Regularization of Inverse Visual Problems Involving Discontinuities

DEMETRI TERZOPOULOS, MEMBER, IEEE

Abstract—**Inverse problems, such as the reconstruction problems that arise in early vision, tend to be mathematically ill-posed. Through regularization, they may be reformulated as well-posed variational principles whose solutions are computable. Standard regularization theory employs quadratic stabilizing functionals that impose global smoothness constraints on possible solutions. Discontinuities present serious difficulties to standard regularization, however, since their reconstruction requires a precise spatial control over the smoothing properties of stabilizers. This paper proposes a general class of controlled-continuity stabilizers which provide the necessary control over smoothness. These nonquadratic stabilizing functionals comprise multiple generalized spline kernels combined with (noncontinuous) continuity control functions. In the context of computational vision, they may be thought of as controlled-continuity constraints. These generic constraints are applicable to visual reconstruction problems that involve both continuous regions and discontinuities, for which global smoothness constraints fail.**

Index Terms—**Controlled-continuity constraints, discontinuities, early vision, inverse problems, reconstruction, regularization.**

I. Introduction

THE well-known rendering problem in the field of computer graphics involves the synthesis of images from explicit representations of 3-D forms. Conversely, early computational vision aims at understanding how explicit geometric representations of the 3-D world may be reconstructed from 2-D images [3], [24]. Visual reconstruction leads to *inverse* mathematical problems. Consequently, it presents formidable theoretical and pragmatic challenges.

Early vision has traditionally been regarded as an array of specialized reconstruction processes operating on images. They include the reconstruction of 2-D image intensity gradient and flow fields, as well as the reconstruction of 3-D surface depth, orientation, and motion fields. A broad range of visual reconstruction problems may be unified mathematically as well-posed variational principles which characterize optimal approximation problems involving a class of generalized multidimensional spline functionals [38, Section 6]. These functionals embody generic smoothness constraints.

The essence of such a unification may be rationalized from first principles. As is generally the case for inverse mathematical problems, visual reconstruction problems tend to be ill-posed in that existence, uniqueness, and stability of solutions cannot be guaranteed in the absence of additional constraints. In this regard, constraints such as smoothness have been useful expressions of generic, *a priori* information about possible solutions.

The utilization of smoothness constraints in vision has often received ad hoc justification that is implicitly based on computational convenience rather than theoretical considerations. However, a formal basis for similar constraints can be found in certain systematic approaches to the mathematical solution of ill-posed inverse problems, particularly the regularization methods pioneered by Tikhonov [44] and others (see [45] and [28] and references therein). A very basic form of regularization, namely, spatial smoothing to suppress detrimental high-frequency noise, is mentioned by Duda and Hart in connection with image restoration, a familiar inverse visual problem [13, Section 7.4]. A broader perspective of regularization theory, as it impacts on early vision, is presented by Poggio and Torre in their review of the Tikhonov regularization approach [31].

Through regularization, a wide range of ill-posed visual reconstruction problems may be reformulated as variational principles. Tikhonov regularization employs a specific class of so-called "stabilizing functionals" to restrict admissible solutions to spaces of smooth functions. Under nonrestrictive conditions, the resulting variational principles can be made well-posed within these spaces; hence, their solutions are effectively computable. Regularization therefore appears to offer a theoretical basis for the smoothness constraints that have been applied to the reconstruction problems of early vision (see [38, Section 6] and [31]).

Smoothness constraints have physical validity inasmuch as the coherence of matter tends to give rise to smoothly varying instrinsic scene characteristics relative to the viewing distance, over some range of spatial resolution. However, most significant, spatially localized physical transitions, such as abrupt changes in surface geometry (e.g., occlusions), surface composition (e.g., tex-

Manuscript received December 20, 1984; revised August 30, 1985. Recommended for acceptance by S. L. Tanimoto. This paper describes research done at the Artificial Intelligence Laboratory of the Massachusetts Institute of Technology. Support for the Laboratory's Artificial Intelligence research 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 and the System Development Foundation. The author was supported by the Natural Sciences and Engineering Research Council of Canada, Ottawa, and the Fonds F.C.A.C., Quebec, Canada.

The author is with Schlumberger Palo Alto Research, 3340 Hillview Ave., Palo Alto, CA 94304.

IEEE Log Number 8406052.

ture), or illumination (e.g., shadows), lead to discontinuities in intrinsic scene characteristics, some of which persist across all scales. Thus, discontinuities inevitably play an important role in early visual reconstruction [40].

Standard Tikhonov regularization theory encounters serious difficulties in application to real life visual problems, since smoothness assumptions clearly do not hold indiscriminately across visual discontinuities. The difficulty is due to the fact that the quadratic stabilizer functionals prescribed in standard regularization offer no spatial control over their smoothness properties.

As an instance of standard regularization, the early work of Terzopoulos [38], [39] in visual surface interpolation (see also [17]) provides vivid demonstrations of the deficiency of global smoothness constraints—globally smooth surface interpolation (with quadratic "thin plate" stabilizers) destroys surface occlusions and creases (i.e., depth and orientation discontinuities). It was clearly necessary to overcome this deficiency in order to compute visible-surface representations [40]. The computational framework developed in [41] and [42] employs a nonquadratic and spatially noninvariant stabilizer (the thin plate surface under tension model) to regularize an ill-posed surface reconstruction problem whose basic mathematical statement is reviewed in Appendix A.

Expanding on the underlying motivation for this nonstandard stabilizer, the need to preserve surface depth and orientation discontinuities, the present paper proposes a general variational approach to visual reconstruction of arbitrary dimensionality that can accommodate visual discontinuities of arbitrary orders. A class of multidimensional *controlled-continuity stabilizers*, composed of generalized spline kernels, is suggested for the regularization of inverse visual problems. By adjusting a constituent set of parametric weighting functions, the continuity properties of these nonquadratic stabilizers can be controlled with spatial precision to reconstruct localized discontinuities.

One can view these more sophisticated stabilizers as representing *controlled-continuity constraints*. Not only do controlled-continuity constraints retain the mathematical elegance and potency of traditional smoothness constraints as applied to continuous problems, but they extend to the reconstruction of visual discontinuities as well where global smoothness constraints fail.

II. Standard Tikhonov Regularization

To better understand Tikhonov regularization, consider the textbook regularization example of reliably estimating point derivatives $v_x(x_i)$ of a function $v(x)$ given only certain approximate samples v_i of $v(x_i)$ for $x_i = ih$, $i = 1, \cdots, N$; $Nh = 1$ [2, Section 5.4]. The errors $\{\epsilon_i = v(x_i) - v_i\}_{i=1}^N$ of the approximate data $\{x_i, v_i\}_{i=1}^N$ are assumed to be independent, normally distributed random variables with zero mean and variance σ^2.

Difficulties arise because no matter how small the errors ϵ_i, the differences between the true point derivatives $v_x(x_i)$ and the numerical derivatives of the data, say $(v_i - v_{i-1})/h$, can be arbitrarily large [10]. Since one cannot guarantee that the solution will be stable with respect to small perturbations of the data, numerical differentiation, unlike numerical integration, is an ill-posed problem.[1] The above problem may therefore be approached through regularization.

The common form of regularization for the numerical differentiation problem seeks an approximating function $u(x)$ which minimizes the functional

$$\mathcal{E}(v) = \int_0^1 v_{xx}^2 \, dx + \frac{1}{\lambda\sigma^2} \sum_{i=1}^{N} [v(x_i) - v_i]^2,$$

where the subscripts on v denote partial derivatives and λ is a nonnegative regularization parameter [2, Sections 5.3, 5.4]. The desired estimates for the derivatives are obtained as $u_x(x_i)$, and these estimates are robust against perturbations in the data. When $\mathcal{E}(u)$ is small, $u(x)$ is a good compromise between smoothness in the sense of minimal stabilizing functional $\int_0^1 v_{xx}^2 \, dx$ and adhesion to the data in the minimal sum of squared deviations sense. The above problem and its variants are relevant to edge detection in images since robust numerical differentiation is an issue in this context [32].

Tikhonov proposed a general stabilizer for univariate regularization, the pth-order weighted Sobolev norm

$$\|v\|_p^2 = \sum_{m=0}^{p} \int_{\mathcal{R}} w_m(x) \left(\frac{d^m v(x)}{dx^m}\right)^2 dx$$

where the $w_m(x)$ are prespecified, nonnegative, and continuous weighting functions [45, pp. 69–70]. The stabilizer in the above example is a particular instance of $\|v\|_p^2$ for $p = 2$, $w_0 = w_1 = 0$, and $w_2 = 1$.

Tikhonov's stabilizer can be viewed as characterizing univariate splines that impose smoothness on the admissible solutions by restricting them to Sobolev spaces of smooth functions. In fact, minimizing $\mathcal{E}(v)$, as given above, is the variational formulation of the smoothing cubic spline problem considered by Reinsch [33]. Tikhonov regularization in this simple case is therefore equivalent to fitting a common univariate cubic spline to the data, and then using the approximating spline to robustly estimate the desired derivatives.

III. Multivariate Regularization as Optimal Approximation

Although univariate problems appear in conjunction with image contours, the bulk of visual reconstruction problems are multivariate. Two-dimensional problems are commonplace in the reconstruction of retinotopic representations of physical scene characteristics, and a rich variety of higher dimensional problems arise if time is incorporated as an ordinary geometric coordinate to supplement the standard spatial coordinates. The formulation in this paper emphasizes the aforementioned correspondence between regularization and optimal spline

[1] The surface reconstruction problem discussed in Appendix A is ill-posed, not merely due to lack of stability, but moreover, in that the solution cannot be guaranteed *a priori* to exist nor to be unique.

approximation and extends uniformly over any number of dimensions. The abstract theory of optimal approximation is well developed and a close connection has been established with variational principles involving the constrained minimization of (semi-)norms in (semi-)Hilbert function spaces [23].

Let $\mathcal{H}$ be a linear, admissible space of smooth functions defined on $\mathcal{R}^d$. Let $\mathcal{S}(v)$: $\mathcal{H} \mapsto \mathcal{R}$ be a functional defined on $\mathcal{H}$ which measures the smoothness of an admissible function $v(x) \in \mathcal{H}$ where $x = [x_1, \cdots, x_d] \in \mathcal{R}^d$. Furthermore, let $\mathcal{P}(v)$: $\mathcal{H} \mapsto \mathcal{R}$ be a functional on $\mathcal{H}$ which measures the discrepancy between the function and given data.

Consider the following variational principle.

VP: Find $u \in \mathcal{H}$ such that

$$\mathcal{E}(u) = \inf_{v \in \mathcal{H}} \mathcal{E}(v),$$

where the functional

$$\mathcal{E}(v) = \mathcal{S}(v) + \mathcal{P}(v).$$

This defines an optimal approximation problem—to find the smoothest admissible function in $\mathcal{H}$ which is most compatible with the data. For instance, in the case of visible-surface reconstruction, u represents a reconstructed depth function (Appendix A).

The necessary (but, in general, not sufficient) condition satisfied by the minimizing function $u(x)$ is given by the vanishing of the first variation δ, which expresses the well known Euler–Lagrange equations $\delta\mathcal{E}(u) = \delta\mathcal{S}(u) + \delta\mathcal{P}(u) = 0$.

A. Penalty Functionals

The role of $\mathcal{P}(v)$ is to "penalize" the discrepancy between admissible functions and given data.

Assuming independent, normally distributed measurement errors ϵ_i with zero means and variances σ_i^2, the optimal $\mathcal{P}(v)$ is a weighted Euclidean norm of the discrepancy between the admissible function and the data c_i. This can be written as

$$\mathcal{P}(v) = \sum_{i=1}^{N} \alpha_i (\mathcal{L}_i(v) - c_i)^2,$$

where the $\mathcal{L}_i$ are measurement functionals and the α_i are nonnegative real-valued weights. Ideally, $\alpha_i = 1/\lambda\sigma_i^2$ (see Section VI-D), but this penalty functional is still useful when the above assumptions do not hold in the strict sense.

For visible-surface reconstruction (see Appendix A) and for other visual problems, the basic measurement functionals of immediate concern are generalized point derivatives of the form

$$\mathcal{L}_i(v) = \mathcal{D}^k_{x_i}(v) = \left.\frac{\partial^k v}{\partial x_1^{j_1} \cdots \partial x_d^{j_d}}\right|_{x_i} \qquad j_1 + \cdots + j_d = k.$$

Note that for $k = 0$, this reduces to the evaluation functional $\mathcal{L}_i(v) = v(x_i)$. It is also possible to incorporate integral functionals of the form $\mathcal{L}_i(v) = \int_{\mathcal{R}^d} K_i(x)\, v(x)\, dx$.

IV. Multivariate Generalized Spline Functionals

To formulate basic multivariate smoothness constraints for visual regularization, we turn to spline approximation. Duchon [11], [12] and Meinguet [26], [27] study the interesting properties of the following class of functionals:

$$|v|_m^2 = \sum_{i_1, \cdots, i_m = 1}^{d} \int_{\mathcal{R}^d} \left(\frac{\partial^m v(x)}{\partial x_{i_1} \cdots \partial x_{i_m}}\right)^2 dx$$

$$= \int_{\mathcal{R}^d} \sum_{j_1 + \cdots + j_d = m} \frac{m!}{j_1! \cdots j_d!} \cdot \left(\frac{\partial^m v(x)}{\partial x_1^{j_1} \cdots \partial x_d^{j_d}}\right)^2 dx$$

defined on d-dimensional functions $v(x)$, $x = [x_1, \cdots, x_d]$.[2]

The positive integer m dictates the order of the partial derivatives that occur in the functional, which in turn determines the order of continuity that the admissible functions v must possess. Since in a single dimension ($d = 1$) the variational principle VP with $\mathcal{S}(v) = |v|_m^2$ reduces to the approximation problem associated with classical smoothing splines which can be generated with piecewise polynomials of degree $2m - 1$ [1], the functionals $|v|_m$ can be viewed as generating multivariate generalized splines.

Generalized spline functionals are invariant under (i.e., commute with) translation, rotation, and similarity transformations of the data in $\mathcal{R}^d$. Such invariance properties are essential in the context of visual reconstruction problems since, e.g., the solutions to visual reconstruction problems should not change shape when objects in the scene translate or rotate parallel to the image plane or when they approach or retreat parallel to the view direction [38], [6].

The Euler–Lagrange equations characterizing functions $u(x)$ which minimize generalized spline functionals are partial differential equations involving iterated Laplacians of $u(x)$. The reader is referred to Appendix B for a more detailed discussion of this and other mathematical properties of generalized spline approximation problems.

A. Special Cases: Two-Dimensional Surface Splines

In two dimensions ($d = 2$), the generalized spline functionals can be written as

$$|v|_m^2 = \int\int_{\mathcal{R}^2} \sum_{i=0}^{m} \binom{m}{i} \left(\frac{\partial^m v}{\partial x^i \partial y^{m-i}}\right)^2 dx\, dy.$$

The two-dimensional case pertains directly to the problem of fitting surfaces to scattered data. This mathematical problem is of considerable concern in numerous applica-

[2]These scalar-valued spline functionals can be extended in a natural way to vector-valued functionals defined on $v(x) = [v_1(x), \cdots, v_n(x)]$ by replacing the partial derivative expressions with $|\partial^m v(x)/\partial x_1^{j_1} \cdots \partial x_d^{j_d}|^2$.

tion areas [35] and, notably in vision, to visible-surface reconstruction.

Surface splines have interesting physical interpretations involving equilibria of elastic bodies with C^{m-1} intrinsic continuity. The two lowest order cases are of particular interest.

For $m = 1$, the functional reduces to

$$|v|_1^2 = \int\int_{\Re^2} (v_x^2 + v_y^2)\, dx\, dy,$$

which is proportional to the small deflection energy of a membrane (e.g., rubber sheet) [9].[3] The associated Euler–Lagrange equation is Laplace's equation, $-\Delta u = 0$ where $\Delta u = u_{xx} + u_{yy}$.

For $m = 2$,

$$|v|_2^2 = \int\int_{\Re^2} (v_{xx}^2 + 2v_{xy}^2 + v_{yy}^2)\, dx\, dy$$

is proportional to the small deflection bending energy of a thin plate (with zero Poisson ratio) [9]. The associated Euler–Lagrange equation is the biharmonic equation $\Delta^2 u = 0$ where $\Delta^2 u = u_{xxxx} + 2u_{xxyy} + u_{yyyy}$. Duchon [11] refers to the minimizers of $|v|_2^2$ as "thin plate splines."

Physically, the membrane spline characterizes a surface of C^0 continuity, a continuous surface which, however, need not have continuous first (and higher) order partial derivatives. The thin plate spline is a C^1 surface, a continuous surface with continuous first partial derivatives, which need not have continuous derivatives of degree greater than one.

As the natural 2-D extensions to the cubic spline, the thin plate spline $|v|_2^2$ is a popular surface interpolant [35], [14]. It has been employed in surface fitting problems such as the interpolation of airfoils [18], the interpolation of digital terrain maps [7], the interpolation of meteorological fields [46], and in visual surface interpolation [17], [6], [39].

V. Controlled-Continuity Stabilizers

The generalized spline functionals are quadratic functionals which specify global smoothness constraints. As stabilizers in regularization, they do not apply to reconstruction problems involving discontinuities. A convenient way to control the continuity properties of a generalized spline stabilizer of order m to accommodate this wider class of problems is to blend it with generalized splines of lower orders. This leads to *controlled-continuity constraints*.

We propose the following class of stabilizing functional for piecewise continuous regularization:

$$|v|_{p,w}^2 = \sum_{m=0}^{p} \int_{\Re^d} w_m(\boldsymbol{x}) \sum_{j_1 + \cdots + j_d = m} \cdot \frac{m!}{j_1! \cdots j_d!} \left(\frac{\partial^m v(\boldsymbol{x})}{\partial x_1^{j_1} \cdots \partial x_d^{j_d}} \right)^2 d\boldsymbol{x}.$$

[3]Minimization of the true surface area leads to the famous Plateau's problem.

The positive integer p is the highest order generalized spline that occurs in the functional, and this determines the maximum order of continuity (C^{p-1}) of the admissible functions v. The weighting functions $\boldsymbol{w}(\boldsymbol{x}) = [w_0(\boldsymbol{x}), \cdots, w_p(\boldsymbol{x})]$ are nonnegative, not necessarily continuous, and not prespecified, in general.

The weighting functions will be allowed over $\Re^d$ to make jump transitions to zero values. This provides the capability of selectively introducing specific discontinuities into the solution, a crucial property that will be demonstrated shortly with examples. The possibility of exercising such precise spatial control over the regularized solution suggests the terms *controlled-continuity stabilizers* for the above class of functionals and *continuity control functions* for the weighting functions.

Euler–Lagrange equations characterizing functions $u(\boldsymbol{x})$ which minimize controlled-continuity stabilizers involve spatially weighted iterated Laplacians of $u(\boldsymbol{x})$. These nonlinear partial differential equations are described in Appendix C.

A. Relationship to Tikhonov Stabilizers

Restricting to the 1-D case, and assuming prespecified and continuously varying weighting functions $w_m(x)$, the controlled-continuity stabilizers reduce to Tikhonov's stabilizers

$$\|v\|_p^2 = \sum_{m=0}^{p} \int_{\Re} w_m(x) \left(\frac{d^m v(x)}{dx^m} \right)^2 dx.$$

For positive $w_m(x)$, Tikhonov stabilizers are, in fact, pth-order weighted versions of Sobolev norms; hence, the admissible spaces $\mathcal{H}$ in which standard regularized formulations to visual reconstruction problems are naturally set are Sobolev spaces of globally smooth functions [38], [39].

Tikhonov stabilizers are quadratic, and define linear, spatially invariant regularization methods. In contrast, the controlled-continuity stabilizers, which are generally nonquadratic, define a class of nonlinear and spatially noninvariant regularization methods.

B. Special Cases: Splines and Surfaces Under Tension

Schweikert [37] introduced splines under tension which can be made to imitate the behavior of cubic interpolating splines, while suppressing the extraneous inflection points that sometimes afflict cubic splines (see also [8]). They may be characterized as interpolatory functions $u(x)$ that minimize the functional

$$\int_{\Re} (\tau v_x^2 + v_{xx}^2)\, dx$$

where τ is a prespecified positive constant, called the tension. The first term influences the "length" of the spline, while the second term influences its "curvature." Increasing the tension tends to eliminate extraneous loops and ripples by reducing the length of the spline.

The spline under tension is a restricted controlled-continuity stabilizer for $d = 1$, $p = 2$, $w(x) = [0, \tau, 1]$. Because the weighting functions are continuous and prespecified, the spline under tension is also a quadratic stabilizer, and it results in linear regularization.

A particular instance of a nonquadratic controlled-continuity stabilizer has been proposed for the regularization of the visible-surface reconstruction problem (Appendix A), a $d = 2$ dimensional problem that requires a C^1 continuous surface spline [40]–[42]. This is provided by the $p = 2$ order stabilizer

$$|v|^2_{2,w} = \mathbb{S}_{\rho\tau}(v) = \iint_{\mathbb{R}^2} \rho(x, y) \cdot \{[1 - \tau(x, y)] (v_x^2 + v_y^2) + \tau(x, y) (v_{xx}^2 + 2v_{xy}^2 + v_{yy}^2)\} \, dx \, dy$$

with $w(x, y) = [0, \rho(x, y) [1 - \tau(x, y)], \rho(x, y) \tau(x, y)]$. The range of $\rho(x, y)$ and $\tau(x, y)$ is assumed to be $[0, 1]$, and they need not vary continuously.

This functional, the natural 2-D extension of the spline under tension functional, will be recognized as a weighted convex combination of the thin plate and membrane generalized spline kernels. The thin plate kernel affects surface "curvature," while the membrane kernel affects surface "area." The continuity control functions $\rho(x, y)$ and $[1 - \tau(x, y)]$ may be interpreted physically as spatially varying surface cohesion and surface tension, respectively. In view of the physical interpretation, solutions to the variational principle involving this functional can be thought of as *thin plate surfaces under tension*.

The above controlled-continuity stabilizer provides the necessary spatial control over its smoothness properties to explicitly reconstruct both depth and orientation discontinuities. The stabilizer is controlled as follows. In continuous regions, $\rho(x, y) = \tau(x, y) = 1$ so that the stabilizer reduces to a thin plate spline and generates a C^1 surface. Along orientation discontinuities, $\rho(x, y) = 1$ and $\tau(x, y) = 0$, i.e., maximum tension is applied so that the stabilizer reduces locally to a membrane functional, thus maintaining only C^0 continuity and allowing the surface to crease freely. Along depth discontinuities, $\rho(x, y) = 0$, thus deactivating all continuity and thereby allowing the surface to fracture freely. Fig. 1 shows examples of reconstructed surfaces and discontinuities using the thin plate surface under tension stabilizer. The details of these and many more examples are presented in [41] and [42].

Higher order controlled-continuity stabilizers can be manipulated analogously to reconstruct smoother surfaces and a wider range of discontinuities. Consider the case $d = 2$, $p = 3$, and $w_0(x, y) = 0$. With the $w_m(x, y)$ positive functions, a C^2 (i.e., continuous curvature) surface results. However, curvature discontinuities can be reconstructed in this surface by setting $w_3(x, y) = 0$ along such discontinuities in the x, y plane. This locally reduces the functional to a C^1 thin plate spline which need not have continuous curvature. Further, it is possible to set $w_3(x, y) = w_2(x, y) = 0$ along orientation discontinuities which

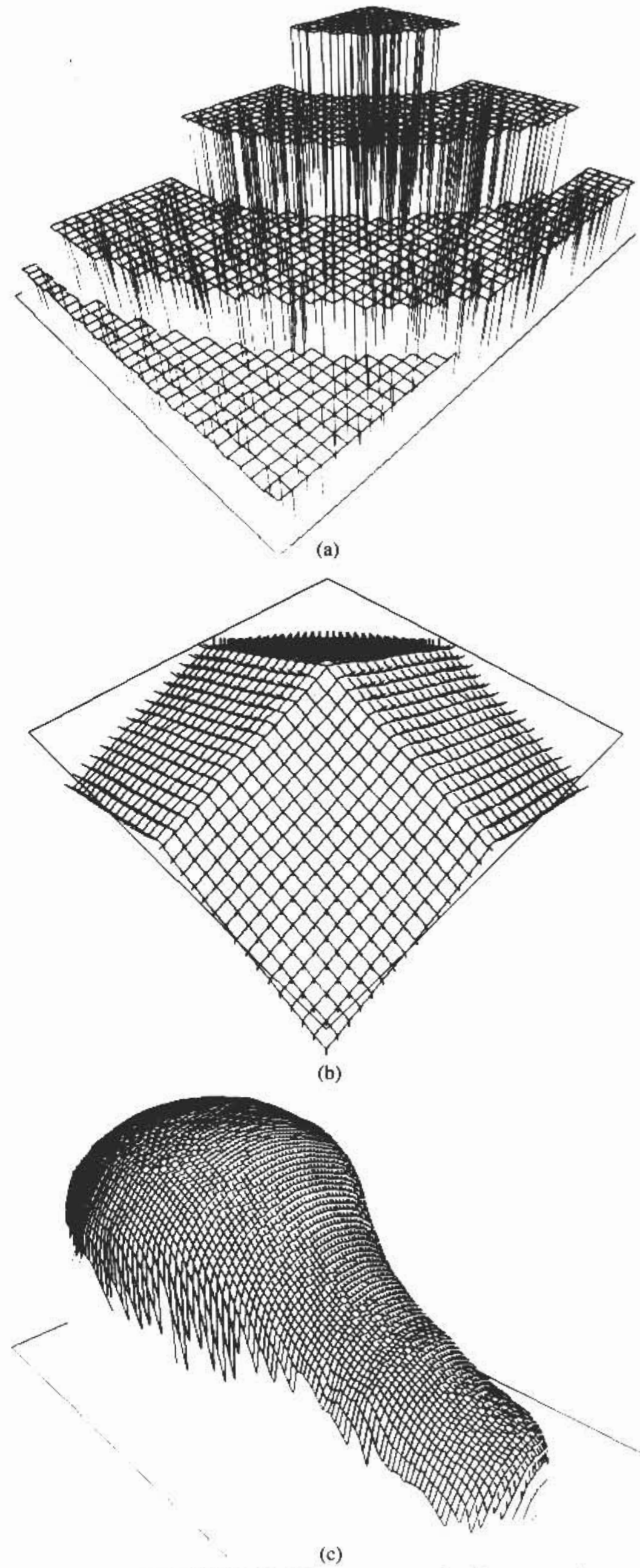

Fig. 1. Surfaces and discontinuities reconstructed with a controlled-continuity stabilizer. (a) Reconstruction of planar surfaces from scattered, synthetic depth data shown as vertical lines. Surface boundaries are reconstructed depth discontinuities. (b) Reconstruction of a surface from synthetic orientation data shown as needles on the surface. Reconstructed orientation discontinuities occur between the faces and depth discontinuities bound the surface. (c) Reconstruction of a light bulb from structured light, range data. The occluding contour is a reconstructed depth discontinuity.

must be reconstructed. Finally, all three continuity control functions can be set to 0 along depth discontinuities that are to be reconstructed.

VI. Discussion

A. Related Geometric Models

The controlled-continuity stabilizers proposed in this paper are related to certain geometric modeling primitives developed in approximation theory and, more recently, in computer graphics.

Schweikert's splines under tension have already been mentioned [37], [8]. The tension in these splines is a global parameter, however, and does not provide local control over shape. The basis functions for splines under tension involve exponentials. Nielson [29] developed *ν splines*, piecewise polynomial alternatives, one of whose advantages is that the tension can be set selectively at each interpolation point.

Pilcher [30] suggested that the idea of splines under tension can be carried over to surfaces. His proposal involved restricting the area of a tensor product of polynomial splines, and he characterized this as a minimization problem. Using tensor products of lower dimensional models to obtain a higher dimensional one is a common approach. This, however, restricts the method's applicability to regular data.

The idea of locally controlling the shape of a continuous curve or surface also appears in the context of piecewise curve and surface models, such as the *B*-spline models [16], notably in the work of Barsky, whose Beta-spline model contains continuously varying bias and tension parameters [4]. The data must be specified at set of control vertices defining surface patches. These vertices are restricted to a regular lattice. The Beta splines are piecewise models; changes within a surface patch have no effect on surface patches further than a certain distance away.

As we have seen, the abstract theory of spline approximation leads to multidimensional generalizations of the classical univariate splines [34], [1], not as tensor products, but rather through physical interpretations concerning equilibria of elastic bodies. Hence, the controlled-continuity stabilizers can be viewed as characterizing global geometric models which place no restrictions on constraint arrangement and provide local continuity control, particularly the capability of introducing arbitrary discontinuities. This is of paramount importance in visual reconstruction where we must routinely deal with irregularly located data comprising both constraints and discontinuities.

B. Computational Considerations

A potent approach to solving variational principles founded on controlled-continuity constraints is to compute a solution as a superposition of basis functions. The choice of basis, particularly whether it comprises local or global support functions, is an important consideration. It depends on the number of constraints and the nature of the discontinuities (either prespecified or to be determined as part of the reconstruction). Moreover, the particular choice dictates the class of algorithm most suited to computing the unknown coefficients of the superposition, a critical issue in computational vision.

A global method for computing generalized splines pursued by Duchon [11] (see also [18]) and developed further in [27] and [46] involves a representation of the solution as a linear combination of N rotationally symmetric, global support basis functions, one centered at each data point (see Appendix D). The basis functions of choice are the fundamental solutions of the iterated Laplacian appearing in the Euler–Lagrange equations associated with the generalized splines. They are no more complicated than logarithms. Computing the solution requires first solving a system of linear equations for the unknown coefficients of the linear combination, and then constructing the superposition of basis functions, restricted to a compact region of interest Ω in which the solution is continuous. The matrix of the linear system, whose size depends on the number of constraints N, is positive definite, symmetric, and full. It can be solved by Cholesky factorization with back substitution, a procedure whose efficiency compares favorably to other direct methods [26], [27].

Local support approximation schemes offer an alternative to computing a global representation. The finite element method is a local scheme that is systematically applicable to the variational principles arising from the regularization of inverse visual problems [38], [39]. The finite element approximation is a linear combination of the local support basis functions (typically low-order piecewise polynomials) of finite element spaces, which may be subspaces of the admissible space $\mathcal{H}$. The number of basis functions depends on the number of finite elements employed to tessellate the continuous domain. The choice of tessellation is very flexible, but a natural tessellation in visual applications follows the image sampling pattern [38], [39]. The finite element representation of generalized spline stabilizers also leads to a positive definite, symmetric system of linear equations that must be solved for the unknown coefficients. While the size of the linear system is usually greater than N, its matrix is sparse due to the local support of the basis functions. This suggests the use of either direct or iterative sparse matrix techniques to compute the solution.

In principle, the global method can extend to controlled-continuity stabilizers provided the continuity control functions are prespecified. For instance, it is possible to construct, in piecewise fashion, solutions restricted to continuous regions from constraints contained within such regions. Restricting the global support basis functions to arbitrary regions is awkward, however, especially when the regions are not known in advance and are irregular in shape. Furthermore, local bases are much better suited to computing continuity control functions containing jump transitions to zero since discontinuities may readily emerge between adjacent elements [41], [42].

To construct the global solution, the Cholesky algorithm computes a sequence of optimal approximations, as each data point is accessed in sequence. This has the advantage that the addition or removal of a single data point requires relatively little computation to determine the new global solution, given the current approximation. Its disadvantage is that the serial, recursive nature of the algorithm appears incompatible with the large-scale parallelism that is characteristic of visual processing. On the other hand, local iterative schemes such as relaxation methods of the Jacobi type for solving finite element systems are readily parallelizable.

Although the global method may be attractive for problems involving a modest number of constraints, it becomes expensive to store and solve an $N \times N$ linear system with full matrix if N exceeds order 10^2–10^3 or so. In addition, the system tends to become ill conditioned for large N. Since early visual processing typically generates a larger number of constraints (on the order of 10^4–10^5 or more is not uncommon), the computational expense can become prohibitive in some machine vision applications. In contrast, sparse matrix techniques for solving finite element equations need store only the nonzero matrix entries and relaxation methods operate on these entries in place. Recently, highly efficient multigrid relaxation methods have been developed in applied numerical analysis. They have been adapted successfully to the solution of finite element systems arising from visual regularization [38], [39], [43]. This class of algorithm was used to compute the examples of Fig. 1.

One final comment. Since the effort required to obtain solutions increases sharply with growth in dimensionality d and order p of the stabilizer, it is prudent to use the smallest possible dimension and order needed to regularize the problem at hand.

C. Discontinuity Detection with Controlled-Continuity Stabilizers

We explained how discontinuities of order less than p may be introduced into the reconstructed solution $u(\boldsymbol{x})$ by regulating pth-order controlled-continuity stabilizers.[4] More interestingly, it is also possible to reconstruct discontinuities, which are not known in advance, as an integral part of controlled-continuity reconstruction. Mathematically, this poses a nonlinear, distributed parameter identification problem. For example, a discontinuity estimation coprocess can detect and localize discontinuities as the regularized solution is being computed [41].

It is possible to augment the energy functional in the variational principle characterizing the solution $u(\boldsymbol{x})$ to optimally reconstruct the continuity control functions $\boldsymbol{w}(\boldsymbol{x})$ according to regularization criteria. The augmented functional takes the form

$$\mathcal{E}(v, \boldsymbol{w}) = \mathcal{S}(v, \boldsymbol{w}) + \mathcal{P}(v) + \mathcal{D}(\boldsymbol{w})$$

where $\mathcal{D}(\boldsymbol{w})$ imposes criteria for good discontinuity structure. The necessary condition for the solution $\{u, \boldsymbol{w}^*\}$ to this discontinuity identification problem expressed by the system of Euler–Lagrange equations

$$\delta_u \mathcal{E}(v, \boldsymbol{w}^*) = 0; \qquad \delta_w \mathcal{E}(v, \boldsymbol{w}^*) = 0.$$

The discontinuity stabilizing functional $\mathcal{D}(\boldsymbol{w})$ is an important consideration. At the very least, it should independently penalize each detected discontinuity; otherwise, an incoherent solution will result [5]. A more sophisticated version can be implemented as a "lookup table" ranking local discontinuity configurations according to rather crude "good continuation" constraints [48], [15], [25], [42], [22]. Better yet, functionals that formally impose controlled-continuity constraints can be constructed from controlled-continuity stabilizers.[5]

The complexity of the discontinuity identification problem is exacerbated by the nonquadratic nature of $\mathcal{E}(v, \boldsymbol{w})$ or, equivalently, the nonlinearity of the Euler–Lagrange system. A direct solution requires the minimization of a nonconvex energy functional having multiple local minima.

Iterative methods can be adapted to perform the minimization. Given a sufficient number of steps, the global minimum can be computed stochastically by simulated annealing [21], [15], [25]. This, however, involves a great deal of computation in practice. A deterministic "graduated nonconvexity" method for finding near optima is described in [5], but the formulation requires dense data.

Alternatively, one can minimize the full, multiargument energy functional as a chain of simpler energy functionals over a number of stages, a useful strategy for nonlinear optimization. For fixed $\boldsymbol{w}(\boldsymbol{x})$, the energy functional becomes quadratic. Hence, treating the continuity control functions as fixed parameters during each stage results in a convex minimization problem for $v(\boldsymbol{x})$, which can be solved by (multigrid) relaxation [41]. At the start of each stage, improved parameter values are computed—either discretely [42] or continuously [22] (see also [19])—from the minimum obtained in the previous stage. Final solutions near the global minimum $\{u(\boldsymbol{x}), \boldsymbol{w}^*(\boldsymbol{x})\}$ evolve over several stages of this relatively efficient, nonlinear dynamic process [42].

D. Correspondence with Bayesian Estimation on Stochastic Processes

The notion of smoothness in visual regularization is intimately connected to prior expectations about the physical world and the underlying processing of low-level visual information. Such prior expectations are emphasized by Bayesian estimation theory. Kimeldorf and Wahba [20] note that there exist stochastic interpretations of spline approximation in which the smoothness properties of splines

[4]Generally speaking, in the context of d-dimensional pth-order controlled-continuity regularization, a visual discontinuity is a piecewise smooth $d-1$ dimensional hypersurface embedded in $\Re^d$. On such a hypersurface, there occurs a discontinuous transition in $d^k u(\boldsymbol{x})/dn^k$ for some $0 \le k < p$ where n is the direction normal to the hypersurface.

[5]Visual discontinuities are recursive in nature, so that $d-1$ dimensional discontinuity hypersurfaces may contain embedded discontinuity subsurfaces of order less than $d-1$. Controlled-continuity stabilizers of dimension $d-1$ can thus be employed to recursively characterize the structure of the jump transitions to zero in the d-dimensional continuity control functions $\boldsymbol{w}(\boldsymbol{x})$.

correspond to the specification of prior probability measures over suitable function spaces.

Consider the problem of estimating a real random function $v(\boldsymbol{x})$ based on prior information and a finite set of fixed observations comprising constraints $\{\boldsymbol{x}_i, c_i\}_{i=1}^N$ having normally distributed errors $\{\epsilon_i = v(\boldsymbol{x}_i) - c_i\}_{i=1}^N$, assumed independent of $v(\boldsymbol{x})$, with zero means and variances $\lambda\sigma_i^2$. One can assume for simplicity and without loss of generality that the observations have zero mean. The prior information is represented by a prior distribution on $v(\boldsymbol{x})$ that is Gaussian with zero mean and covariance function $K(\boldsymbol{s}, \boldsymbol{t})$.

Following [20], it can be shown that, for fixed $\boldsymbol{x}$, the estimate $u(\boldsymbol{x}) = E[v(\boldsymbol{x})|c_1, \cdots, c_N]$, the mean of the marginal conditional distribution of the unknown $v(\boldsymbol{x})$, given the observations, may be written as

$$u(\boldsymbol{x}) = [c_1, \cdots, c_N]\,(\boldsymbol{K} + \lambda\Sigma^2)^{-1} \cdot [K(\boldsymbol{x}_1, \boldsymbol{x}), \cdots, K(\boldsymbol{x}_N, \boldsymbol{x})]^T$$

where the $N \times N$ matrix $\boldsymbol{K} = [K(\boldsymbol{x}_i, \boldsymbol{x}_j)]$ and the $N \times N$ diagonal matrix $\Sigma = [\sigma_i]$.

However, the Bayesian estimate $u(\boldsymbol{x})$ corresponds with the solution to the variational principle resulting from regularization using generalized spline functionals, i.e., the minimizer of

$$\mathcal{E}(v) = |v|_m^2 + \sum_{i=1}^{N} \frac{1}{\lambda\sigma_i^2}\,[v(\boldsymbol{x}_i) - c_i]^2.$$

The covariance function $K(\boldsymbol{s}, \boldsymbol{t})$ corresponds to the reproducing kernel associated with the reproducing kernel Hilbert space in which the variational principle is set (see Appendix D).

Roughly speaking, the above correspondence extends to the case of controlled-continuity stabilizers. However, an improper prior distribution is implied in which $v(\boldsymbol{x})$, for fixed $\boldsymbol{x}$, has infinite prior variance. Moreover, the prior distribution will not be wide-sense stationary, particularly when the continuity control functions contain jump transitions to zero, leading to discontinuities in $u(\boldsymbol{x})$.

E. Correspondence with Filtering

The Bayesian estimation correspondence suggests that $u(\boldsymbol{x})$ can also be viewed as a solution to a filtering prediction problem. Indeed, for the case of continuous data $c(\boldsymbol{x})$ on $\mathcal{R}^d$, the generalized spline functionals correspond to linear convolution filters. To see this, consider the Euler–Lagrange equation associated with the variational principle (Appendix B) for the special case $\alpha(\boldsymbol{x}) = \alpha$, a constant, and $k = 0$ (note that $\mathcal{D}^0 u = u$):

$$(-1)^m \Delta^m u(\boldsymbol{x}) + \alpha u(\boldsymbol{x}) = \alpha c(\boldsymbol{x}).$$

A Fourier transformation of the Euler–Lagrange equation leads to the expression

$$U(\boldsymbol{\omega}) = \left(1 + \frac{1}{\alpha}\,|\boldsymbol{\omega}|^{2m}\right)^{-1} C(\boldsymbol{\omega})$$

where $\boldsymbol{\omega} = [\omega_1, \cdots, \omega_d]$ is the d-dimensional frequency domain variable and where $U(\boldsymbol{\omega})$ and $C(\boldsymbol{\omega})$ denote the Fourier transforms of $u(\boldsymbol{x})$ and $c(\boldsymbol{x})$, respectively.

The transfer function in brackets will be recognized as having the characteristic of a doubly cascaded, d-dimensional, mth-order Butterworth low-pass filter. Letting $f = |\boldsymbol{\omega}|/\sqrt[2m]{\alpha}$, the frequency radius of the filter's half-amplification band is given by $f = 1$, while the attenuation in gain for $f \gg 1$ is $-40m \log_{10} f$ dB. Therefore, the effective half-amplification radius is determined (reciprocally) by the regularization parameter λ where $\alpha = \lambda^{-2m}$, while the effective rolloff is determined by the order m of the generalized spline.

An analogous convolution filter interpretation is possible for the controlled-continuity stabilizers under similar restrictive conditions. For instance, in the above special case and with constant $w_m(\boldsymbol{x}) = w_m$, one readily obtains

$$U(\boldsymbol{\omega}) = \left(1 + \frac{1}{\alpha}\sum_{m=0}^{p} w_m|\boldsymbol{\omega}|^{2m}\right)^{-1} C(\boldsymbol{\omega}).$$

Although the correspondence with linear convolution filters may readily be shown to extend to the case of regularly sampled data over $\mathcal{R}^d$ (or over a compact region Ω^d with periodic continuation on the exterior), it is important to realize that the convolution property does not hold for irregularly scattered data $\{\boldsymbol{x}_i, c_i\}_{i=1}^N$. Moreover, all spatial invariance is destroyed when the continuity control functions $\boldsymbol{w}(\boldsymbol{x})$ make jump transitions, so that the simple linear convolution filter interpretation breaks down in the presence of discontinuities, as one would expect. In the most general case then, the controlled-continuity stabilizer can be thought of as defining a complex nonlinear filter whose characteristics vary spatially in accordance with the stabilizer order p, the regularization parameter λ, the structure of the constraints, and the continuity control functions $\boldsymbol{w}(\boldsymbol{x})$.

Intriguing possibilities come to mind for applying controlled-continuity stabilizers to signal processing problems where standard filtering has been used in the past (such as for edge detection), as well as in situations where standard smoothing filters (e.g., Gaussians) would not apply. An example of the latter is discontinuity preserving smoothing of noisy, piecewise continuous signals [41]. Another is "scale space filtering" [47], but on irregularly sampled signals, with the regularization parameter continuously sweeping the "bandwidth" of the regularizer.

VII. Conclusion

Global smoothness constraints intrinsic to standard (Tikhonov) stabilizers are inadequate near discontinuities. Since discontinuities play a principal role in inverse visual problems, less restrictive constraints are necessary for visual regularization. This paper proposed controlled-continuity constraints whose smoothness properties may be regulated spatially to preserve visual discontinuities. They unify the treatment of regions and boundaries for the purposes of visual reconstruction or segmentation.

These generic constraints were formulated as multidimensional controlled-continuity stabilizers. The stabilizers comprise generalized spline kernels combined with noncontinuous continuity control functions. The latter need not be prespecified, but can be determined as part of the variational solution to the inverse problem.

Appendix A
The Visible-Surface Reconstruction Problem

An important reconstruction problem in computer vision arises in the computation of visible-surface representations. The task is to reconstruct dense 3-D representations of the shapes of visible surfaces from initial measurements provided by multiple specialized low-level visual processes. These include estimates of surface depth and orientation, as well as their discontinuities [40]–[42].

Assuming parallel projection onto the image plane for simplicity, let the true distance from the viewer to visible surfaces be given by the function $Z(x, y)$ where x and y are the image coordinates. Over much of the visual field $Z(x, y)$ is continuous and has continuously varying surface orientation. In addition, it generally exhibits orientation discontinuities at surface corners and creases and depth discontinuities along occluding contours.

Low-level visual processes generate a set of noise corrupted surface shape estimates (i.e., constraints) which can be expressed in the abstract notation

$$c_i = \mathcal{L}_i(Z(x, y)) + \epsilon_i$$

where $\mathcal{L}_i$ denotes measurement functionals of $Z(x, y)$ and ϵ_i denotes associated measurement errors. The visible-surface reconstruction problem is to reconstruct as faithfully as possible from the available constraints $\{\mathcal{L}_i, c_i\}_{i=1}^{N}$ the depth function $Z(x, y)$ with an explicit representation of its discontinuities.

Measurement functionals for surface reconstruction may be represented by point evaluation of generalized kth-order derivatives:

$$\mathcal{L}_i(Z(x, y)) = \left.\frac{\partial^k Z(x, y)}{\partial x^j \partial y^{k-j}}\right|_{(x_i, y_i)} \qquad j = 0, 1, \cdots, k.$$

The simplest measurement functionals, evaluation functionals of order $k = 0$, represent local depth constraints

$$c_i = Z(x_i, y_i) + \epsilon_i = d_{(x_i, y_i)}.$$

For orientation constraints, the components of the surface normal $\boldsymbol{n}(x, y) = [Z_x(x, y), Z_y(x, y), -1]$ are represented by order $k = 1$ derivatives:

$$c_i = Z_x(x_i, y_i) + \epsilon_i = p_{(x_i, y_i)}$$

$$c_i = Z_y(x_i, y_i) + \epsilon_i = q_{(x_i, y_i)}.$$

Visible-surface reconstruction is a fundamentally ill-posed problem due to the nature of the constraints. First, the initial measurements are contributed not by one, but by multiple specialized low-level visual processes. Since coincident measurements that are even slightly inconsistent will locally overdetermine surface shape, a solution need not exist. Second, the measurements are not dense, but scattered sparsely over the visual field. While they constrain surface shape locally, they generally do not determine a unique solution. Third, the measurements are inexact, subject to errors and noise. Because additive noise of high enough frequency, regardless how small its rms amplitude, can radically perturb local surface orientation, the solution is unstable with respect to perturbations of the measurements.

Appendix B
Mathematics of Generalized Spline Approximation

A. Euler–Lagrange Equations

Consider the variational principle VP with

$$\mathcal{E}(v(\boldsymbol{x})) = |v(\boldsymbol{x})|_m^2 + \int_{\Re^d} \alpha(\boldsymbol{x})\,(\mathcal{D}^k v(\boldsymbol{x}) - c(\boldsymbol{x}))^2.$$

This is the case of continuous data $c(\boldsymbol{x})$ and $\mathcal{L}v = \mathcal{D}^k v = \partial^k v / \partial x_1^{j_1} \cdots \partial x_d^{j_d}$ for $j_1 + \cdots + j_d = k$. The vanishing of the first variation $\delta\mathcal{E}(u) = 0$ can be shown to yield the following linear Euler–Lagrange equation:

$$(-1)^m \Delta^m u(\boldsymbol{x}) + (-1)^k \alpha(\boldsymbol{x})\,(\mathcal{D}^{2k} u(\boldsymbol{x}) - \mathcal{D}^k c(\boldsymbol{x})) = 0$$

where $\mathcal{D}^{2k} u = \mathcal{D}^k \mathcal{D}^k u$ and

$$\Delta^m = \sum_{j_1 + \cdots + j_d = m} \frac{m!}{j_1! \cdots j_d!} \left(\frac{\partial^{2m}}{\partial x_1^{2j_1} \cdots \partial x_d^{2j_d}}\right)$$

denotes the mth-order iterated Laplacian operator [36], a linear operator. Note that $m = 1$ yields the simple Laplacian operator

$$\Delta = \sum_{k=1}^{d} \frac{\partial^2}{\partial x_k^2}.$$

For discrete data $\{\mathcal{D}_{x_i}^k, c_i\}_{i=1}^{N}$, the second term of the Euler–Lagrange equation will involve Dirac distributions (delta functions) and their derivatives, e.g., for $k = 0$, it becomes $\Sigma_{i=1}^{N} \alpha_i (u(\boldsymbol{x}) - c_i)\, \delta(\boldsymbol{x} - \boldsymbol{x}_i)$.

B. Hilbert Space Characterization

Variational principles involving generalized splines are naturally set in Beppo–Levi spaces [11], [27]. The Beppo–Levi space of order m over $\Re^d$ is defined as the vector space of all the functions (i.e., Schwartz distributions) in $\Re^d$ whose partial derivatives (in the distributional sense) of total order m are square integrable; symbolically,

$$B^m(\Re^d) = \{v(\boldsymbol{x}) | (\partial^m v / \partial x_1^{j_1} \cdots \partial x_d^{j_d}) \in L_2 \quad \text{for } j_1 + \cdots + j_d = m\}.$$

Generalized spline functionals constitute the natural semi-norms of Beppo–Levi spaces. Equipped with these semi-norms, $B^m(\Re^d)$ is a semi-Hilbert space of continuous functions when $2m > d$. The null spaces $\mathfrak{N}$ of the semi-norms are simply the $M = \binom{d+m-1}{d}$ dimensional spaces of all polynomials $\{p_i\}_{i=1}^{M}$ over $\Re^d$ of total degree less than m [36, p. 60].

With $\mathbb{S}(v) = |v|_m^2$ a generalized spline functional, $\mathcal{E}(v)$ in the variational principle VP can, under certain conditions, become a norm in $\mathcal{H} = B^m(\mathcal{R}^d)$. This guarantees existence, uniqueness, and stability of the solution $u(x)$ to VP. A possible set of conditions is that the $\{\mathcal{L}_i\}_{i=1}^N$ include evaluation functionals at an $\mathfrak{N}$-unisolvent set of points, i.e., a set of M points in $\mathcal{R}^d$ which define a unique polynomial in the null space of the smoothness functional.

Given such a set of M points, an inner product of $B^m(\mathcal{R}^d)$ is defined by

$$(u, v)_m = \int_{\mathcal{R}^d} \sum_{j_1 + \cdots + j_d = m} \frac{m!}{j_1! \cdots j_d!} \cdot \left(\frac{\partial^m u(x)}{\partial x_1^{j_1} \cdots \partial x_d^{j_d}}\right)\left(\frac{\partial^m v(x)}{\partial x_1^{j_1} \cdots \partial x_d^{j_d}}\right) dx + \sum_{i=1}^M u(x_i)\, v(x_i).$$

Equipped with this inner product, $B^m(\mathcal{R}^d)$ becomes a Hilbert space. The corresponding norm $\|v\|_m$ is given by

$$\|v\|_m^2 = (v, v)_m = |v|_m^2 + \sum_{i=1}^M v^2(x_i).$$

Appendix C
Euler–Lagrange Equation for the Controlled-Continuity Stabilizer

Consider for the case of continuous data, as in Appendix B, the variational principle VP with

$$\mathcal{E}(v(x)) = |v(x)|_{p,w}^2 + \int_{\mathcal{R}^d} \alpha(x)\, (\mathcal{D}^k v(x) - c(x))^2$$

where $|v(x)|_{p,w}^2$ is the controlled-continuity stabilizer. Assuming that the continuity control functions $w(x)$ are differentiable to order p, the vanishing of the first variation $\delta\mathcal{E}(u) = 0$ yields the following Euler–Lagrange equation:

$$\sum_{m=0}^{p} (-1)^m \Delta_{w_m}^m(x)\, u(x) + (-1)^k\, \alpha(x) \cdot (\mathcal{D}^{2k} u(x) - \mathcal{D}^k c(x)) = 0$$

where

$$\Delta_{w_m}^m = \sum_{j_1 + \cdots + j_d = m} \frac{m!}{j_1! \cdots j_d!} \frac{\partial^m}{\partial x_1^{j_1} \cdots \partial x_d^{j_d}} \cdot \left(w_m(x) \frac{\partial^m}{\partial x_1^{j_1} \cdots \partial x_d^{j_d}}\right)$$

is a spatially weighted mth-order iterated Laplacian operator.

Generally, the continuity control functions will not be differentiable, which implies that the general operator is nonlinear and includes additional components involving Dirac distributions and their derivatives. Once again, for discrete data $\{\mathcal{D}_{x_i}^k, c_i\}_{i=1}^N$, the second term of the Euler–Lagrange equation will also involve these distributions.

Appendix D
Representation in Reproducing Kernel Hilbert Space

Consider the variational principle VP with

$$\mathcal{E}(v(x)) = |v(x)|_m^2 + \sum_{i=1}^N \frac{1}{\lambda\sigma_i^2} (\mathcal{L}_i(v) - c_i)^2$$

set in the space $B^m(\mathcal{R}^d)$. Assume that $\{\mathcal{L}_i\}_{i=1}^N$ are N linearly independent functionals. For the case $\mathcal{L}_i(v) = \mathcal{D}_{x_i}^k(v)$ with $2(m - k) > d$, and for $\mathcal{L}_i(v) = \int_{\Omega^d} K_i(x)\, v(x)\, dx$ with, e.g., Ω^d a bounded region and $\int_{\Omega^d} |K_i(x)|\, dx < \infty$, the $\mathcal{L}_i$ will be bounded (continuous) linear functionals.

When $2m > d$, the evaluation functional $\mathcal{L}_i(u) = u(x_i)$ is a bounded linear functional in $B^m(\mathcal{R}^d)$. This property makes $B^m(\mathcal{R}^d)$ a reproducing kernel Hilbert space. It follows that a unique (bounded) solution $u(x)$ will exist provided $N \geq M$, and $\mathcal{L}_i(\Sigma_{j=1}^M a_j p_j) = 0$, for $i = 1, 2, \cdots, N$ where $\{p_j\}_{j=1}^M$ are a basis for the $M = \binom{d+m-1}{d}$ dimensional space of polynomials of total degree less than m implies that all the a_j are 0 [26].

The solution has a representation

$$u(x) = \sum_{i=1}^N a_i \xi_i(x) + \sum_{j=1}^M b_j p_j(x)$$

where

$$\xi_i(x) = \mathcal{L}_{i(s)} E_m(r(x - s)), \qquad i = 1, 2, \cdots, N$$

and the subscript (s) indicates that the functional is applied to what follows considered as a function of s [46].

$E_m(x)$ is the fundamental solution of the iterated Laplacian Δ^m in $\mathcal{R}^d$, i.e., it satisfies the equation $\Delta^m E_m(x) = \delta(x)$ where δ is the Dirac delta distribution. It turns out to be the rotation invariant function defined on $\mathcal{R}^d - \{0\}$ given by

$$E_m(x) = \begin{cases} \dfrac{(-1)^{d/2+1}}{2^{2m-1}\pi^{d/2}(m-1)!\,(m-d/2)!}\, r^{2m-d} \ln r, \\ \qquad \text{if } 2m \geq d \text{ and } d \text{ even} \\ \dfrac{(-1)^m \Gamma(d/2 - m)}{2^{2m}\pi^{d/2}(m-1)!}\, r^{2m-d}, \\ \qquad \text{otherwise} \end{cases}$$

where $r(x) = |x| = (\Sigma_{i=1}^d x_i^2)^{1/2}$ denotes the Euclidean norm of x [36, p. 288].

The coefficients $\boldsymbol{a} = [a_1, \cdots, a_N]^T$ and $\boldsymbol{b} = [b_1, \cdots, b_M]^T$ are determined by the linear systems

$$(\boldsymbol{K} + \lambda\Sigma^2)\, \boldsymbol{a} + \boldsymbol{T}\boldsymbol{b} = \boldsymbol{c}$$

$$\boldsymbol{T}^T \boldsymbol{a} = 0$$

where $\boldsymbol{c} = [c_1, \cdots, c_N]^T$, the $N \times N$ symmetric matrix $\boldsymbol{K} = [\mathcal{L}_{i(x)} \mathcal{L}_{j(s)} E_m(d(\boldsymbol{x} - \boldsymbol{s}))]$, the $N \times M$ matrix $\boldsymbol{T} = [\mathcal{L}_i p_j]$, and the $N \times N$ diagonal matrix $\Sigma = [\sigma_i]$ [46].

Incidentally, the reproducing kernel in $B^m(\mathcal{R}^d)$ with the inner product given in Appendix B may be regarded as a continuous function on $\mathcal{R}^d \times \mathcal{R}^d$ defined by

$$K(s, t) = E_m(s, t) - \sum_{i=1}^{M} \pi_i(s)\, E_m(t, s_i) - \sum_{j=1}^{M} \pi_j(t)\, E_m(s_j, s) + \sum_{i,j=1}^{M} \pi_i(s)\, \pi_j(t) \cdot E_m(s_i, s_j) + \sum_{i=1}^{M} \pi_i(s)\, \pi_i(t)$$

where $E_m(s, t) = E_m(r(s - t))$ and $\{\pi_i\}_{i=1}^{M}$ are the polynomials of total degree less than m satisfying $\pi_i(s_j) = 1$ if $u = v$ and equal to zero otherwise [11], [26], [46].

Acknowledgment

I thank T. Poggio for discussions about regularization and J. Bliss for his comments on regularizing filters.

References

[1] J. H. Ahlberg, E. N. Nilson, and J. L. Walsh, *The Theory of Splines and their Applications*. New York: Academic, 1967.
[2] N. S. Bakhvalov, *Numerical Methods*. Moscow: Mir, 1977.
[3] H. G. Barrow and J. M. Tenenbaum, "Recovering intrinsic scene characteristics from images," in *Computer Vision Systems*, A. R. Hanson and E. M. Riseman, Eds. New York: Academic, 1978, pp. 3–26.
[4] B. A. Barsky and J. C. Beatty, "Local control of bias and tension in beta-splines," *ACM Trans. Graphics*, vol. 2, pp. 109–134, 1983.
[5] A. Blake, "The least-disturbance principle and weak constraints," *Pattern Recognition Lett.*, vol. 1, pp. 393–399, 1983.
[6] J. M. Brady and B. K. P. Horn, "Rotationally symmetric operators for surface interpolation," *Comput. Vision, Graphics, Image Processing*, vol. 22, pp. 70–94, 1983.
[7] I. C. Briggs, "Machine contouring using minimum curvature," *Geophys.*, vol. 39, pp. 39–48, 1974.
[8] A. K. Cline, "Scalar- and planar-valued curve fitting using splines under tension," *Commun. ACM*, vol. 17, pp. 218–220, 1974.
[9] R. Courant and D. Hilbert, *Methods of Mathematical Physics, Vol. I*. London: Interscience, 1953.
[10] T. F. Dolgopolova and V. K. Ivanov, "On numerical differentiation," *Zh. Vych. Mat. Fiz. (USSR Comp. Math. and Math. Phys.)*, vol. 6, pp. 570–576, 1966.
[11] J. Duchon, "Interpolation des fonctions de deux variables suivant le principe de la flexion des plaques minces," *R.A.I.R.O. Analyse Numérique*, vol. 10, pp. 5–12, 1976.
[12] ——, "Splines minimizing rotation-invariant semi-norms in Sobolev spaces," in *Constructive Theory of Functions of Several Variables*, A. Dodd and B. Eckmann, Eds. Berlin: Springer-Verlag, 1977, pp. 85–100.
[13] R. O. Duda and P. E. Hart, *Pattern Classification and Scene Analysis*. New York: Wiley, 1973.
[14] R. Franke, "Scattered data interpolation: Tests of some methods," *Math. Comput.*, vol. 38, pp. 181–199, 1982.
[15] S. Geman and D. Geman, "Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. PAMI-6, pp. 721–741, 1985.
[16] W. J. Gordon and R. F. Riesenfeld, "B-spline curves and surfaces," in *Computer Aided Geometric Design*, R. E. Barnhill and R. F. Riesenfeld, Eds. New York: Academic, 1974, pp. 95–126.
[17] 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.
[18] R. L. Harder and R. N. Desmarais, "Interpolation using surface splines," *J. Aircraft*, vol. 9, pp. 189–191, 1972.
[19] J. J. Hopfield, "Neurons with graded response have collective computational properties like those of two-state neurons," *Proc. Nat. Acad. Sci., USA*, vol. 81, pp. 3088–3092, 1984.
[20] G. Kimeldorf and G. Wahba, "A correspondence between Bayesian estimation on stochastic processes and smoothing by splines," *Ann. Math. Statist.*, vol. 41, pp. 495–502, 1970.
[21] S. Kirkpatrick, C. D. Gelatt, Jr., and M. P. Vecchi, "Optimization by simulated annealing," *Science*, vol. 220, pp. 671–680, 1983.
[22] C. Koch, J. L. Marroquin, and A. L. Yuille, "Analog 'neuronal' networks in early vision," M.I.T. A.I. Lab., Cambridge, MA, AI Memo 751, 1985.
[23] P. J. Laurent, *Approximation et Optimisation*. Paris: Hermann, 1972.
[24] D. Marr, "Representing visual information," in *Computer Vision Systems*, A. R. Hanson and E. M. Riseman, Eds. New York: Academic, 1978.
[25] J. L. Marroquin, "Surface reconstruction preserving discontinuities," M.I.T. A.I. Lab., Cambridge, MA, AI Memo 792, 1984.
[26] J. Meinguet, "An intrinsic approach to multivariate spline interpolation at arbitrary points," in *Polynomial and Spline Approximation: Theory and Applications*, B. N. Sahney, Ed. Dordrecht, The Netherlands: Reidel, 1979.
[27] ——, "Multivariate interpolation at arbitrary points made simple," *J. Appl. Math. Phys. (ZAMP)*, vol. 30, pp. 292–304, 1979.
[28] V. A. Morozov, *Methods for Solving Incorrectly Posed Problems*. New York: Springer-Verlag, 1984.
[29] G. M. Nielson, "Some piecewise polynomial alternatives to splines under tension," in *Computer Aided Geometric Design*, R. E. Barnhill and R. F. Riesenfeld, Eds. New York: Academic, 1974, pp. 209–235.
[30] D. Pilcher, "Smooth parametric surfaces," in *Computer Aided Geometric Design*, R. E. Barnhill and R. F. Riesenfeld, Eds. New York: Academic, 1974, pp. 237–253.
[31] T. Poggio and V. Torre, "Ill-posed problems and regularization analysis in early vision," M.I.T. A.I. Lab., Cambridge, MA, AI Memo 773; reprinted in *Proc. DARPA Image Understanding Workshop*, New Orleans, LA, L. S. Baumann, Ed., 1984, pp. 257–263.
[32] T. Poggio, H. Voorhees, and A. Yuille, "A regularized solution to edge detection," M.I.T. A.I. Lab., Cambridge, MA, AI Memo 833, 1985.
[33] G. Reinsch, "Smoothing by spline functions," *Numer. Math.*, vol. 10, pp. 177–183, 1967.
[34] I. J. Schoenberg, "Spline functions and the problem of graduation," *Proc. Nat. Acad. Sci.*, vol. 52, pp. 947–950, 1964.
[35] L. L. Schumaker, "Fitting surfaces to scattered data," in *Approximation II*, G. G. Lorentz, C. K. Chui, and L. L. Schumaker, Eds. New York: Academic, 1976, pp. 203–267.
[36] L. Schwartz, *Théorie des Distributions*. Paris: Hermann, 1966.
[37] D. G. Schweikert, "An interpolation curve using a spline in tension," *J. Math. Phys.*, vol. 45, pp. 312–317, 1966.
[38] D. Terzopoulos, "Multilevel reconstruction of visual surfaces: Variational principles and finite element representations," M.I.T. A.I. Lab., Cambridge, MA, 1982, AI Memo 671; reprinted in *Multiresolution Image Processing and Analysis*, A. Rosenfeld, Ed. New York: Springer-Verlag, 1984, pp. 237–310.
[39] ——, "Multilevel computational processes for visual surface reconstruction," *Comput. Vision, Graphics, Image Processing*, vol. 24, pp. 52–96, 1983.
[40] ——, "The role of constraints and discontinuities in visible-surface reconstruction," in *Proc. 8th Int. Joint Conf. AI*, Karlsruhe, W. Germany, 1983, pp. 1073–1077.
[41] ——, "Multiresolution computation of visible-surface representations," Ph.D. dissertation, Dep. Elec. Eng. Comput. Sci., M.I.T., Cambridge, MA, 1984.
[42] ——, "Computing visible-surface representations," M.I.T. A.I. Lab., Cambridge, MA, AI Memo 800, 1985.
[43] ——, "Image analysis using multigrid relaxation methods," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. PAMI-8, pp. 129–139, Mar. 1986.
[44] A. N. Tikhonov, "Regularization of incorrectly posed problems," *Sov. Math. Dokl.*, vol. 4, pp. 1624–1627, 1963.
[45] A. N. Tikhonov and V. A. Arsenin, *Solutions of Ill-Posed Problems*. Washington, DC: Winston, 1977.

[46] G. Wahba and J. Wendelberger, "Some new mathematical methods for variational objective analysis using splines and cross validation," *Monthly Weather Rev.*, vol. 108, pp. 1122–1143, 1980.
[47] A. P. Witkin, "Scale-space filtering," in *Proc. 8th Int. Joint Conf. AI*, Karlsruhe, W. Germany, 1983, pp. 1019–1022.
[48] S. W. Zucker, R. A. Hummel, and A. Rosenfeld, "An application of relaxation labeling to line and curve enhancement," *IEEE Trans. Comput.*, vol. C-26, pp. 394–403, 1977.

Demetri Terzopoulos (S'78–M'85), for a photograph and biography, see p. 139 of the March 1986 issue of this TRANSACTIONS.