chemical turnover rate for chemical systems). KCL and KVL provide the unifying structure of network theory. A large body of theoretical results is available about networks satisfying them, including classical thermodynamics [52]. In particular, KCL and KVL imply Tellegen's theorem. Tellegen's theorem captures the basic constraints provided by KCL and KVL. It is one of the most general and powerful results of network theory and is independent of any assumptions about constitutive relations or stationarity. (*Tellegen's theorem*: If $\mathbf{U}$ is the vector of branch potentials—with one component for each branch—and $\mathbf{J}$ is the vector of branch flows, then $\mathbf{U}^t \cdot \mathbf{J} = 0$. Thus the flow and the potential variables are orthogonal at any instant in time.)

For a network containing only sources and linear resistances, Tellegen's theorem implies Maxwell's minimum heat theorem: *the distribution of voltages and currents is such that it minimizes the total power dissipated as heat*. These results can be extended to nonlinear circuit components [39, 51, 53], but in the following we will restrict ourselves to linear networks (possibly with negative resistances). The power dissipated by each linear resistance in the circuit is a quadratic term of the form

$$I_k V_k \tag{11}$$

where I_k and V_k are the current and the voltage respectively, corresponding to the resistive process r_k. It follows that any network consisting of linear resistances and voltage sources E_k minimizes the following associated quadratic functional

$$\sum_k r_k I_k^2 - \sum_i E_i I_i, \tag{12a}$$

where the second sum includes all the batteries. For a network of resistances and current sources I_i, the functional is given by

$$\sum_k g_k V_k^2 - \sum_i I_i V_i, \tag{12b}$$

where the second sum includes all the current sources and $g_k = 1/r_k$.

It is then easy to show the equivalence of Eqs. (12) and the regularization principle Eqs. (4), (5). Thus, electrical networks of linear resistances and batteries (or current sources) can *solve* quadratic variational principles of the form of Eqs. (4), (5). The solution is unique when Eqs. (4), (5) yields a unique solution (which is usually the case, see [54]).

Electrical networks of resistances and batteries do not have any dynamics. In practice, however, small capacitances will be present and the *stability* of the network must then be considered. It turns out that networks implementing regularization principles of the form of Eqs. (4), (5) are indeed stable, under the same conditions that ensure a unique solution [53].

An equivalent way to see how electrical networks can implement variational principles of the form of Eqs. (4), (5) is to consider the associated Euler–Lagrange equations (the equivalence of variational principles with PDE also shows how to map them into parallel digital architecture). Since the functional to be minimized is quadratic, the Euler–Lagrange equations are linear, of the form $Qz = b$. They have a unique solution z, corresponding to the unique solution of the variational principle. In the discrete case, these equations correspond to n linear, coupled algebraic equations. These equations can be implemented in a network containing only linear resistances and sources. More precisely, the vector b, which depends on the data $(b = A^* y)$, can always be represented in terms of current or voltage sources. The matrix Q corresponds to the symmetric, real matrix of the network resistances [53].

Although their procedure will always yield an electrical network with linear elements implementing $Q_z = b$, its physical realization might require negative resistances (if the corresponding term in Q is negative). An alternative implementation of variational principles, common on analog computers, involves operational amplifiers [29].

As pointed out by Terzopoulos in the context of vision (earlier, Horn [24] proposed an analog implementation of the lightness computation) a significant advantage of analog networks is their extreme parallelism and speed of convergence. Furthermore, resistive networks are robust against random errors in the values of the resistances [31]. A disadvantage is the limited precision of the analog signals.

2.1. *An Example. Circuits for the Velocity Field Computation*

We will consider next some specific networks for solving the optical flow computation. The simpler case is when the measurements of the perpendicular component of the velocity, $v_i^\perp$, at n points along the contours, are exact. In this case, the discretized Euler–Lagrange equations, corresponding to the regularization solution, Eq. (3), are [20]

$$\left(2 + \kappa_i^2\right) v_i^{\mathrm{T}} - v_{i+1}^{\mathrm{T}} - v_{i-1}^{\mathrm{T}} = d_i, \tag{13}$$

where κ is the curvature of the curve at location i, d_i is a function of the data $v_i^\perp$ and the curve and v_i^{T} is the unknown tangential component of the velocity v_i at location i to be computed. Figures 3a and 3b show two simple networks that solve Eq. (13), where one network is the dual of the other. The equation describing the ith node, in the case of Fig. 3b, is

$$(2g + g_i) V_i - g V_{i+1} - g V_{i-1} = I_i \tag{14}$$

where V_i is the voltage—corresponding to the unknown v_i^{T}—and I_i the injected current at node i—corresponding to the measurement $v_i^\perp$. It is surprising that this implementation does not require negative resistances. When the constraints are satisfied only approximately (Eq. (5)), the equations are

$$\begin{aligned} \left(2 + l_{xi}^2\right) V_{x_i} - V_{x_{i+1}} - V_{x_{i-1}} + c_i V_{y_i} &= d_{x_i} \\ \left(2 + l_{y_i}^2\right) V_{y_i} - V_{y_{i+1}} - V_{y_{i-1}} + c_i V_{x_i} &= d_{y_i} \end{aligned} \tag{11}$$

where l_i depends on the contour and V_{x_i} and V_{y_i} denote the x and y component of

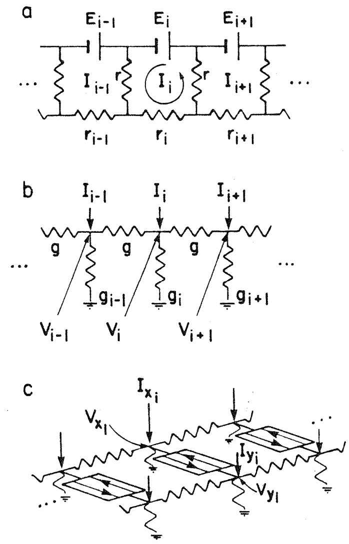

FIG. 3. Resistive networks computing the smoothest velocity field. The first two networks correspond to the situation where the constraints imposed by the data are to be satisfied exactly. The equation for the current, which corresponds to the desired v^{T} in mesh i (for Fig. 1a), is given by $(2r + r_i)I_i - rI_{i+1} - rI_{i-1} = E_i$, where the value of the battery E_i depends on the velocity data $v_i^{\perp}$ at location i. The voltage at node i, corresponding to v_i^{T}, for the network 1b, the dual of network 1a, is given by $(2g + g_i)V_i - gV_{i+1} - gV_{i-1} = I_i$, where the injected current I_i depends on the velocity data. Sampling the voltage between nodes corresponds to linear interpolation between the node values. Network 1c, consisting of two interconnected networks of the type shown in 1b, solves the velocity field problem when the data are not exact. The equations for the ith nodes are $(2g_x + g_{x_i})V_{x_i} - g_x V_{x_{i+1}} - g_x V_{x_{i-1}} + c_i V_{y_i} = d_{x_i}$ and $(2g_y + g_{y_i})V_{y_i} - g_y V_{y_{i+1}} - g_y V_{y_{i-1}} + c_i V_{x_i} = d_{y_i}$. However, unlike the two purely passive networks shown above, an active element may be required, since the cross-term c_i, relating the x and the y components of velocity, can be negative. Such a negative resistance can be mimicked by operational amplifiers [53].

the unknown velocity v_i at location i. The corresponding network is shown in Fig. 3c. The resistances c_i can be either positive or negative, and may therefore require active components such as operational amplifiers. More precisely, physically realizable linear resistances, whether in electrical or in chemical systems, must dissipate energy, i.e., they are constrained to the upper right and the lower left quadrant in the I–V plane and can thus only be positive. There are at least three options for implementing negative resistances using basic circuit components: (i) The positive and negative resistances can be replaced in a purely resistive network by inductances and capacitances, with impedance $i\omega L$ and $-i/(\omega C)$, respectively. The network equations are then formulated in terms of the currents and voltages at the fixed frequency ω. (ii) The negative resistance can be implemented by the use of operational amplifiers or similar active circuit elements. (iii) One may exploit the negative impedance regions in such highly nonlinear systems as the tunnel diode.

In the limit, as the meshes of the circuit become infinitesimally small, the network solves the continuous variational problem, and not simply its discrete approximation.

We have devised similar analog networks for solving other variational problems [53] arising from regularization analysis of several early vision problems such as edge detection [55] and surface interpolation [60, 62]. These networks are analog solutions to certain kinds of spline interpolation and approximation problems. For instance, in the case of surface interpolation the analog network solves the *biharmonic* equation which is the Euler–Lagrange equation corresponding to the variational problem associated with thin-plate splines [60]. The stabilizing functionals used in regularization analysis of vision problems typically lead to local and limited connections between the components of the network.

2.2. *Solving Ill-posed Problems with Biological Hardware*

Analog electrical networks are a natural hardware for computing the class of variational principles suggested by regularization analysis. Because of the well-known isomorphism between electrical and chemical networks (see, e.g., [7 or 10]) that derives from the common underlying mathematical structure, appropriate sets of chemical reactions can be devised, at least in principle, to "simulate" exactly the electrical circuits. Figure 4 shows chemical networks that are equivalent (in the steady state) to the electrical circuit of Figs. 3b and c.

Electrical and chemical systems of this type therefore offer a computational model for early vision that is quite difference from the digital computer. Equations are "solved" in an implicit way, exploiting the physical constraints provided by Kirchhoff's laws. It is not difficult to imagine how this model of computation could be extended to mixed electrochemical systems by the use of transducers, such as chemical synapses, that can decouple two parts of a system, similarly to operational amplifiers [53].

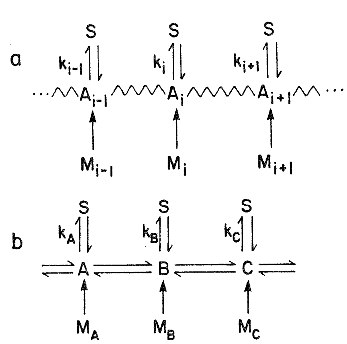

FIG. 4. Two examples of chemical networks solving the motion problem for exact measurements. They are equivalent, under steady-state conditions, to the electric circuit of Fig. 3b. Figure 4a illustrates a diffusion-reaction system. A substance A (the concentration of which corresponds to the desired v^T) diffuses along a cable while reacting with an extracellular substance S (first-order kinetics). The corresponding on-rate k_i varies from location to location. This could be achieved by a differential concentration of an enzyme catalyzing the reaction or by varying the properties of the membrane where the reaction has to take place. The off-rates can either be constant or vary with location. The inputs are given by the influxes of substance A. Figure 2b shows a lumped chemical network, where n different, well-mixed substances, interact with each other and with the substrate S. Assuming first-order kinetics, these reactions can mimick a linear positive resistance under steady-state conditions. The input is given by the influx M_X and the output by the concentration of X [53].

Could neural hardware exploit this model of computation? Increasing evidence shows that electronic potentials play a primary role in many neurons [58] and that membrane properties such as resistance, capacitance, and equivalent inductance (arising through voltage and time-dependent conductances; see, e.g., [9, or 34] may be effectively modulated by various types of neurotransmitters, acting over very different time scales [41]. Dendrodendritic synapses and gap junctions serve to mediate graded, analog interactions between neurons and do not rely on all-or-none action potentials [14].

When implementing electrical networks in equivalent neuronal hardware, one can exploit a number of elementary circuit elements (for possible neuronal implementations; see Fig. 5). Patches of neuronal membrane or cytoplasm can be treated as resistance and capacitance. Voltage sources may be mimicked by synapses on dendritic spines [35] or on small dendrites, whereas synapses on large dendrites act as current sources. Chemical synapses could effectively serve to decouple different parts of a network (see [53]). Chemical processes such as the reactions associated with postsynaptic effects or with neuropeptides could also be thought as part of a complex electrochemical network. Obviously, the analogy cannot be taken too literally. It would be very surprising to find the exact neural analog of the circuit of Fig. 5 somewhere in the CNS. We are convinced, however, that the *style* of computation represented by analog circuits represents a very useful model for neural computations as well as a challenge for future VLSI circuit designs.

3. CONCLUSION

The concept of ill-posed problems and the associated regularization theories seem to provide a satisfactory theoretical framework for much of early vision. This perspective justifies the use of variational principles of a certain type for solving specific problems, and suggests how to approach other early vision problems. It provides a link between the computational (ill-posed) nature of the problems and the computational structure of the solution (as a variational principle). It also suggests computational "hardware" that is natural for solving variational problems of the type implied by regularization methods.

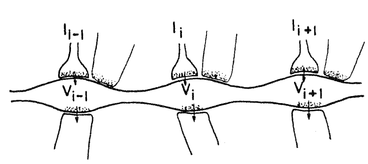

FIG. 5. This schematic figure illustrates a hypothetical neuronal implementation of the regularization solution of the motion problem. A dendrite, acting both as pre- and post-synaptic element has a membrane resistance that can vary with location. It can implement under steady-state conditions the circuit 3b. The inputs—corresponding to the measurements $v^{\perp}$—are given by synaptic mediated currents, while the output voltages—corresponding to the desired $v^{\top}$—are sampled by dendro-dendritic synapses. The membrane resistance can be locally controlled by suitable synaptic inputs—corresponding to the curvature of the contour—from additional synapses that open channels with a reversal potential close to the resting potential of the dendrite. This scheme can be extended to the case where the measurements of the perpendicular velocities are not exact, by having a similar, second dendrite (see also Fig. 3c). The interaction between both dendrites takes place via two reciprocal chemical synapses. If the corresponding cross-term in Eq. (15) is negative, the chemical synapses must be inverting, presynaptic depolarization leading to a hyperpolarization [53].

Despite its attractions, this theoretical synthesis of early vision also shows the limitations that are intrinsic to the variational solutions proposed so far, and in any case to the *standard* (Tikhonov's) forms of the regularization approach. The basic problem is the degree of smoothness required for the unknown function z that has to be recovered. If z is very smooth, then it will be robust against noise in the data, but it may be too smooth to be physically plausible. For instance, in visual surface interpolation, the degree of smoothness obtained from a specific form of Eqs. (4), (5) —corresponding to so-called thin plate splines—smoothes depth discontinuities too much and often leads to unrealistic results (but see [60, 62]). An interesting approach to this problem is to parametrize with an additional parameter λ—a function of position—the order of the Tikhonov stabilizer. The question is then how to determine the optimal value of the parameter.

Different (e.g., nonquadratic) variational principles may be used to attack the general problem of discontinuities. Nonstandard variational principles may also arise in another one of the most fundamental problems in early vision, the problem of integrating different sources of information, such as stereo, motion, shape from shading, etc. This problem is ill-posed, not just because the solution is not unique (the standard case), but because the solution is usually overconstrained and may not exist (because of noise in the data). For instance, the problem of combining several different sources of surface information may easily lead to nonquadratic regularization expressions (though different "noninteracting" constraints can be combined in a convex way, see Terzopoulos, [60, 62]). These minimization problems will in general have multiple local minima.

Again, analog networks may be used to solve these minimization problems, with multiple local minima corresponding to the zeros of the *mixed potential* [6, 52]. Schemes similar to *annealing* [47, 33, 22] may be easily implemented by appropriate sources of Gaussian noise driving the analog network. The associated differential equation describing the dynamics of the system is then a stochastic differential equation. The stochastic differential equations ("Langevin" equations) describing an electrical or a chemical system with a source of Gaussian noise (e.g., voltage or the presence of a chemical reactive substance) can be formulated in terms of Ito or Stratonovitch calculus [13]. They can be solved with the Fokker–Planck or the Kolmogorov method. A "solution" of a stochastic differential equation is a characterization in terms of probability distributions of the "output" process. For linear networks, simpler correlation methods can also be used. If the noise is white and Gaussian, its spectral density is proportional to the "temperature" T. In a chemical network "noise" may be introduced in various, simple ways.

Needless to say, a number of biophysical mechanisms, such as somatic and dendritic action potentials, interactions between conductance changes, voltage, and time-dependent conductances, etc., are likely to be used by neurons and patches of membrane to perform a variety of nonlinear operations.

I conclude with a caution note, that hopefully will turn out to be too conservative. The range of applicability of variational principles is related to the deep question of the computational organization of a visual processor and its control structure. It is unlikely that variational principles alone could have enough flexibility to control and coordinate the different modules of early vision and their interaction with higher level knowledge. This also hints at the basic limitation of present regularization methods that makes them suitable only for the first stages of vision. They derive numerical representations—surfaces—from numerical representations—images. However, it is not difficult to see how the computation of the more symbolic type of representations that are essential for a powerful vision processor represent a form of regularization. The restriction of the solution space to a set of "symbols" regularizes an ill-posed problem. Standard regularization methods restrict the solution space to the set of generalized splines.

ACKNOWLEDGMENTS

This written version of my lecture is a porridge of two previous technical reports [54, 53]. The idea of vision problems as ill-posed problems originated from a conversation with Professor Mario Bertero of the University of Genoa. We are grateful to A. Yuille, E. Hildreth, D. Terzopoulos, J. Marroquin, T. Collett, and C. Koch for many discussions and comments. C. Bonomo made many less appreciated comments along the way as well. This paper reviews research done within the Artificial Intelligence Laboratory and the Center for Biological Information Processing (Whitaker College) at the Massachusetts Institute of Technology. Support for the A.I. Laboratory's research in artificial intelligence is provided in part by the Advanced Research Projects Agency of the Department of Defense under Office of Naval Research contract N00014-80-C-0505.

REFERENCES

1. H. G. Barrow and J. M. Tennenbaum, Interpreting line drawings as three dimensional surfaces, *Artif. Intell.* **17**, 1981, 75–117.
2. M. V. L. Benett, A comparison of electrically and chemically mediated transmission, in *Structure and Function of Synapses* (G. D. Pappas and D. P. Purpura, Eds.) Raven, New York, 1972.
3. M. Bertero, Problemi lineari non ben posti e metodi di regolarizzazione, in *Problem non ben posti ed inversi*, Istituto di Analisi Globale, Florence, Italy, 1982.
4. J. M. Brady, W. E. L. Grimson, and D. J. Langridge, Shape encoding and subjective contours, in *First Annual Natl. Conf. on Artificial Intelligence*, 1980, pp. 15–17.
5. J. M. Brady and A. Yuille, An extremum principle for shape from contour, *IEEE Trans. Pattern Anal. Mach. Intell.* **PAMI-288**, 1984, 301.
6. R. K. Brayton and J. K. Moser, A theory of nonlinear networks, I, *Quart. Appl. Math.* **22**, 1964, 1–33.
7. H. Busse and B. Hess, Information transmission in a diffusion-coupled oscillatory chemical system, *Nature (London)* **244**, 1973, 203–205.
8. J. F. Canny, Finding Edges and Lines in Images, MIT Artificial Intelligence Laboratory Technical Report TR-720, June 1983.
9. K. S. Cole, *Membranes Ions and Impulses*, Univ. of California Press, Berkeley, 1968.
10. M. Eigen, Molecules, information, and memory: From molecules to neural networks, in *The Neurosciences: Third Study Program* (F. O. Schmitt and F. G. Worden, Eds.), MIT Press, Cambridge, Mass. 1974.
11. S. A. Ellias and J. K. Stevens, The dendritic varicosity: A mechanism for electrically isolating the dendrites of cat retinal amacrine cells? *Brain Res.* **196**, 1972, 365–372.
12. S. Geman and D. Geman, Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images, IEEE *Trans. Pattern Anal. Mach. Intell.*, **6**, 1984, 721–741.
13. I. I. Gihman and A. V. Skorohod, *Stochastic Differential Equations*, Springer-Verlag, Berlin/New York, 1972.
14. K. Graubard and W. H. Calvin, Presynaptic dendrites: Implications of spikeless synaptic transmission and dendritic geometry, in *The Neurosciences: Fourth Study Program* (F. O. Schmitt and F. G. Worden, Eds.), MIT Press, Cambridge, Mass., 1979.
15. W. E. L. Grimson, A computer implementation of a theory of human stereo vision, *Philos. Trans. R. Soc. London B* **292**, 1981, 217–253.

16. W. E. L. Grimson, *From Images to Surfaces: A Computational Study of the Human Early Visual System*, MIT Press, Cambridge, Mass., 1981.
17. W. E. L. Grimson, A computational theory of visual surface interpolation, *Philos. Trans. R. Soc. London B* **298**, 1982, 395–427.
18. J. Hadamard, *Lectures on the Cauchy Problem in Linear Partial Differential Equations*, Yale Univ. Press, New Haven, Conn., 1923.
19. E. C. Hildreth, Implementation of a theory of edge detection, S.M. thesis, Department of Computer Science and Electrical Engineering, MIT, 1980; AI Lab. Technical Report 597, MIT, Cambridge, Mass. 1980.
20. E. C. Hildreth, *The Measurement of Visual Motion*, MIT Press, Cambridge, Mass. 1984.
21. E. C. Hildreth, Computation of the velocity field, *Proc. R. Soc. London B* **221**, 1984, 189–220.
22. G. E. Hinton and T. J. Sejnowski, Optimal perceptual inference, in *Proc. IEEE 1983 Conf. Computer Vision and Pattern Recognition*, Washington, D.C., 1983.
23. J. J. Hopfield, Neurons with graded response have collective computational properties like those of two-state neurons, in *Proc. Natl. Acad. Sci. USA*, 1984, in press.
24. B. K. P. Horn, Determining lightness from an image, *Comput. Graphics Image Process.* **3**, 1974, 111–299.
25. B. K. P. Horn, The Least Energy Curve, AI Memo 612, MIT AI Lab., Cambridge, Mass., 1981.
26. B. K. P. Horn and B. G. Schunck, Determining optical flow, *Artif. Intell.* **17**, 1981, 185–203.
27. R. Hummel and S. Zucker, On the Foundations of Relaxation Labeling Processes, Computer Vision and Graphics Laboratory TR 80-7, McGill University, 1980.
28. K. Ikeuchi and B. K. P. Horn, Numerical shape from shading and occluding boundaries, *Artif. Intell.* **17**, 1981, 141–184.
29. D. S. Jackson, *Analog Computation*, McGraw–Hill, New York, 1960.
30. M. Kac, Can one hear the shape of a drum? *Amer. Math. Monthly* **73**, No. 4, Part II, 1966, 1–23.
31. W. J. Karplus, *Analog Simulation: Solution of Field Problems*, McGraw–Hill, New York, 1958.
32. M. H. Kass, Computating Stereo Correspondence, M.S. thesis, MIT, 1984.
33. S. Kirkpatrick, C. D. Gelatt, Jr., and M. P. Vecchi, Optimization by simulated annealing, *Science* **220**, 1983, 671–680.
34. C. Koch, Cable theory in neurons with active, linearized membranes, *Biol. Cybern.* **50**, 1984, 15–33.
35. C. Koch and T. Poggio, A theoretical analysis of electrical properties of spines, *Proc. R. Soc. London B* **218**, 1983, 455–477.
36. G. Kron, Electric circuit models of the Schrodinger equation, *Phys. Rev.* **67**, 1945, 39–43.
37. B. C. Kuo, *Linear Networks and Systems*, McGraw–Hill, New York, 1967.
38. D. G. Luenberger, *Optimization by Vector Space Methods*, Wiley, New York, 1969.
39. A. G. J. MacFarlane, *Dynamical System Models*, Harrap, London, 1970.
40. R. H. MacNeal, The solution of elastic plate problems by electrical analogies, *J. Appl. Mech.* **18**, 1951, 59–67; in *The Bending and Stretching of Plates*, Macmillan, New York.
41. E. Marder, Mechanisms underlying neurotransmitter modulation of a neuronal circuit, *Trends Neurosci.* (*Pers. Ed.*) **7**, 1984, 48–53.
42. D. Marr and E. C. Hildreth, Theory of edge detection, *Proc. R. Soc. London B* **207**, 1980, 187–217.
43. D. Marr and T. Poggio, Cooperative computation of stereo disparity, *Science* (*Washington, D.C.*) **194**, 1976, 283–287.
44. D. Marr and T. Poggio, A theory of human stereo vision, *Proc. Roy. Soc. London B* **204** 1979, 301–328; earlier version, MIT AI Lab Memo 451, 1977.
45. J. E. W. Mayhew and J. P. Frisby, Psychophysical and computational studies towards a theory of human stereopsis, *Artif. Intell.* **17**, 1981, 349–385.
46. W. S. McCulloch and W. Pitts, A logical calculus of the ideas immanent in nervous activity, *Bull. Math. Biophys.* **5**, 1943, 115–133.
47. N. Metropolis, A. Rosenbluth, M. Rosenbluth, A. Teller, and E. Teller, Equation of state calculations by fast computing machines, *J. Chem. Phys.* **21**, No. 6, 1953, 1087–1092.
48. M. Z. Nashed, Approximate regularized solutions to improperly posed linear integral and operator equations, in *Constructive and Computational Methods for Differential and Integral Equations* (G. Auger, Ed.), pp. 289–296, Akademie Verlag, Berlin, 1974.
49. M. Z. Nashed, (Ed.), *Generalized Inverses and Applications*, Academic Press, New York, 1976.
50. H. K. Nishihara and T. Poggio, Stereo vision for robotics, in *Proceedings of the First International Symposium of Robotics Research*, in press.

51. G. F. Oster and C. A. Desoer, Tellegen's theorem and thermodynamic inequalities, *J. Theor. Biol.* **32**, 1971, 219–241.
52. G. F. Oster, A. Perelson, and A. Katchalsky, Network thermodynamics, *Nature* **234**, 1971, 393–399.
53. T. Poggio and C. Koch, Analog Networks: A New Approach to Neuronal Computation, Artificial Intelligence Lab. Memo No. 783, MIT, Cambridge, Mass., 1984.
54. T. Poggio and V. Torre, Ill-posed Problems and Regularization Analysis in Early Vision, Artificial Intelligence Lab. Memo No. 773, MIT, Cambridge, Mass., April 1984.
55. T. Poggio, H. Voorhees, and A. Yuille, Regularizing Edge Detection, Artificial Intelligence Lab. Memo No. 776, MIT, Cambridge, Mass., 1984.
56. I. Prigogine, *Thermodynamics of Irreversible Processes*, Wiley–Interscience, New York, 1967.
57. C. H. Reinsch, Smoothing by spline functions, *Numer. Math.* **10**, 1967, 177–183.
58. F. O. Schmitt, P. Dev, and B. H. Smith, Electrotonic processing of information in brain cells, *Science* (Washington, D.C.), **193**, 1976, 114–120.
59. B. G. Schunck and B. K. P. Horn, Constraints on optical flow computation, in *Proc. IEEE Conf. Pattern Recognition and Image Processing*, pp. 205–210, 1981.
60. D. Terzopoulos, Multiresolution Computation of Visible-Surface Representations, Ph.D. thesis, Dept. of Electrical Engineering and Computer Science, MIT, Cambridge, Mass., 1984.
61. D. Terzopoulos, Multilevel computational processes for visual surface reconstruction, *Comput. Vision Graphics, Image Process.* **24**, 1983, 52–96.
62. D. Terzopoulos, Multilevel Reconstruction of Visual Surfaces: Variational Principles and Finite Element Representations, Artificial Intelligence Lab. Memo 671, MIT; in *Multiresolution Image Processing and Analysis* (A. Rosenfeld, Ed.), pp. 237–310, Springer-Verlag, New York/Berlin, 1984.
63. D. Terzopoulos, Integrating visual information from multiple sources for the cooperative computation of surface shape, in *From Pixels to Predicates: Recent Advances in Computational and Robotic Vision*, (A. Pentland, Ed.), Ablex, Norwood, N.J., 1985.
64. A. N. Tikhonov, Solution of incorrectly formulated problems and the regularization method, *Soviet Math. Dokl.* **4**, 1963, 1035–1038.
65. A. N. Tikhonov and V. Y. Arsenin, *Solutions of Ill-Posed Problems*, Winston & Sons, Washington, D.C., 1977.
66. V. Torre and T. Poggio, On Edge Detection, Artificial Intelligence Lab. Memo No. 768, MIT, Cambridge, Mass., March 1984.
67. S. Ullman, Filling in the gaps: The shape of subjective contours and a model for their generation, *Biol. Cybern.* **25**, 1976, 1–6.
68. S. Ullman, The interpretation of structure from motion, *Proc. R. Soc. London B* **203**, 1979, 405–426.
69. G. Wahba, Ill-Posed Problems: Numerical and Statistical Methods for Mildly, Moderately, and Severely Ill-Posed Problems with Noisy Data, Tech. Report No. 595, Univ. of Wisconsin, Madison, 1980.
70. A. Yuille, The Smoothest Velocity Field and Token Matching Schemes, Artificial Intelligence Lab. Memo No. 724, MIT, Cambridge, Mass., August 1983.