MASSACHUSETTS INSTITUTE OF TECHNOLOGY
ARTIFICIAL INTELLIGENCE LABORATORY

A.I. Memo No. 800 March, 1985

COMPUTING VISIBLE-SURFACE REPRESENTATIONS

Demetri Terzopoulos

The low-level interpretation of images provides constraints on 3D surface shape at multiple resolutions, but typically only at scattered locations over the visual field. Subsequent visual processing can be facilitated substantially if the scattered shape constraints are immediately transformed into visible-surface representations that unambiguously specify surface shape at every image point. The required transformation is shown to lead to an ill-posed surface reconstruction problem. A well-posed variational principle formulation is obtained by invoking "controlled continuity," a physically nonrestrictive (generic) assumption about surfaces which is nonetheless strong enough to guarantee unique solutions. The variational principle, which admits an appealing physical interpretation, is locally discretized by applying the finite element method to a piecewise, finite element representation of surfaces. This forms the mathematical basis of a unified and general framework for computing visible-surface representations. The computational framework unifies formal solutions to the key problems of (i) integrating multiscale constraints on surface depth and orientation from multiple visual sources, (ii) interpolating these scattered constraints into dense, piecewise smooth surfaces, (iii) discovering surface depth and orientation discontinuities and allowing them to restrict interpolation appropriately, and (iv) overcoming the immense computational burden of fine resolution surface reconstruction. An efficient surface reconstruction algorithm is developed. It exploits multiresolution hierarchies of cooperative relaxation processes and is suitable for implementation on massively parallel networks of simple, locally interconnected processors. The algorithm is evaluated empirically in a diversity of applications.

This report 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 N00014-75-C-0643 and the System Development Foundation. The author was supported by the Natural Sciences and Engineering Research Council of Canada and the Fonds F.C.A.C., Quebec, Canada.

1. Introduction

Over thirty years ago, J.J. Gibson [1950] made the seminal conjecture that natural human perception amounts to the perception of *visible surfaces*. The explicit representation of visible surfaces, an intermediate goal of computational vision, has since attracted considerable interest.

The computational framework offered in this paper addresses, in a unified way, certain visual information processing tasks involved in the representation of visible surfaces. Particular emphasis is placed on utilizing highly parallel, cooperative processing to integrate surface shape information over multiple visual sources, to fuse it across a multiplicity of spatial resolutions, and to maintain the global consistency of the resulting distributed shape representations. The issues are first investigated in terms of a surface reconstruction model rooted in mathematical physics. This formal analysis is augmented by an empirical study of the resulting algorithms, which feature multiresolution iterative processing within hierarchical surface shape representations. The approach is guided by current knowledge of how humans perceive visible surfaces, while applications in machine vision provide a testbed for the algorithms.

The remainder of this introductory section examines the role of surface representations in early visual processing, outlines the key computational problems that will be of primary concern, and reviews some relevant prior work.

1.1. Early Visual Processing and Visible-Surface Representations

Early vision comprises a set of processes which specialize in recovering the physical properties of visible surfaces in a 3D scene from 2D images of the scene. They apply generic assumptions about the physical world and the imaging process to infer 3D surface shape constraints by interpreting specific image cues, such as stereoscopic disparity, motion, texture, contours, and shading. These conceptually independent shape estimation processes fall into two broad categories.

The first category comprises what are commonly referred to as correspondence processes. They operate over multiple image frames of a scene taken across space or over time. Paradigm examples are stereopsis and structure from motion (see, e.g., the review articles [Poggio and Poggio, 1984] and [Ullman, 1983]). Stereopsis is driven by computations on typically two image frames taken simultaneously, but from different spatial positions. The basic structure from motion computation involves frames taken from the same position, but at different times. If correspondences can be established across the frames, between image features which originate from the same point on a visible surface (not a trivial problem), then the *depth* (i.e., 3D distance) to such points can be estimated by triangulation, given the disparity (i.e., 2D displacement) between corresponding features as well as some knowledge of the imaging geometry.

The second category of shape estimation processes involve computations on a single static frame. Perspective projection of 3D scenes onto images imparts a systematic distortion to imaged surface properties such as shading, texture, and contours. A major part of this distortion can be attributed to the relative orientations of visible surfaces with respect to the viewer. In principle, it is possible to estimate surface *orientation* by measuring and interpreting such distortions in the image. This is the basis of practical approaches to recovering surface shape from shading, texture, and contours [Ikeuchi and Horn, 1981; Horn and Brooks, 1985; Kender, 1980; Witkin, 1981; Brady and Yuille, 1984].

The combined output of the shape estimation processes is best collected into intermediate representations of the 3D shapes and configurations of visible surfaces, which we will refer to as *visible-surface representations*. Notable among proposed visible-surface representations are the

depth and needle maps of Horn [1982], the intrinsic images of Barrow and Tenenbaum [1978], and the $2\frac{1}{2}$-D sketch of Marr and Nishihara [1978]. For humans, the perception of visible surfaces is generally immediate, involuntary, and seems to precede (object) recognition. This strongly suggests the existence of a visual process that autonomously computes visible-surface representations. Aside from the perceptual evidence, the availability of explicit visible-surface representations can also substantially facilitate subsequent surface analysis tasks in machine vision.

Since early visual processing provides relative surface shape estimates with respect to the viewer, it is most natural to define the basic shape primitives of visible-surface representations in a viewer-centered coordinate system. Moreover, the primitives should be computationally compatible with the local depth and orientation measurements (as well as discontinuities) that are provided by the various shape estimation processes. These criteria are satisfied by a particularly appealing class of local, piecewise shape primitives known as finite elements.

A crucial realization is that shape estimates can be provided at multiple resolutions. Indeed, multiresolution spatial frequency channels have been identified psychophysically in the human visual system (e.g., [Braddick *et al.*, 1978]). Their existence has influenced the design of early visual algorithms (e.g., [Marr, 1982]). In addition, machine vision research has demonstrated that multiresolution processing effectively bridges fine and coarse image structure, while it simultaneously increases computational efficiency (e.g., [Rosenfeld, 1984]). Hence, a multiresolution organization of visible-surface representations is most desirable [Terzopoulos, 1982, 1983a].

1.2. Key Problems of Visible-Surface Reconstruction

The main topic of concern in this paper is the development of a *visible-surface reconstruction process* responsible for generating and dynamically maintaining visible surface representations. Whether the intention is to model human vision or to design competent artificial vision systems, this process must solve four key problems [Terzopoulos, 1983b, 1984]: (i) the constraint integration problem, (ii) the interpolation problem, (iii) the discontinuity problem, and (iv) the computational efficiency problem. We elaborate on each of these problems next.

(i) ***The Constraint Integration Problem:*** Each specialized visual process may be thought of as a quasi-independent source of information partially constraining the shapes of visible surfaces. The human visual system is reliable and robust because it integrates the various processes, enabling them to complement one another. The integration of multiple sources of information introduces redundancy, which is necessary not only to resolve potential ambiguities, but also to overcome the detrimental effects of noise and inaccuracies in the initial shape estimates. The constraint integration problem is fundamentally one of devising an effective means of integrating all available surface depth and orientation constraints (and discontinuities) within a cooperative visible-surface reconstruction process.

(ii) ***The Interpolation Problem:*** It is widely accepted that initial descriptions of images ought to make explicit the occurrence and local 2D structure of image features that are correlated to salient events on physical surfaces (markings, boundaries, etc.). This is the essence, for instance, of Marr's "primal sketch" representation of significant image irradiance changes (edges) [Marr, 1982]. Generally, such salient features do not occur everywhere over the visual field. The initial representation of images as a sparse set of features implies that surface shape constraints generated by the specialized processes will also be scattered over a subset of image points. It is fascinating, however, that the human visual system systematically interprets visual stimuli such as sparse random dot stereograms as coherent 3D surfaces [Julesz, 1971]. Indeed, these stereograms continue to elicit perceptions of dense surfaces, even when the density of dots carrying disparity information has been reduced until depth is unspecified over 98 percent of

the visible surface area (see Fig. 1). It therefore appears that the surface reconstruction process is smoothly "filling in the gaps." This phenomenon has been the subject of some psychophysical investigation (e.g., [Collett, 1984]). The interpolation problem of visible-surface reconstruction challenges us to devise a scheme, consistent with human perception, for propagating shape information into indeterminate regions (devoid of shape estimates) from places where it is available.

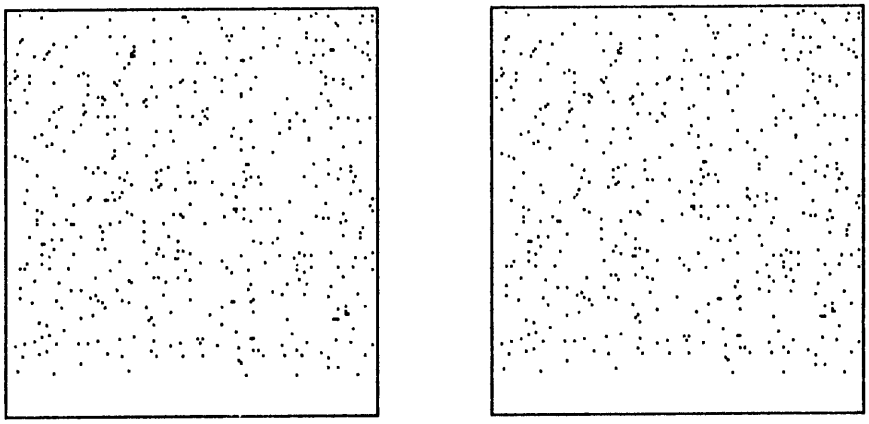

Figure 1. A sparse random dot stereogram. Binocular fusion of this stereogram reveals a planar surface as a central, opaque, textured square suspended nearer in depth over a similarly textured background. Vivid depth discontinuities separate the dense surfaces.

(iii) ***The Discontinuity Problem:*** Visual discontinuities result from significant, spatially-localized changes in the physical world, particularly abrupt changes in surface structure. Both depth and orientation discontinuities are perceptually relevant and provide vital boundary conditions for surface reconstruction. Discontinuities in depth occur at occluding contours, along which a surface in the scene occludes itself or another surface. Orientation discontinuities occur at creases or cusps of an otherwise continuous surface. In addition to the perception of coherent surfaces, random dot stereograms elicit vivid perceptions of surface discontinuities at abrupt disparity changes (see Fig. 1). The discontinuity problem amounts to (1) finding both depth and orientation discontinuities in surfaces, and (2) dealing with their presence during visible-surface reconstruction; i.e., allowing discontinuities to limit the otherwise smooth interpolation of shape constraints.

(iv) ***The Computational Efficiency Problem:*** Visible-surface reconstruction at the resolution of the image imposes an immense computational burden on both biological and artificial vision systems. Nevertheless, visible-surface representations must be computed quickly if they are to be of any practical value. It is generally accepted that to achieve the necessary performance, visual algorithms and mechanisms must emphasize parallelism [Ballard *et al.*, 1983]; however, visible-surface reconstruction is compute bound to the point where the fundamental limitations of massively parallel mechanisms, particularly with respect to global interprocessor communications, lead to severe inefficiencies. The computational efficiency problem is to develop a

visible-surface reconstruction process that not only exploits parallelism, but also overcomes cooperative communication bottlenecks to compute visible-surface representations quickly, given suitable architectures. For the reasons outlined in the foregoing section and following our previous work [Terzopoulos, 1982, 1983a], our solution to this problem hinges on the idea of multiresolution structuring of visual representations and associated cooperative processes.

1.3. Prior Work

There has been some prior work relating to the computation of visible-surface representations. Barrow and Tenenbaum [1979] describe an approach to reconstructing smooth surfaces from noisy visual data. This approach did not apply to general classes of surfaces, however, and the proposed relaxation algorithms were not supported by a firm mathematical analysis. Nevertheless, Barrow and Tenenbaum's [1978] basic model of intrinsic images and much of the philosophy underlying their computation seems appropriate, and it has influenced our approach.

The interpolation problem is related to classical spline approximation. A number of well-known surface approximation methods for scattered data are reviewed by Schumaker [1976]. Grimson [1983] employed one of these methods for the continuous interpolation of visual surfaces from depth constraints; a minimization scheme involving a particular functional containing second derivatives (he referred to it as the "quadratic variation"). Brady and Horn [1983] observe that this functional is related to the bending energy of a thin plate (a connection noted by Duchon [1977]), and the thin plate model was developed further by Terzopoulos [1983a] (see also [Blake, 1984]).

Interestingly, thin plate interpolants have appeared in other areas, including the interpolation of aircraft wing deflections [Harder and Desmarais, 1972], interpolation of meteorological fields [Wahba and Wendelberger, 1980], and the interpolation of digital terrain maps [Briggs, 1974; Bolondi *et al.*, 1976]. In this latter paper there is some concern for the presence of discontinuities (faults).

Following Ullman [1979] and others, Grimson [1983] pursued "biologically feasible," parallel and iterative algorithms for surface interpolation. A serious drawback of algorithms which satisfy these criteria is that they often converge excruciatingly slowly for problems of reasonable size. The idea of multiresolution surface reconstruction exploiting multigrid relaxation methods was shown to overcome this problem while adhering to biological feasibility [Terzopoulos, 1982, 1983a]. The multiresolution methodology yields efficient algorithms not only for the surface reconstruction problem but for other visual problems as well [Terzopoulos, 1984].

In retrospect, although progress has been made, a satisfactory computational theory of visible-surface representations has been elusive. This is largely a consequence of the significant technical obstacles encountered in devising formal solutions to all four key problems of visible-surface reconstruction within a unified computational framework. The difficulty of the task appears to have evoked some skepticism as to the actual computability (hence, even the usefulness) of intrinsic surface representations [Witkin and Tenenbaum, 1983]. Based on the theoretical generality of our approach and the accompanying empirical results, however, we believe such skepticism to be premature.

2. Mathematical Analysis of Visible-Surface Reconstruction

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. Low-level visual processes generate a set of noise corrupted

surface shape estimates (i.e., constraints) $\{c_i\}$ which can be expressed in the abstract notation

$$c_i = \mathcal{L}_i[Z(x,y)] + \epsilon_i, \tag{1}$$

where $\mathcal{L}_i$ denote measurement functionals of $Z(x,y)$ and ϵ_i denote associated measurement errors. Stated simply, the visible-surface reconstruction problem is to reconstruct, as faithfully as possible, the depth function $Z(x,y)$ from the available constraints $\{c_i\}$.

2.1. The Ill-Posed Nature of the Problem and Regularization

The problem is made nontrivial by the nature of the constraints. First, constraints are contributed not by one, but by multiple specialized early visual processes. Hence, slightly inconsistent measurements provided by different processes that happen to coincide will locally overdetermine surface shape. Second, constraints are not dense, but scattered sparsely over the visual field. Therefore, while they may restrict surface shape locally, they do not determine it uniquely everywhere; there remain very many feasible surfaces that are consistent with the constraints. Third, the measurements are subject to errors and noise. High spatial frequency additive noise, regardless how small its (RMS) amplitude, can locally perturb the surface (orientation) radically.

In view of the above three considerations, we cannot conclude in general that the solution will exist, nor that it will be unique, nor that it will be stable with respect to measurement errors. Mathematical problems for which the existence, uniqueness, or stability of solutions cannot be guaranteed *a priori* are said to be *ill-posed* [Tikhonov and Arsenin, 1977]. Visible-surface reconstruction can thus be characterized as a fundamentally ill-posed problem.

Ill-posed reconstruction (or inverse) problems are the rule rather than the exception in early vision [Poggio and Torre, 1984]. Ill-posed problems cannot be solved in general, without imposing some additional restrictions on possible solutions. This is the basis of a number of systematic approaches, notably the *regularization methods* introduced by Tikhonov and others (see [Tikhonov and Arsenin, 1977] and references therein). Duda and Hart [1973, Sec. 7.4] mention a basic form of regularization (essentially spatial smoothing) for combating the effects of noise in images. A more sophisticated class of regularization methods is discussed in the context of low-level vision by Poggio and Torre [1984].

Through regularization, ill-posed problems can be solved by reformulating them as *variational principles* that are effectively computable. Unlike the original problems, the variational principle formulations are *well-posed*; i.e., it is possible to guarantee the existence, uniqueness, and stability of their solutions under nonrestrictive conditions. Reformulation proceeds with the introduction of suitable *stabilizing functionals*, notably the class of stabilizers proposed by Tikhonov and Arsenin [1977, pp. 69–70]. These stabilizers can be interpreted as spline functionals that impose smoothness assumptions on the admissible solutions (by restricting them to Sobolev spaces of smooth functions).[1] Pragmatically then, this type of regularization is essentially equivalent to optimal approximation by generalized splines [Terzopoulos, 1985a]. We pursue the generalized spline approximation point of view, since splines are familiar and since they suggest helpful physical interpretations.

2.2. A Variational Principle

The abstract theory of optimal spline 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 [Laurent, 1972]. Let $\mathcal{H}$ be a linear space of smooth functions and let $\mathcal{S}(v)$ be a functional defined on $\mathcal{H}$ which measures the (lack of) smoothness of a function in

[1] Generic smoothness assumptions are generally the weakest (least committal) assumptions that one can make about feasible solutions and still obtain well-posed formulations.

$\mathcal{H}$. Furthermore, let P be a functional on $\mathcal{H}$ which provides a measure of the discrepancy between the function and the given constraints. Consider the following variational principle:

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

$$\mathcal{E}(u) = \inf_{v \in \mathcal{H}} \mathcal{E}(v), \tag{2}$$

where the energy functional

$$\mathcal{E}(v) = S(v) + P(v). \tag{3}$$

This variational principle will serve as a formal statement of the visible-surface reconstruction problem: The best reconstruction of the depth function $Z(x, y)$ from the available constraints will be given by the solution $u(x, y)$, the smoothest function in the admissible space $\mathcal{H}$ which is most compatible with the constraints.

Before proceeding to specify the *smoothness functional* $S(v)$ and the *penalty functional* $P(v)$, it should be noted that, if the solution exists, it satisfies the necessary condition for the minimum given by the vanishing of the first variation δ,

$$\delta\mathcal{E}(u) = \delta S(u) + \delta P(u) = 0, \tag{4},$$

which expresses the so called *Euler–Lagrange equations.*

2.3. Generalized Spline Functionals

For an appropriate smoothness functional $S(v)$, we turn to the multidimensional splines studied by Duchon [1977] and Meinguet [1979], generalizations of the classical univariate splines [Ahlberg, *et al.*, 1967]. The subclass of (2D) surface splines relevant to our problem can be characterized as members of a suitable space of admissible functions $v(x, y)$ which minimize the functional

$$|v|_m^2 = \int\int_{\Re^2} \sum_{j=0}^{m} \binom{m}{j} \left(\frac{\partial^m v}{\partial x^j \partial y^{m-j}} \right)^2 dx\, dy. \tag{5}$$

The positive integer m dictates the order of the partial derivatives that occur in the functional, which in turn determines the order of continuity possessed by the admissible functions. The Euler–Lagrange equation satisfied by the minimizing function $u(x, y)$ is an iterated version of Laplace's equation: $(-1)^m \Delta^m u = 0$, where $\Delta u = u_{xx} + u_{yy}$ is the Laplacian of u.

Low order surface splines have interesting physical interpretations involving equilibria of elastic bodies. Two special cases are of interest. For $m = 1$ the functional reduces to

$$|v|_1^2 = \int\int \left(v_x^2 + v_y^2 \right) dx\, dy, \tag{6}$$

which is proportional to the small deflection energy of a membrane (e.g., rubber sheet), while for $m = 2$,

$$|v|_2^2 = \int\int \left(v_{xx}^2 + 2v_{xy}^2 + v_{yy}^2 \right) dx\, dy, \tag{7}$$

is proportional to the small deflection bending energy of a thin plate (with zero Poisson ratio) [Courant and Hilbert, 1953]. Duchon [1977] refers to the minimizers of $|v|_2^2$ as thin plate splines. Since thin plate splines are the natural 2D analogs of cubic splines, $|v|_2^2$ finds frequent usage in surface interpolation problems [Schumaker, 1976]. In particular, it has been employed for visual surface interpolation [Grimson, 1983; Terzopoulos, 1983a].

The physical interpretations make it clear that membrane splines offer a lower order of continuity than thin plate splines. Since the physical forces in the membrane are due primarily to its surface tension, it generates minimal area surfaces. Although minimal area surfaces are continuous, they need not have continuous first partial derivatives; i.e., they are C^0 surfaces. For

instance, a sharp corner would result readily if an idealized physical membrane were subjected to the deflecting force of a knife edge. In contrast, the restoring forces in a physical thin plate are due primarily to its flexural rigidity. A thin plate would not crease when deflected by a knife edge. Thin plate splines therefore maintain continuity as well as continuous first partial derivatives; i.e., they generate C^1 surfaces.

2.4. Controlled Continuity and the Thin Plate Surface Under Tension

Generic smoothness assumptions are justified in pursuing a regularization approach to the visible-surface reconstruction problem, inasmuch as the coherence of matter tends to give rise to smoothly varying surfaces relative to the viewing distance, over some range of scales; however, smoothness assumptions clearly do not hold arbitrarily across surface discontinuities, some of which persist across all scales. This introduces significant complications for classical spline approximation or regularization methods; the continuity of spline functionals (or stabilizers) must be controlled at discontinuities in order to preserve them.

A stabilizer providing the necessary local continuity control can be realized as a weighted combination of generalized spline functionals of more than one order m [Terzopoulos, 1985a]. We propose the following smoothness functional:

$$S_{\rho\tau}(v) = \frac{1}{2}\int\int_{\Omega} \rho(x,y)\Big\{\tau(x,y)(v_{xx}^2 + 2v_{xy}^2 + v_{yy}^2) + [1-\tau(x,y)](v_x^2 + v_y^2)\Big\}\,dx\,dy, \qquad (8)$$

where Ω denotes the image domain, and $\rho(x,y)$ and $\tau(x,y)$ are real-valued weighting functions whose range is $[0,1]$. This controlled-continuity stabilizer is a weighted convex combination of the thin plate spline functional $|v|_2^2$ and membrane spline functional $|v|_1^2$ integrands. The associated Euler–Lagrange equation is

$$\frac{\partial}{\partial x^2}(\mu u_{xx}) + \frac{\partial}{\partial x \partial y}(2\mu u_{xy}) + \frac{\partial}{\partial y^2}(\mu u_{yy}) - \frac{\partial}{\partial x}(\eta u_x) - \frac{\partial}{\partial y}(\eta u_y) = 0, \qquad (9)$$

where $\mu(x,y) = \rho(x,y)\tau(x,y)$ and $\eta(x,y) = \rho(x,y)[1-\tau(x,y)]$, with natural (i.e., free) boundary conditions. The functional $S_{\rho\tau}(v)$ can be thought of as a *thin plate surface under tension*, where $\rho(x,y)$ is a spatially varying "rigidity" and $[1-\tau(x,y)]$ is the spatially varying "surface tension." It generalizes the unidimensional splines under tension of Schweikert (see [Ahlberg, *et al.*, 1967]).

The local continuity properties of the thin plate surface under tension functional can be controlled at any point $(x,y) \in \Omega$ by specifying the values of the *continuity control functions* $\rho(x,y)$ and $\tau(x,y)$ at that point. As τ approaches 1 the functional tends to a thin plate spline (a C^1 surface) whereas towards the other extreme, 0, the functional tends to a membrane spline (a C^0 surface) with intermediate values characterizing a hybrid C^1 surface that blends the properties of both constituent splines. ρ determines the overall potency of the smoothness functional.

Reconstructed surfaces must be able to faithfully preserve known depth and orientation discontinuities, while not introducing spurious discontinuities at other locations. This can be accomplished if (i) away from known depth and orientation discontinuities, the reconstructed surface possesses (at least) the C^1 smoothness of a thin plate, maintaining both continuity and continuous derivatives, (ii) at known orientation discontinuities, it exhibits just the C^0 smoothness of a membrane, maintaining continuity only, and (iii) at known depth discontinuities, the smoothness functional is deactivated so that the reconstructed surface is free to "fracture" locally. Hence, $S_{\rho\tau}(v)$ will be manipulated as follows: At all non-discontinuity points (x,y), $\rho(x,y)$ and $\tau(x,y)$ should be nonzero. At orientation discontinuity points, $\tau(x,y)$ is set to zero. At depth discontinuity points, $\rho(x,y)$ is set to zero. Mechanisms for automatically detecting discontinuities by computing continuity control functions optimally according to local criteria are considered in a subsequent section.

2.5. Penalty Functionals

Assuming independently distributed measurement errors ϵ_i with zero means and variances σ_i^2, the optimal measure of incompatibility is a weighted Euclidean norm of the discrepancy between the admissible function and the data c_i:

$$P(v) = \frac{1}{2}\sum_i \alpha_i(\mathcal{L}_i[v] - c_i)^2, \tag{10}$$

where the α_i are nonnegative real-valued weights (ideally α_i is inversely proportional to σ_i^2; i.e., $\alpha_i = 1/\lambda\sigma_i^2$) [Kimeldorf and Wahba, 1970]. This penalty functional can also be employed (suboptimally) when the above assumptions do not hold strictly.

Appropriate measurement functionals $\mathcal{L}_i$ for surface reconstruction may be synthesized from generalized k^{th}-order derivatives:

$$\mathcal{L}_i[v] = \left.\frac{\partial^k v}{\partial x^j \partial y^{k-j}}\right|_{(x_i,y_i)}, \qquad j = 0, 1, \ldots, k. \tag{11}$$

$k = 0$ yields simple evaluation functionals $\mathcal{L}_i[v(x,y)] = v(x_i, y_i)$, which will be employed to model the local depth constraints

$$c_i = v(x_i, y_i) + \epsilon_i = d_{(x_i,y_i)}. \tag{12}$$

The components of the local surface normal $\mathbf{n}(x_i, y_i) = [v_x(x_i, y_i), v_y(x_i, y_i), -1]$, which determine local surface orientation, can be handled by the first order ($k = 1$) derivative functionals $\mathcal{L}_i[v(x,y)] = v_x(x_i, y_i)$ and $\mathcal{L}_i[v(x,y)] = v_y(x_i, y_i)$ and yield analogous expressions for the local orientation constraints:

$$\begin{aligned} c_i &= v_x(x_i, y_i) + \epsilon_i = p_{(x_i,y_i)} \\ c_i &= v_y(x_i, y_i) + \epsilon_i = q_{(x_i,y_i)}. \end{aligned} \tag{13}$$

Other potentially relevant functionals such as directional derivatives can be accommodated straightforwardly with the above notation.

It is convenient to separate the various constraints into three sets; the set $i \in D$ of image points at which depth constraints $d_{(x_i,y_i)}$ occur, and the sets $i \in P$ and $i \in Q$ at which orientation constraints $p_{(x_i,y_i)}$ and $q_{(x_i,y_i)}$ occur respectively. The penalty functional can then be expressed as a sum of three components

$$P(v) = \frac{1}{2}\sum_{i\in D} \alpha_{d_i}[v(x_i, y_i) - d_{(x_i,y_i)}]^2 + \frac{1}{2}\sum_{i\in P} \alpha_{p_i}[v_x(x_i, y_i) - p_{(x_i,y_i)}]^2 + \frac{1}{2}\sum_{i\in Q} \alpha_{q_i}[v_y(x_i, y_i) - q_{(x_i,y_i)}]^2, \tag{14}$$

where the α_i parameters are now distinguished as α_{d_i}, α_{p_i}, and α_{q_i}.

2.6. A Physical Model for Visible-Surface Reconstruction

The variational principle formulation of the surface reconstruction problem has an appealing physical interpretation which is illustrated in Fig. 2. The thin plate surface under tension may be visualized as an elastic surface, planar in its natural state, whose elastic bending energy $S_{pr}(v)$ stabilizes surface shape so that it varies smoothly in between constraints (but not at discontinuities). Constraints deflect the surface according the penalty functional $P(v)$, which can be interpreted as the total stretching energy of a set of ideal springs attached to the constraints. The left part of the figure shows the elastic surface whose deflection $u(x,y)$ at equilibrium is determined by an infrastructure of scattered depth constraints. The local depth estimate is encoded as the vertical height of the constraint and the tightness of each constraint is controlled by associated spring stiffness α_{d_i}. The right part of the figure illustrates an orientation constraint coercing the local surface normal. The spring stiffness is determined by the constraint parameters α_{p_i} and α_{q_i}.

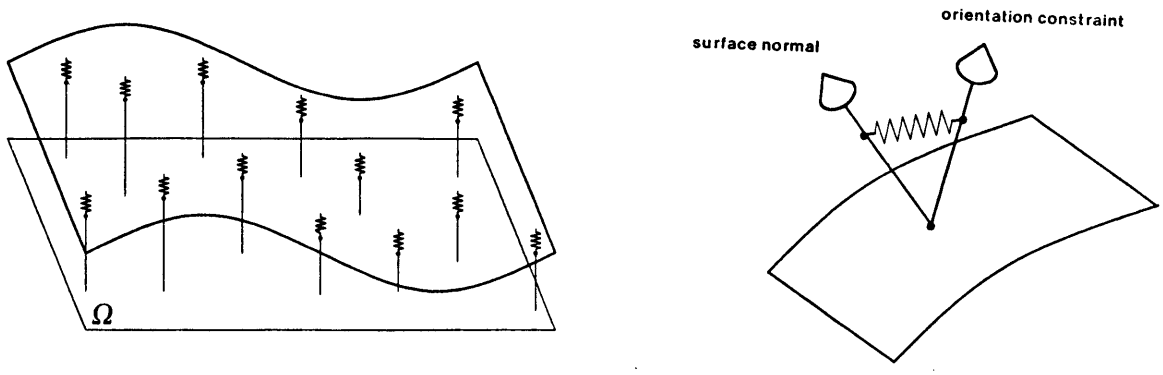

Figure 2. The physical model. Thin plate surface under tension and depth constraints (left). Local influence of an orientation constraint (right).

2.7. Existence, Uniqueness, and Stability of the Solution

Existence, uniqueness, and stability of the solution $u(x, y)$ to the variational principle VP are guaranteed when $\mathcal{E}_{\rho\tau}(v) = S_{\rho\tau}(v) + P(v)$ is a norm in the admissible space $\mathcal{H}$. Unfortunately, generalized spline functionals $|v|_m^2$ are *a priori* only semi-norms (of a particular class of Sobolev spaces). The null spaces $\mathcal{N}$ of functions that map to zero under the semi-norms are simply the ($M = \binom{m+1}{2}$ dimensional) spaces of all polynomials over $\Re^2$ of degree less than or equal to $m-1$ [Meinguet, 1979]. The penalty functional $P(v)$ can force $\mathcal{E}_{\rho\tau}(v)$ to be a norm, however, if it at least constrains $\mathcal{N}$ to a unique polynomial. A possible set of conditions for this to occur is that the $\mathcal{L}_i$ include evaluation functionals at an $\mathcal{N}$-unisolvent set of points (i.e., a set of M points which define a unique polynomial in the null space of the smoothness functional). In particular, since the maximum order of generalized splines in the stabilizer $S_{\rho\tau}(v)$ is $m = 2$, its null space is the space of linear polynomials. Thus, the following proposition can be proven [Terzopoulos, 1984]:

Proposition. *The solution $u(x, y)$ will exist, be unique, and stable given any one of the following minimal conditions*

(i) three noncolinear depth constraints,

(ii) two depth constraints as well as a single p or q constraint,

(iii) a single depth constraint as well as a single p and a single q constraint,

(iv) a single p and a single q constraint with the "center of gravity" of the surface fixed.

These minimal conditions will hold in practice, due to the large number of constraints typically available from early shape estimation processes (the fixed center of gravity condition can be imposed when necessary). Consequently, the visible-surface reconstruction problem may be considered well-posed, hence effectively computable in general.

Satisfying the conditions for a well-posed problem essentially guarantees that a unique state of stable equilibrium will exist for the plate/spring system (the minimal energy state $\mathcal{E}_{\rho\tau}(u)$). In this context, the controlled continuity assumption about surfaces, as embodied by the thin plate surface

under tension model, is physically nonrestrictive but nonetheless powerful enough to guarantee the existence of unique solutions to the variational principle.

3. Discretization

It is extremely difficult, if not impossible, to obtain an analytic solution to the variational principle due to the irregular occurrence and geometry of constraints and discontinuities. For our purposes, the only viable approach is to convert the continuous surface reconstruction problem to an equivalent discrete problem whose solution can be computed numerically. To this end, *finite elements* make ideal local surface shape primitives for use in visible-surface representations [Terzopoulos, 1982, 1983a]. The finite element method [Strang and Fix, 1973] is a general, powerful, and mathematically rigorous approximation technique which guides the selection of appropriate elements and governs their interactions according to the nature of the variational principle.

The finite element method offers substantial flexibility in discretizing domains with irregular shaped boundaries. Although the use of irregularly shaped elements to discretize such domains may not present a feasibility problem with regard to distributed biological mechanisms, it makes nontrivial the mapping of elemental computations onto regularly interconnected processing networks typically provided by VLSI technology. In this paper we restrict ourselves to regular finite elements in order to facilitate such mappings. Since the goal is to obtain a particularly fine discretization, at the resolution of the image, the restriction to fine regular elements will not jeopardize our ability to accommodate the irregular occurrence of constraints or discontinuities.

3.1. The Discrete Equations

The domain Ω is tessellated into square element subdomains with sides of length h. Nodes are located at element corners and shared by adjacent elements. This results in a planar and uniform square grid of nodes that is ideally suited to VLSI implementation. The nodes are naturally indexed by (i, j) for $i = 1, \ldots, N_x$ and $j = 1, \ldots, N_y$, where N_x and N_y are the number of nodes along the x and y axis respectively of the (rectangular) domain Ω. The total number of nodes is $N = N_x \times N_y$. The reconstructed surface is represented by an assembly of (nonconforming) finite elements, each of which is a six-point (full) quadratic interpolant defined locally within its particular subdomain. The unknown displacement (surface depth), at node (i, j) is denoted by the variable $v^h_{i,j} = v(ih, jh)$. Taken together, the displacement variables are denoted by the vector $\mathbf{v}^h \in \Re^N$. Once this vector is determined by solving a discrete version of the variational principle, the local interpolants are known exactly and, consequently, they explicitly represent depth and orientation everywhere over the surface.

The proposed square, quadratic element leads to the following $O(h^2)$ formulas for the required

partial derivatives at an arbitrary node (i, j) [Terzopoulos, 1983a]:

$$\begin{aligned}
v_{xx}^h &\Rightarrow \frac{1}{h^2}\left(\mathrm{v}_{i+1,j}^h - 2\mathrm{v}_{i,j}^h + \mathrm{v}_{i-1,j}^h\right);\\
v_{yy}^h &\Rightarrow \frac{1}{h^2}\left(\mathrm{v}_{i,j+1}^h - 2\mathrm{v}_{i,j}^h + \mathrm{v}_{i,j-1}^h\right);\\
v_{xy}^h &\Rightarrow \frac{1}{h^2}\left(\mathrm{v}_{i+1,j+1}^h - \mathrm{v}_{i,j+1}^h - \mathrm{v}_{i+1,j}^h + \mathrm{v}_{i,j}^h\right);\\
v_x^h &\Rightarrow \frac{1}{h}\left(\mathrm{v}_{i+1,j}^h - \mathrm{v}_{i,j}^h\right);\\
v_y^h &\Rightarrow \frac{1}{h}\left(\mathrm{v}_{i,j+1}^h - \mathrm{v}_{i,j}^h\right);
\end{aligned} \tag{15}$$

Note that the formulas are finite difference expressions. Their appearance is due to the uniformity and low order of the element and their relative simplicity will facilitate the calculations substantially.

Substituting the above expressions with the constant approximations $\rho(x, y) \Rightarrow \rho_{i,j}^h$ and $\tau(x, y) \Rightarrow \tau_{i,j}^h$ into (8), and noting that the area of each element is h^2, we obtain the discrete functional

$$\begin{aligned}
S_{\rho\tau}^h(\mathrm{v}^h) = \frac{1}{2}\sum_{i,j}\rho_{i,j}^h\Bigg\{\frac{\tau_{i,j}^h}{h^2}\Bigg[&\left(\mathrm{v}_{i+1,j}^h - 2\mathrm{v}_{i,j}^h + \mathrm{v}_{i-1,j}^h\right)^2\\
&+ 2\left(\mathrm{v}_{i+1,j+1}^h - \mathrm{v}_{i,j+1}^h - \mathrm{v}_{i+1,j}^h + \mathrm{v}_{i,j}^h\right)^2\\
&+ \left(\mathrm{v}_{i,j+1}^h - 2\mathrm{v}_{i,j}^h + \mathrm{v}_{i,j-1}^h\right)^2\Bigg]\\
+\left[1 - \tau_{i,j}^h\right]&\left[\left(\mathrm{v}_{i+1,j}^h - \mathrm{v}_{i,j}^h\right)^2 + \left(\mathrm{v}_{i,j+1}^h - \mathrm{v}_{i,j}^h\right)^2\right]\Bigg\}.
\end{aligned} \tag{16}$$

Although by no means a necessity, it is both natural (in view of common image discretization) and convenient to assume that the constraints coincide with nodes (i, j) of the grid. Hence, to obtain a discrete expression for $P(v)$, we collect the nodes at which the various constraints occur into three sets; the set $(i, j) \in D$ at which depth constraints $\mathrm{d}_{i,j}^h$ occur, and the sets $(i, j) \in P$ and $(i, j) \in Q$ at which orientation constraints $\mathrm{p}_{i,j}^h$ and $\mathrm{q}_{i,j}^h$ occur. Using symmetric difference approximations for the partial derivatives in (13), the discrete penalty functional may be written in terms of the nodal variables as

$$\begin{aligned}
P^h(\mathrm{v}^h) = &\frac{1}{2}\sum_{(i,j)\in D}\alpha_{d_{i,j}}^h\left(\mathrm{v}_{i,j}^h - \mathrm{d}_{i,j}^h\right)^2\\
&+ \frac{1}{2}\sum_{(i,j)\in P}\alpha_{p_{i,j}}^h\left(\frac{1}{2h}\left(\mathrm{v}_{i+1,j}^h - \mathrm{v}_{i-1,j}^h\right) - \mathrm{p}_{i,j}^h\right)^2\\
&+ \frac{1}{2}\sum_{(i,j)\in Q}\alpha_{q_{i,j}}^h\left(\frac{1}{2h}\left(\mathrm{v}_{i,j+1}^h - \mathrm{v}_{i,j-1}^h\right) - \mathrm{q}_{i,j}^h\right)^2.
\end{aligned} \tag{17}$$

The energy-minimizing vector of nodal displacements $\mathbf{u}^h$ satisfies the equilibrium condition

$$\nabla\mathcal{E}_{\rho\tau}^h(\mathbf{u}^h) = \nabla S_{\rho\tau}^h(\mathbf{u}^h) + \nabla P^h(\mathbf{u}^h) = 0, \tag{18}$$

where ∇ is the gradient operator. Since the discrete functional $\mathcal{E}_{\rho\tau}^h(\mathbf{u}^h)$ is a quadratic form in the $\mathrm{u}_{i,j}^h$, the above equation defines a linear system of simultaneous equations that are satisfied by $\mathbf{u}^h$. The discrete problem amounts to solving these *nodal equations*.

3.2. Computational Molecules

To progress towards explicit expressions for the nodal equations, we first determine the partial derivatives of $S^h_{\rho\tau}(\mathbf{u}^h)$ and $\mathcal{P}^h(\mathbf{u}^h)$ with respect to an arbitrary nodal variable $u^h_{i,j}$. Letting

$$\mu^h_{i,j} = \rho^h_{i,j}\tau^h_{i,j}/h^2 \qquad \text{and} \qquad \eta^h_{i,j} = \rho^h_{i,j}(1-\tau^h_{i,j}), \tag{19}$$

we obtain

$$\begin{aligned}
\frac{\partial S^h_{\rho\tau}(\mathbf{u}^h)}{\partial u^h_{i,j}} = \Big\{ & \left(u^h_{i,j} - 2u^h_{i-1,j} + u^h_{i-2,j}\right)\mu^h_{i-1,j} \\
& + \left(-2u^h_{i+1,j} + 4u^h_{i,j} - 2u^h_{i-1,j}\right)\mu^h_{i,j} \\
& + \left(u^h_{i+2,j} - 2u^h_{i+1,j} + u^h_{i,j}\right)\mu^h_{i+1,j} \\
& + \left(2u^h_{i,j} - 2u^h_{i-1,j} - 2u^h_{i,j-1} + 2u^h_{i-1,j-1}\right)\mu^h_{i-1,j-1} \\
& + \left(-2u^h_{i+1,j} + 2u^h_{i,j} + 2u^h_{i+1,j-1} - 2u^h_{i,j-1}\right)\mu^h_{i,j-1} \\
& + \left(-2u^h_{i,j+1} + 2u^h_{i-1,j+1} + 2u^h_{i,j} - 2u^h_{i-1,j}\right)\mu^h_{i-1,j} \\
& + \left(2u^h_{i+1,j+1} - 2u^h_{i,j+1} - 2u^h_{i+1,j} + 2u^h_{i,j}\right)\mu^h_{i,j} \\
& + \left(u^h_{i,j} - 2u^h_{i,j-1} + u^h_{i,j-2}\right)\mu^h_{i,j-1} \\
& + \left(-2u^h_{i,j+1} + 4u^h_{i,j} - 2u^h_{i,j-1}\right)\mu^h_{i,j} \\
& + \left(u^h_{i,j+2} - 2u^h_{i,j+1} + u^h_{i,j}\right)\mu^h_{i,j+1}\Big\} \\
+ \Big\{ & \left(u^h_{i,j} - u^h_{i-1,j}\right)\eta^h_{i-1,j} + \left(u^h_{i,j} - u^h_{i+1,j}\right)\eta^h_{i,j} \\
& + \left(u^h_{i,j} - u^h_{i,j-1}\right)\eta^h_{i,j-1} + \left(u^h_{i,j} - u^h_{i,j+1}\right)\eta^h_{i,j}\Big\},
\end{aligned} \tag{20}$$

the discrete version of (9). Next, for $(i,j) \in D \cap P \cap Q$,

$$\begin{aligned}
\frac{\partial \mathcal{P}^h(\mathbf{u}^h)}{\partial u_{i,j}} = & \left(\alpha^h_{d_{i,j}} u^h_{i,j} - \alpha^h_{d_{i,j}} d^h_{i,j}\right) \\
& + \left(\frac{\alpha^h_{p_{i-1,j}}}{4h^2}\left(u^h_{i,j} - u^h_{i-2,j}\right) - \frac{\alpha^h_{p_{i-1,j}}}{2h} p^h_{i-1,j}\right) \\
& + \left(\frac{\alpha^h_{p_{i+1,j}}}{4h^2}\left(u^h_{i,j} - u^h_{i+2,j}\right) + \frac{\alpha^h_{p_{i+1,j}}}{2h} p^h_{i+1,j}\right) \\
& + \left(\frac{\alpha^h_{q_{i,j-1}}}{4h^2}\left(u^h_{i,j} - u^h_{i,j-2}\right) - \frac{\alpha^h_{q_{i,j-1}}}{2h} q^h_{i,j-1}\right) \\
& + \left(\frac{\alpha^h_{q_{i,j+1}}}{4h^2}\left(u^h_{i,j} - u^h_{i,j+2}\right) + \frac{\alpha^h_{q_{i,j+1}}}{2h} q^h_{i,j+1}\right).
\end{aligned} \tag{21}$$

The above expressions specify the nodal equations implicitly. Each constituent term in (round) parentheses can be represented graphically as a basic *computational molecule*. Computational molecules will be interpreted both as spatial representations of the nonzero coefficients in the nodal

equations, and as local computations involving multiplications and additions of specific proximal nodal variables. The former interpretation will facilitate the construction of individual nodal equations given some local structure of constraints and discontinuities, while the latter will lead directly to local iterative algorithms for solving the resulting simultaneous linear system.

3.2.1. Basic Molecules

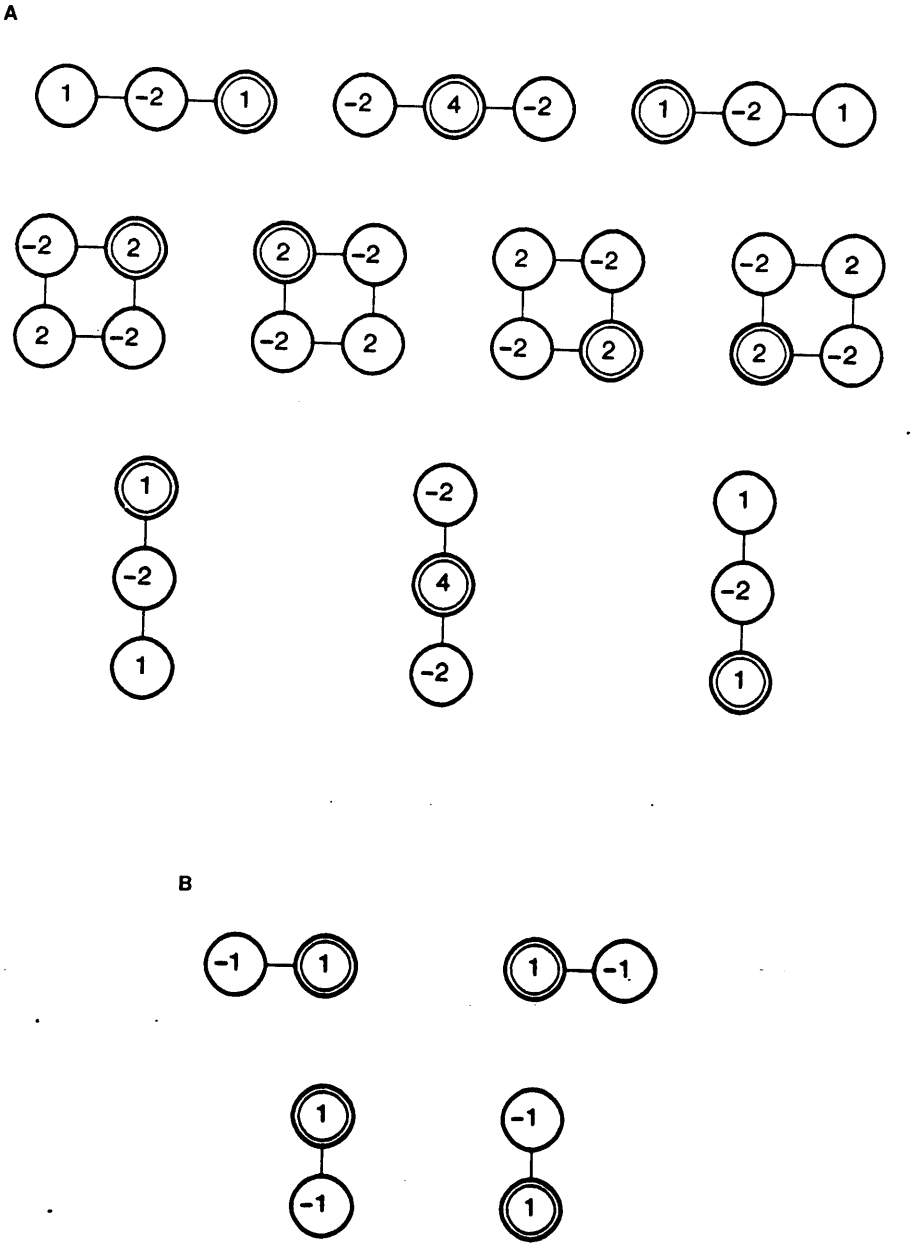

Figure 3. Plate molecules (a) and membrane molecules (b).

Eq. (20) is a convex combination of two components; the first stemming from the thin plate energy functional, the second, from the membrane energy functional. Each constituent term yields a basic computational molecule (see Fig. 3), a set of linked *atoms* indicated by circles. The central node (i, j) is indicated by a double circle in each molecule. Fig. 3(a) illustrates the ten *plate molecules* obtained from the terms of the first component, while Fig. 3(b) shows the four *membrane molecules* obtained from the terms of the second component. Each atom contains the coefficient of the associated nodal variable (aside from the $\mu^h_{i,j}$ and $\eta^h_{i,j}$ factors).

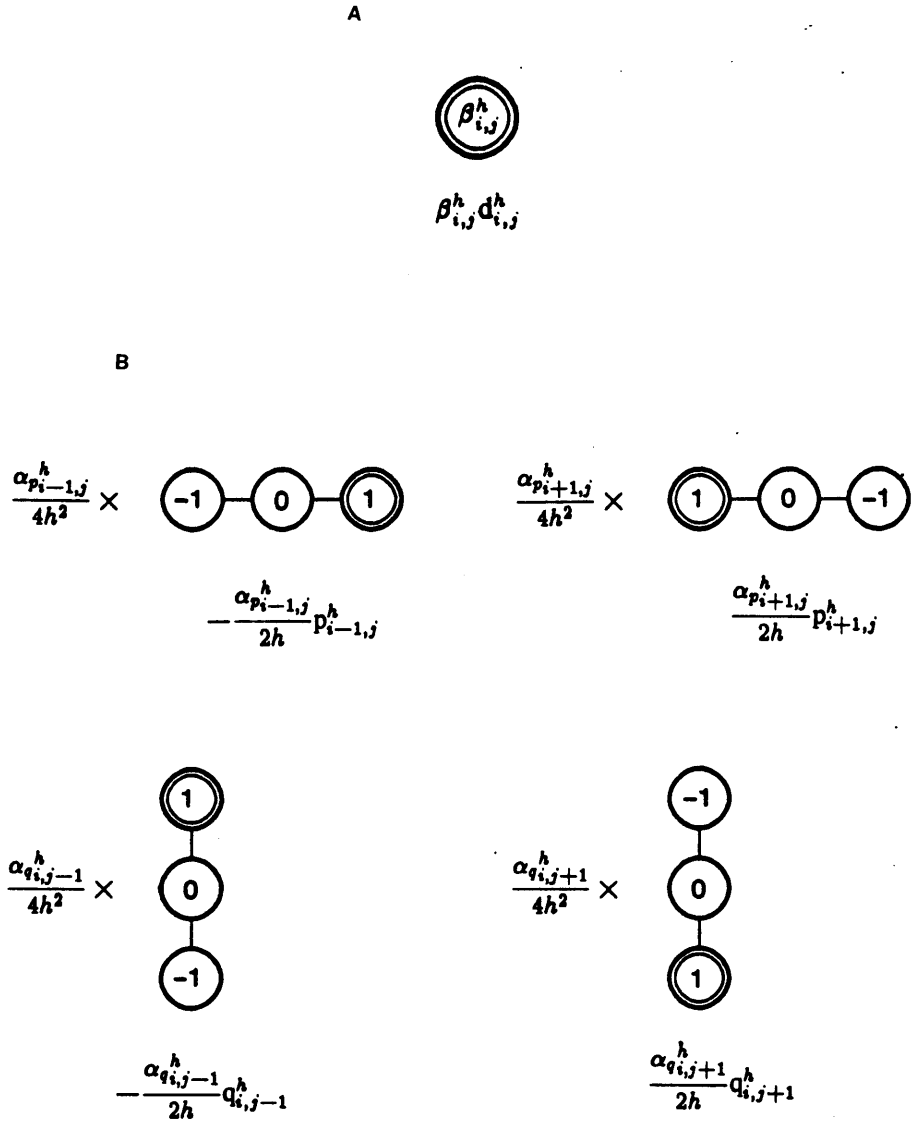

Figure 4. Depth constraint molecule (a) and orientation constraint molecules (b).

Similarly, the depth constraint term in (21) can be represented by the *depth constraint molecule* shown in Fig. 4(a). Associated with it is the factor $\alpha_{d^h_{i,j}} d^h_{i,j}$, which is indicated underneath the molecule. The remaining orientation constraint terms of (21) are represented by the *orientation constraint molecules* and associated factors shown in Fig. 4(b).

3.2.2. Molecular Summation within Smooth Regions

The formation of nodal equations within continuous regions can be visualized as a process of *molecular summation.* During molecular summation, the basic molecules combine at the central node, coincident atoms summing together.

When (i, j) is an interior node, away from constraints and discontinuities, $\rho^h_{i,j} = \tau^h_{i,j} = 1$, and only the plate component of (20) contributes to the expression for the partial derivative. Hence, the equilibrium condition (18) reduces to the nodal equation

$$\begin{aligned} 0 = \frac{20}{h^2}u^h_{i,j} - \frac{8}{h^2}\left(u^h_{i-1,j} + u^h_{i+1,j} + u^h_{i,j-1} + u^h_{i,j+1}\right) \\ + \frac{2}{h^2}\left(u^h_{i-1,j-1} + u^h_{i+1,j-1} + u^h_{i-1,j+1} + u^h_{i+1,j+1}\right) \\ + \frac{1}{h^2}\left(u^h_{i-2,j} + u^h_{i+2,j} + u^h_{i,j-2} + u^h_{i,j+2}\right). \end{aligned} \tag{22}$$

This equation can be represented by the composite *nodal molecule* illustrated in Fig. 5(a), which results from the summation of the plate molecules in Fig. 3.

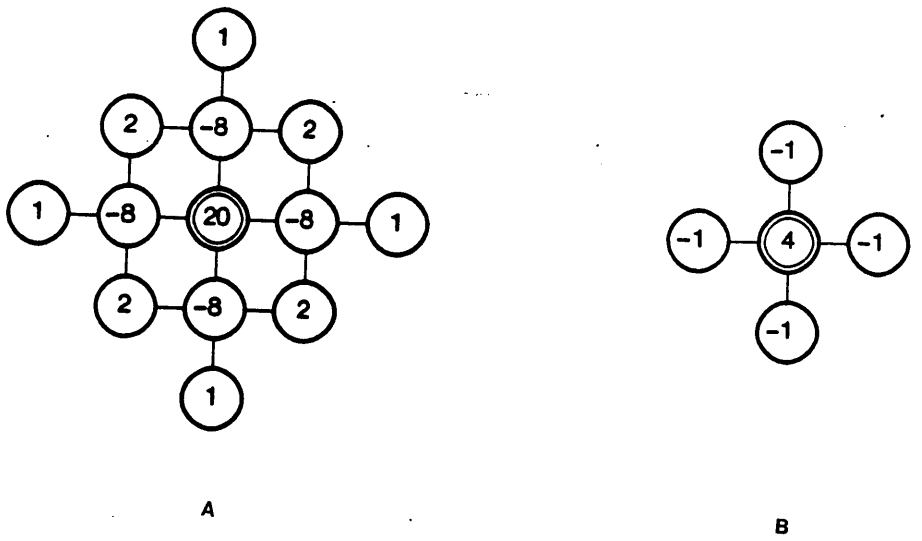

Figure 5. Interior node molecules. (a) Away from discontinuities. (b) At interior orientation discontinuities.

Note that the computational molecule for the center of the region is a factor of h^2 (due to the elemental area) times an order $O(h^2)$ finite difference approximation for the biharmonic operator [Abramowitz and Stegun, 1965, p. 885], the Euler–Lagrange equation associated with the thin plate spline. This is an expected consequence of the particular element employed which yielded finite difference approximations for the second partial derivatives of v^h.

If node (i, j) is a depth constraint, the first term in (21) takes part in the nodal equation. The effect can be represented as a summation of the depth constraint molecule and associated constraint factor with the nodal molecule for (i, j) shown in Fig. 5(a).

Similarly, if $(i-1, j)$ or $(i+1, j)$ are p constraints, or $(i, j-1)$ or $(i, j+1)$ are q constraints, the other terms in (21) participate in the nodal equation. Again this can be represented as the

summation of computational molecules. Specifically, the upper left molecule in Fig. 4(b) sums with the nodal molecule if $(i-1,j) \in P$, the upper right only if $(i+1,j) \in P$, the lower left only if $(i,j-1) \in Q$, and the lower right only if $(i,j+1) \in Q$.

3.2.3. Molecular Inhibition at Discontinuities

If (i,j) is a discontinuity, either $\mu^h_{i,j}$ or $\eta^h_{i,j}$ or both may be zero, thus nullifying the summation of specific molecules. This crucial influence of the discrete continuity control functions near known discontinuities will be referred to as *molecular inhibition.* It was convenient for expressing (20) and (21) to discretize $u(x,y)$, $\rho(x,y)$, and $\tau(x,y)$ over the same set of nodes. Although orientation discontinuities can be situated at these nodes (since $\mathrm{u}^h_{i,j}$ is defined at an orientation discontinuity), it is better to position depth discontinuities on the links half way between nodes (since $\mathrm{u}^h_{i,j}$ is undefined at a depth discontinuity). Discretizing $\tau(x,y)$ on links does not present a problem in practice. As a general rule, a discontinuity may inhibit a molecule only if it coincides with a constituent atom or link.

First consider orientation discontinuities. At an orientation discontinuity, $\tau^h_{i,j} = 0$ and only the second component of (20) contributes to the nodal equation. In effect, the plate molecules are inhibited and replaced by the membrane molecules of Fig. 3. At an interior orientation discontinuity (i,j), away from depth discontinuities, all four membrane molecules superpose to yield the nodal molecule shown in Fig. 5(b), which represents the nodal equation

$$4\mathrm{u}^h_{i,j} - \mathrm{u}^h_{i-1,j} - \mathrm{u}^h_{i+1,j} - \mathrm{u}^h_{i,j-1} - \mathrm{u}^h_{i,j+1} = 0. \tag{23}$$

The equation will be recognized as $-h^2$ times a standard finite difference equation for the Laplacian [Abramowitz and Stegun, 1965, p. 885]. It too appears as a consequence of the Euler–Lagrange equation associated with the membrane spline.

Since an orientation constrain cannot meaningfully coincide with an orientation discontinuity, orientation discontinuity nodes inhibit orientation constraint molecules. On the other hand, depth constraint molecules are not inhibited by orientation discontinuities since it is perfectly reasonable to locally constrain a membrane spline in depth.

Because smoothness constraints are unsuitable at a depth discontinuity node (i,j) (i.e., $\rho_{i,j} = 0$), a nodal equation cannot involve nodal variables separated by or coinciding with a depth discontinuity. Consequently, depth discontinuities inhibit all of the basic computational molecules. Fig. 6(a) illustrates examples (disregarding constraints) of nodal molecules for boundary nodes (marked as double circles) which are near depth discontinuity nodes (marked by X's). Examples of nodal molecules at boundary orientation discontinuities (double circles) next to depth discontinuities (X's) are shown in Fig. 6(b).

4. Detection and Localization of Surface Discontinuities

An important feature of our framework for computing visible-surface representations is the uniform treatment of constraints and discontinuities, essentially as localized and independent surface shape primitives. This facilitates the parallel integration of discontinuity information, along with shape constraints, over the various early shape estimation processes. It is convenient to think of discontinuity information as being collected into a *discontinuity map* which is in registration with

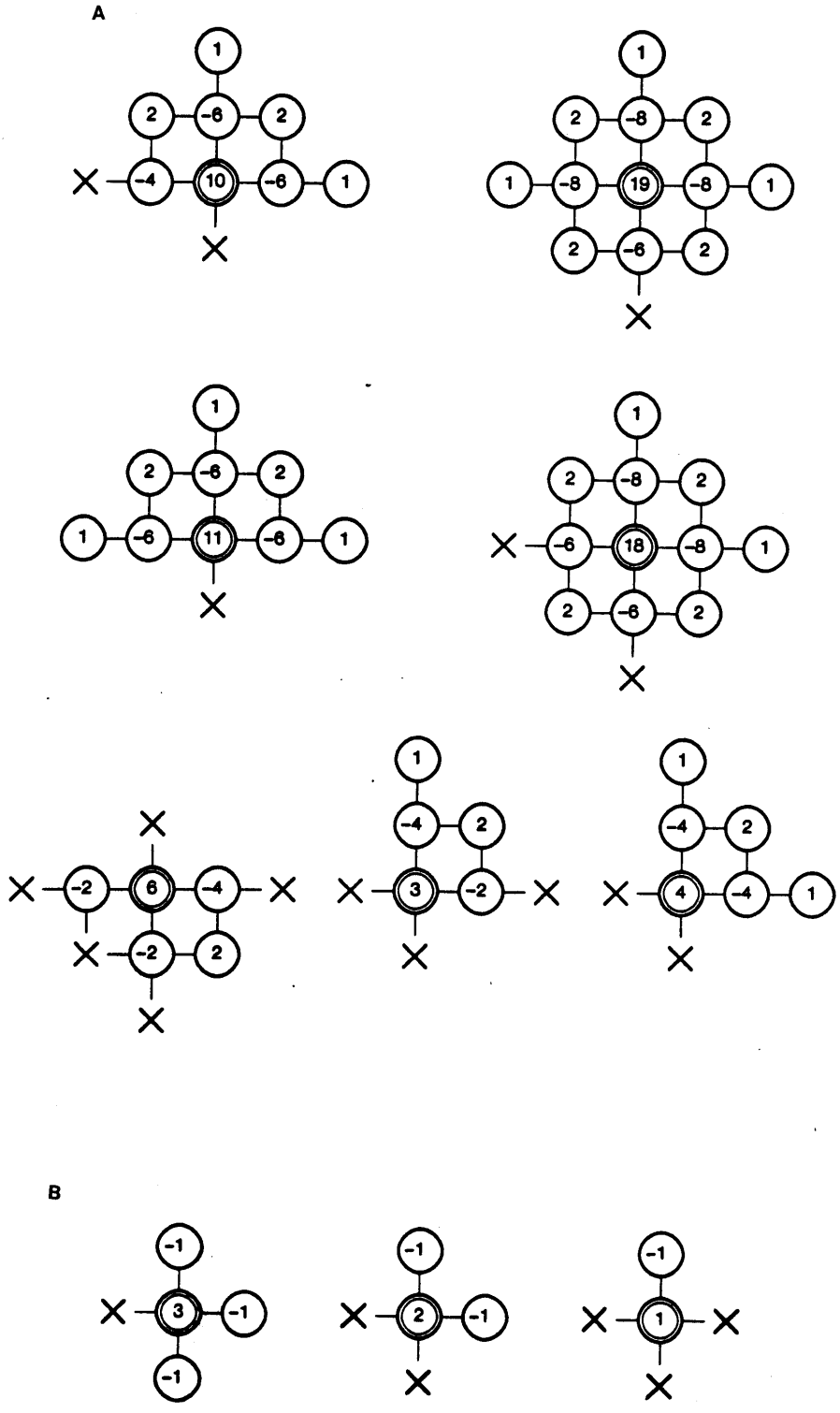

Figure 6. Molecular inhibition at discontinuities. (a) Nodal molecules at boundary nodes (double circles) near depth discontinuity nodes (X's). (b) Nodal molecules at boundary orientation discontinuities (double circles) next to depth discontinuities (X's).

the reconstructed surface. Technically, the map comprises the nodal variables $\{\rho^h_{i,j}\}$ and $\{\tau^h_{i,j}\}$ representing the discrete continuity control functions.

Any early visual process can participate in initializing the discontinuity map according to its own local hypotheses about the occurrence of discontinuities. In general, this prior discontinuity information will be partly incomplete and inconsistent, since it derives from narrowly specialized 2D image analysis. In evolving a globally consistent surface, the visible-surface reconstruction process performs a crucial task: it brings the prior discontinuity information into consonance with the 3D shape constraints collected from all the early processes. This raises the problem of detecting and localizing both depth and orientation discontinuities during reconstruction. The current section investigates this problem for the (impoverished) case in which no prior discontinuity information is available. We first propose a straightforward scheme which exploits the regularizing properties of the surface model, then a more sophisticated approach that extends the variational principle to optimally estimate discontinuities according to generic expectations about their local structure.

4.1. Regularization Based Discontinuity Detection

From one perspective, surface discontinuity detection shares much in common with traditional approaches to image intensity edge detection. In particular, it is possible to detect discontinuities by applying thresholded local differencing operations to the reconstructed surface which, like the image, is a regularly sampled function. Because they are easily corrupted by image noise, however, local edge operators such as Laplacians perform poorly [Rosenfeld and Kak, 1982] without a smoothing prefilter, say a Gaussian [Marr and Hildreth, 1980]. Interestingly, the thin plate surface under tension performs the necessary smoothing on the sparse and noisy shape constraints (standard low-pass filters such as Gaussians are inapplicable to sparse data). This regularizing effect permits the reliable computation of numerical derivatives for detecting discontinuities [Bakhvalov, 1977, Sec. 5.4; Poggio and Torre, 1984; Terzopoulos, 1985a]. In addition to exploiting the regularizing effect of the thin plate surface under tension, the discontinuity detection scheme described next is easily accommodated within the distributed computational structure of our framework, and it permits relevant criteria such as psychophysically measured limits on stereofusion to impact on discontinuity detection.

Consider the random dot stereogram in Fig. 7 which depicts a set of planar surfaces stacked in depth. Fig. 8 shows a single continuous surface generated by the surface reconstruction algorithm from sparse stereoscopic disparities provided the Marr–Poggio–Grimson (MPG) stereo algorithm [Grimson, 1985]. Fig. 9 dramatizes a portion of the reconstructed surface in cross section as it passes across a depth discontinuity. The C^1 surface overshoots constraints near the discontinuity because its smoothness conflicts with the sudden change in depth. The surface is clearly inappropriate as a final solution near depth discontinuities, but the local incompatibility can signal the occurrence of these discontinuities.

Opposing bending moments are imparted to the surface by the constraints on either side of the discontinuity. The surface inflection (see Fig. 9), where the bending moment undergoes a sign change, localizes the depth discontinuity. For a thin plate spline $u(x, y)$, the bending moment per unit length parallel to the x–z plane is proportional to $-u_{xx}$, while its counterpart parallel to the y–z plane is proportional to $-u_{yy}$ [Szilard, 1974]. The sum of orthogonal bending moments gives the total moment $M = -(u_{xx} + u_{yy}) = -\Delta u$, the negative Laplacian of the deflection function. It can be computed readily at a node (i, j) of the discrete surface using the standard approximation:

$$M_{i,j} = -\frac{1}{h^2}\left(\mathrm{u}^h_{i-1,j} + \mathrm{u}^h_{i+1,j} + \mathrm{u}^h_{i,j-1} + \mathrm{u}^h_{i,j+1} - 4\mathrm{u}^h_{i,j}\right). \tag{24}$$

Figure 7. Synthesized random dot stereogram. When fused, the stereogram depicts four planar surfaces stacked one atop the other in depth.

The zero crossings of M for the reconstructed surface in Fig. 8 are shown on the left in Fig. 10 as black contours. Most of these correspond to weak inflections due to slight ripples in the reconstructed surface. A measure of significance is therefore needed to detect true discontinuities while weeding out spurious, weak inflections. The magnitude of the local depth gradient (surface tilt) is a suitable significance measure for depth discontinuities. Hence, an inflection point will be considered significant if $G = |\nabla u| = \sqrt{u_x^2 + u_y^2}$ exceeds a limit t_d (it is more efficient to use the square of this expression $u_x^2 + u_y^2$, or even $|u_x| + |u_y|$). Employing the usual discrete approximations, we obtain

$$G_{i,j} = \frac{1}{4h^2}\left[\left(\mathrm{u}^h_{i+1,j} - \mathrm{u}^h_{i-1,j}\right)^2 + \left(\mathrm{u}^h_{i,j+1} - \mathrm{u}^h_{i,j-1}\right)^2\right]. \tag{25}$$

The right half of Fig. 10 shows the significant inflection points where $G_{i,j} \geq t_d$ with $t_d = 1$. Adding these significant points to the discontinuity map (by setting the associated $\rho^h_{i,j}$ to zero) fractures the continuous surface to yield as a solution the reconstructed stack of surfaces shown in Fig. 11.

The limit t_d must be large enough so that weak inflection points are rejected as possible discontinuities, while not so large as to miss many true depth discontinuities. A possible criterion for choosing t_d in applications to stereopsis of opaque surfaces is suggested by Panum's limiting case; i.e., when a surface is tilted so much from the viewer that it begins to occlude itself from one eye, causing stereopsis to fail. Human stereofusion limits have been measured psychophysically. Using pairs of points at different orientations, Burt and Julesz [1980] measured a roughly isotropic disparity gradient limit of approximately 1 between fusion and diplopia. Interestingly, this is only half the Panum limit.

It is not inconsistent with these findings to use the disparity gradient limit t_d to detect significant depth discontinuities in conjunction with the isotropic bending moments $M_{i,j}$ to localize these discontinuities. The required local support computations can be performed in parallel at each

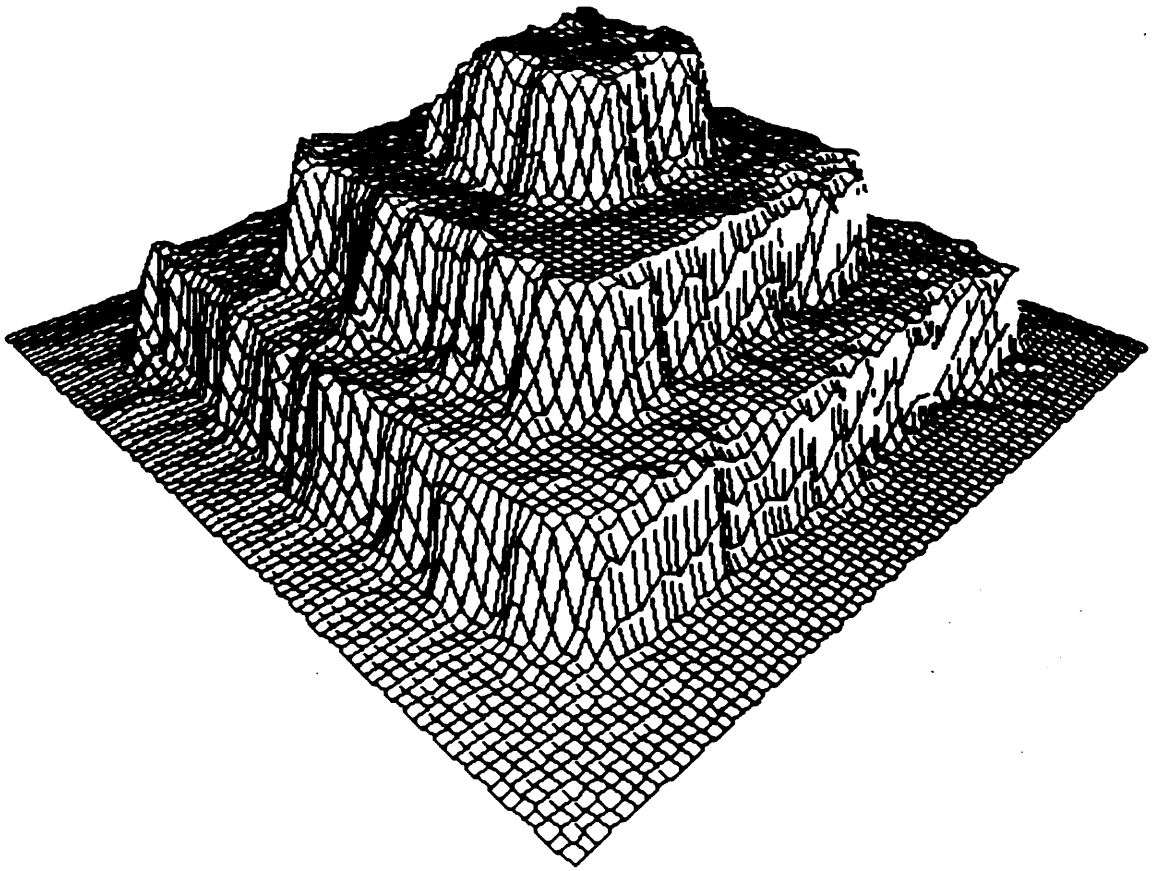

Figure 8. Single surface reconstructed from the stereogram of Fig. 7.

grid point over the surface. Analogously, significant orientation discontinuities may be detected when the magnitude of the bending moment $|M_{i,j}|$ of the surface exceeds a limit t_o (points of high curvature), and they may be localized at relative extrema of the bending moment (positions of locally highest curvature). The sign of a bending moment extremum indicates the sense of the orientation discontinuity; negative signals a concave crease, and positive, a convex crease. Curvature peaks were also employed in a scheme for detecting surface orientation discontinuities proposed by Langridge [1984].

4.2. Discontinuity Detection by Variational Continuity Control

On the one hand, experimentation on natural data with the regularized approach to discontinuity detection demonstrates the feasibility of discovering many of the more significant discontinuities during surface reconstruction (results are presented later). On the other hand, certain inherent inadequacies of this simple scheme can often lead to poor surface reconstructions. The shortcoming are due to a basic conflict caused by smoothing. While regularization eliminates noise, making reasonable estimation of surface derivatives possible in continuous regions, it tends to obscure discontinuities [Terzopoulos, 1985a]. It can result in poor detectability and localization of the more subtle discontinuities, a common problem with smoothing edge operators in general [Leclerc and Zucker, 1984].

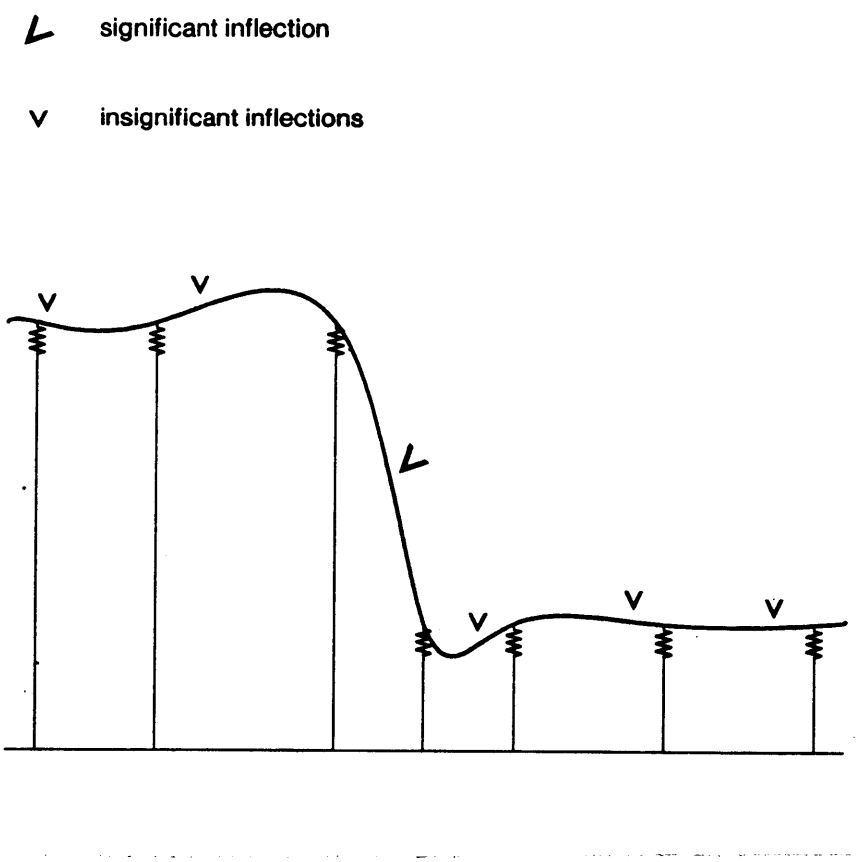

Figure 9. Cross-section of a reconstructed surface across a depth discontinuity. The significant and insignificant inflections of the surface are indicated.

The problem can be resolved by exploiting more fully the controlled continuity model to preserve surface discontinuities and, moreover, to incorporate *a priori* expectations about discontinuity structure into the variational principle for surface reconstruction. So augmented, the variational principle establishes a beneficial cooperation between the interpolation process, which smoothly propagates shape information across regions, and the complementary discontinuity process, which delimits these regions. Thus it optimally reconstructs the piecewise continuous surfaces and discontinuities simultaneously to achieve the best possible surface shape.

As was mentioned in the previous section, the smoothness of the thin plate surface under tension is incompatible with any sudden transitions imposed by the scattered shape constraints. This implies that its potential energy of deformation is generally greater at what ought to be interpreted as surface discontinuities. Any local reduction in the continuity of the surface reduces the incompatibility and locally reduces potential energy. This can be seen from (8); $\mathcal{S}_{\rho\tau}(v)$ considered as a function of (v, ρ, τ), decreases as either $\rho(x, y)$ or $\tau(x, y)$ are made zero over more of Ω. This suggests that discontinuities can be discovered in the course of solving the variational principle, by allowing the surface to crease and fracture as needed to reduce the total energy below the minimum obtainable with a single smooth surface. The insertion of discontinuities must, however, incur some energy increase, otherwise $\rho(x, y) = 0$ everywhere would trivially minimize the energy.

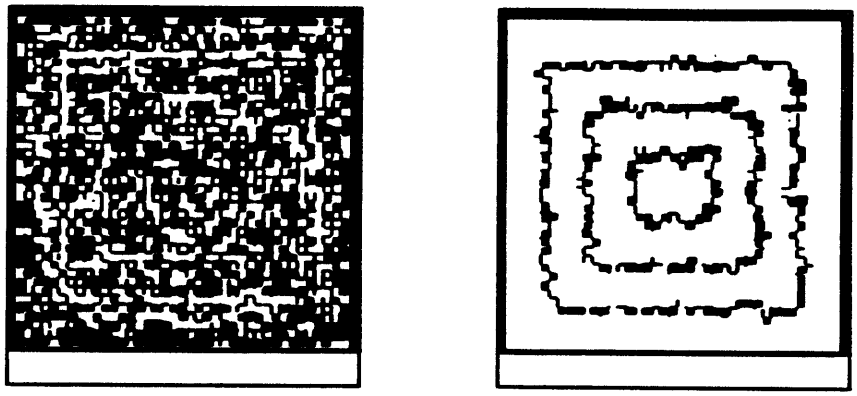

Figure 10. Bending moment zero crossings (left) and detected depth discontinuities (right) for the reconstructed surface in Fig. 8.

The variational continuity control approach to detecting discontinuities involves augmenting the original energy functional $\mathcal{E}_{\rho\tau}(v) = S_{\rho\tau}(v) + P(v)$ with the discontinuity functional $D(\rho, \tau)$ (explained shortly) to obtain the new variational principle

Find u, ρ, and τ such that

$$\mathcal{E}(u, \bar{\rho}, \bar{\tau}) = \inf_{v,\rho,\tau} \mathcal{E}(v, \rho, \tau), \tag{26}$$

where the energy functional

$$\mathcal{E}(v, \rho, \tau) = S(v, \rho, \tau) + P(v) + D(\rho, \tau). \tag{27}$$

The solutions $u(x, y)$, $\bar{\rho}(x, y)$, and $\bar{\tau}(x, y)$, satisfy the three coupled Euler–Lagrange equations, which express the vanishing of the first variation with respect to each independent function

$$\begin{aligned}
\delta_u \mathcal{E}(u, \bar{\rho}, \bar{\tau}) = 0 =& \frac{\partial}{\partial x^2}(\bar{\mu} u_{xx}) + \frac{\partial}{\partial x \partial y}(2\bar{\mu} u_{xy}) + \frac{\partial}{\partial y^2}(\bar{\mu} u_{yy}) - \frac{\partial}{\partial x}(\bar{\eta} u_x) - \frac{\partial}{\partial y}(\bar{\eta} u_y); \\
\delta_\rho \mathcal{E}(u, \bar{\rho}, \bar{\tau}) = 0 =& \bar{\tau}(v_{xx}^2 + 2v_{xy}^2 + v_{yy}^2) + [1 - \bar{\tau}](v_x^2 + v_y^2) + \delta_\rho D(\bar{\rho}, \bar{\tau}); \\
\delta_\tau \mathcal{E}(u, \bar{\rho}, \bar{\tau}) = 0 =& \bar{\rho}\Big[(v_{xx}^2 + 2v_{xy}^2 + v_{yy}^2) - (v_x^2 + v_y^2)\Big] + \delta_\tau D(\bar{\rho}, \bar{\tau}).
\end{aligned} \tag{28}$$

Note that the first equation is identical to (9).

The functional $D(\rho, \tau)$ maps the depth and orientation discontinuity configurations $\rho(x, y)$ and $\tau(x, y)$ into positive energies (this is analogous to the role of $S_{\rho\tau}(v)$ with respect to u). In its simplest form, the functional can increase monotonically with the total number of discontinuities; e.g., $D(\rho, \tau) = \int\int_\Omega \beta_d[1 - \rho(x, y)] + \beta_o[1 - \tau(x, y)]\, dx\, dy$, where β_d and β_o are positive energy scaling parameters for the depth and orientation discontinuity contributions, respectively.

More interestingly, significant advantages accrue in the detection of weak or subtle discontinuities if the functional can be designed so as to bias the solution according to generic constraints about the local structure of discontinuities. Useful constraints can, for example, be based on Gestalt principles of good continuation — discontinuities tend to be arranged along contours, these contours tend to be continuous, etc. This may be accomplished readily by assigning potential

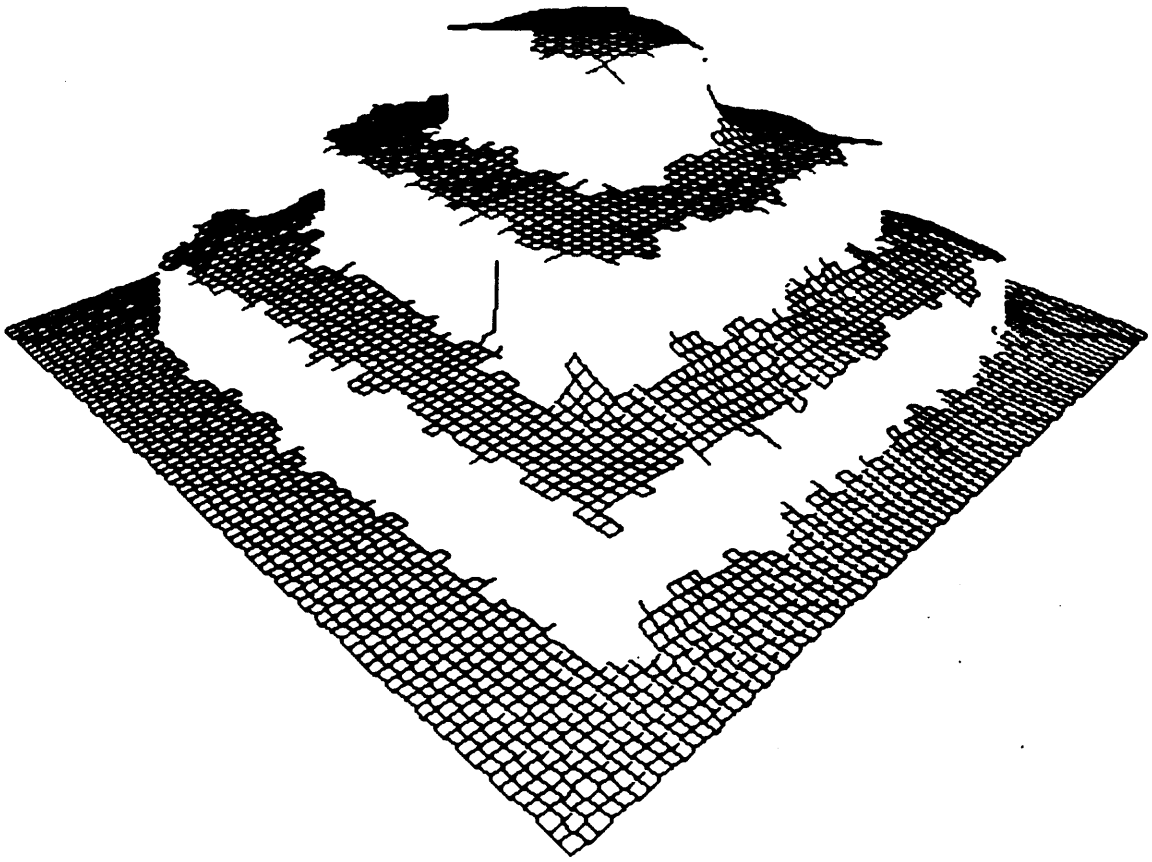

Figure 11. Reconstructed surfaces and discontinuities.

energies to various local discontinuity configurations on the (i, j) grid of nodes for the discrete problem. Encodings of local edge configurations that favor good continuation have been employed, for instance, in relaxation labeling curve enhancement processes [Zucker *et al.*, 1977] and in Markov random field image restoration models [Geman and Geman, 1985].

In our current implementation, the discrete discontinuity functional is a weighted nodal sum of potential energy quanta $D^h_{i,j}$ and $O^h_{i,j}$ over depth and orientation discontinuity configurations respectively:

$$\mathcal{D}^h(\rho^h, \tau^h) = \sum_{i,j} [\beta^h_d D^h_{i,j}(\rho^h) + \beta^h_o O^h_{i,j}(\tau^h)]. \tag{29}$$

We employ a heuristic encoding which favors the formation of continuous and smoothly curving contours by locally assigning higher energies to isolated discontinuities, terminations, sharp bends, junctions, and regions. Fig. 12 illustrates some of the configurations, and the (numeric) energies associated with them and their rotationally symmetric counterparts. Just as for computational molecules, the circles denote nodes (i, j), while the X's denote discontinuities (positions where ρ^h or τ^h are 0). The quanta in Fig. 12(a) constitute $D^h_{i,j}$ for depth discontinuities, which occur on links between nodes as explained previously. They are equivalent to the configurations found in [Geman and Geman, 1985]. Fig. 12(b) depicts some of the orientation discontinuity configurations encoded by $O^h_{i,j}$. Orientation discontinuities coincide with nodes.

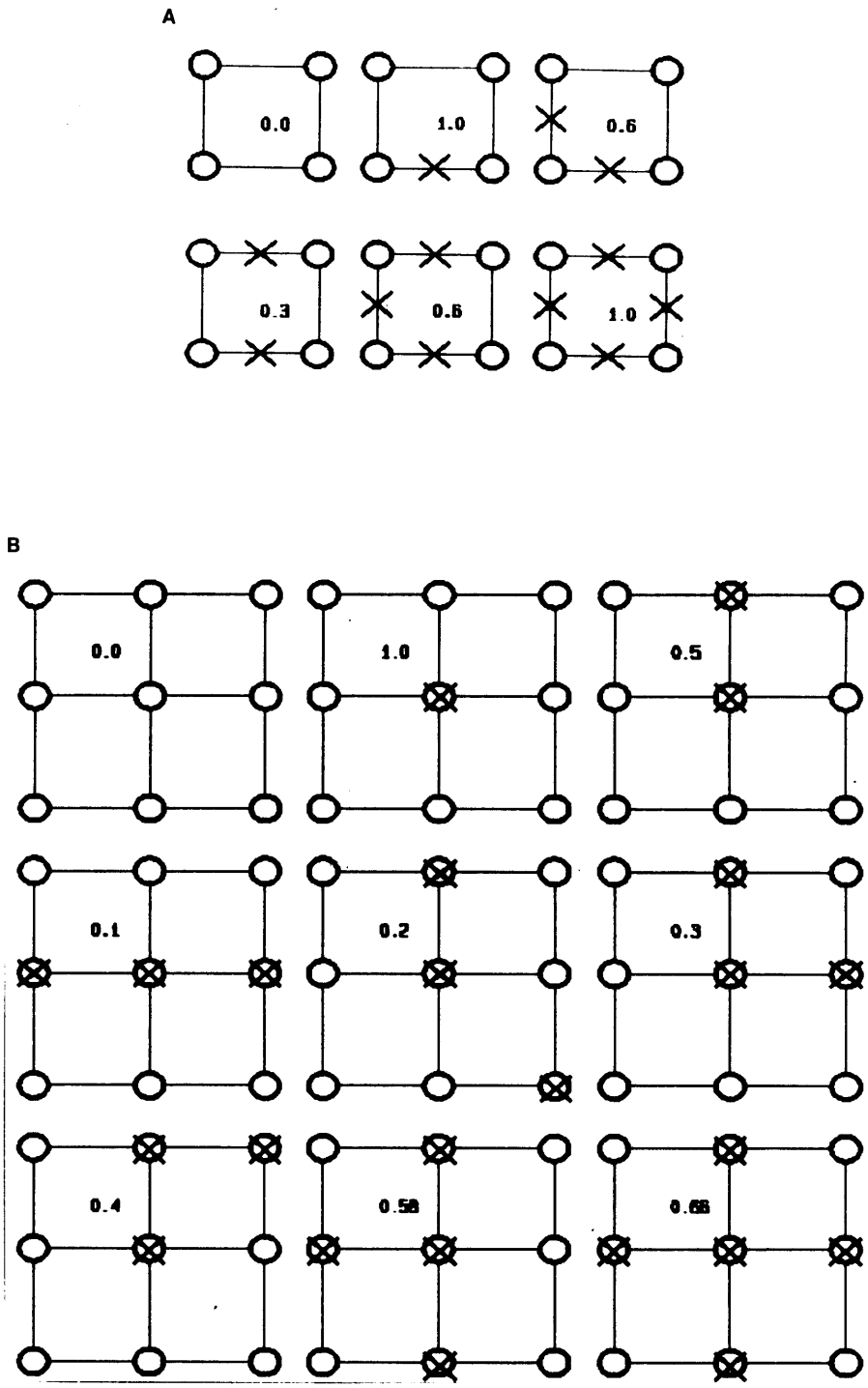

Figure 12. Some local configurations and associated energy quanta for depth discontinuities (a) and orientation discontinuities (b).

The discrete variational principle simultaneously governs the values of the displacement nodal variables of the surface as well as the nodal variables in the discontinuity map. Although the energy functional $\mathcal{E}_{\rho,\tau}(v)$ has a unique minimum (given the conditions of Sec. 2.7) for *fixed* ρ and τ, this is no longer the case for $\mathcal{E}(v,\rho,\tau)$ which allows *variation* of the continuity control functions in the

minimization. The nonconvexity of the energy landscape makes this a much more difficult problem to solve numerically. In the discontinuity detection experiments to be presented in Sec. 6.5, we propose a strategy for efficiently obtaining good, though not necessarily optimal solutions.

5. Overview of the Multiresolution Surface Reconstruction Algorithm

Application of the finite element yields the discrete problem of solving a linear system of simultaneous equations. This system has computationally desirable properties; i.e., its matrix is sparse, banded, symmetric, as well as positive definite (for fixed $\rho(x,y)$ and $tau(x,y)$) when the available constraints satisfy the conditions for a well-posed problem. The sparseness of the matrix, a direct consequence of the local support of the finite element, is evident from the nodal equation of an interior node; rows associated with interior nodes have only 13 nonzero entries, while nodes at and near discontinuities have even fewer. The $N \times N$ matrix however, tends to be extremely large in practice, since the number of pixels N in a typical image can range from 10^4 to 10^6 or greater. This combination of properties suggests the application of iterative techniques such as (parallel) Jacobi or (sequential) Gauss-Seidel relaxation methods [Hageman and Young, 1981]. Relaxation methods lead to distributed algorithms, and the parallel variants may be implemented concurrently on networks of many simple, locally-interconnected processors.

5.1. Nodal Relaxation Computations

A local-support *nodal relaxation computation* can be obtained at node (i,j) by expressing $u^h_{i,j}$ in terms of the remaining variables in the nodal equation determined by the local structure of constraints and discontinuities. The nodal relaxation computation may be constructed automatically by applying our simple rules governing the summation of basic computational molecules:

(i) Plate, depth constraint, and orientation constraint molecules sum at interior (non-discontinuity) nodes.

(ii) Membrane and depth constraint molecules sum at orientation discontinuity nodes.

(iii) Orientation discontinuities inhibit plate and orientation constraint molecules.

(iv) Depth discontinuities inhibit all basic molecules.

For instance, at a depth constraint node away from discontinuities, the Gauss-Seidel relaxation computation becomes

$$u^{(n+1)}_{i,j} = \frac{1}{\frac{20}{h^2} + \alpha_{di,j}} \left[\frac{8}{h^2} \left(u^{(n+1)}_{i-1,j} + u^{(n)}_{i+1,j} + u^{(n+1)}_{i,j-1} + u^{(n)}_{i,j+1} \right) \right.$$
$$- \frac{2}{h^2} \left(u^{(n+1)}_{i-1,j-1} + u^{(n+1)}_{i+1,j-1} + u^{(n)}_{i-1,j+1} + u^{(n)}_{i+1,j+1} \right) \qquad (30)$$
$$\left. - \frac{1}{h^2} \left(u^{(n+1)}_{i-2,j} + u^{(n)}_{i+2,j} + u^{(n+1)}_{i,j-2} + u^{(n)}_{i,j+2} \right) + \alpha_{di,j} d_{i,j} \right],$$

where we have suppressed the discretization superscript h and instead introduced the bracketed iteration indices. At an unconstrained depth discontinuity, we obtain

$$u^{(n+1)}_{i,j} = \frac{1}{4} \left(u^{(n+1)}_{i-1,j} + u^{(n)}_{i+1,j} + u^{(n+1)}_{i,j-1} + u^{(n)}_{i,j+1} \right). \qquad (31)$$

Note that the nodal relaxation computations do not change from one iteration to the next, so long as the influencing constraints or discontinuities remain unperturbed.

5.2. Multiresolution Relaxation

A serious problem with iterative techniques, in general, is their slow convergence rates for large problems. This inherent inefficiency is due to the fact that information must propagate incrementally across large representations from nodes to their near neighbors in accordance with the nodal relaxation formulas.[2] We have developed highly efficient iterative algorithms that overcome this problem for surface reconstruction [Terzopoulos, 1982, 1983a] as well as for certain other visual problems [Terzopoulos, 1984]. These algorithms achieve efficiency by exploiting multiresolution relaxation methods [Fedorenko, 1961; Brandt, 1977; Hackbusch and Trottenberg, 1982].

Briefly, the multiresolution surface reconstruction algorithm features (i) multiple representations of surface shape over a range of spatial resolutions, (ii) local, iterative (relaxation) processes that propagate smoothness constraints within each representational level, (iii) local coarse-to-fine (prolongation) processes that allow coarser representations to constrain finer ones, (iv) fine-to-coarse (restriction) processes that allow finer representations to constrain and improve the accuracy coarser ones, and (v) a multilevel coordination strategy that enables the hierarchy of representations and component processes to cooperate towards increasing the computational efficiency, usually by orders of magnitude.

Fig. 13 depicts the structure of the algorithm schematically. In this particular case, only three levels are shown. Note the 2:1 resolution reduction between adjacent levels. Not only does this ratio simplify the component processes considerably, but it is also nearly optimal with regard to total computation to convergence (this is conveniently measured in machine independent *work units*, where a work unit is the amount of computation required for a relaxation iteration on the finest level) [Brandt, 1977]. The diagram illustrates the intralevel relaxation processes, as well as the fine-to-coarse restriction and coarse-to-fine prolongation processes that communicate between levels. The figure shows synthetically generated scattered orientation and depth constraints consistent with a hemispherical surface. The algorithm reconstructs a dense representation of surface at three resolutions. The sparse information at any particular scale can be thought of as a set of constraints which defines a discrete surface approximation problem at that level. It is natural then to view the multiresolution surface reconstruction algorithm as iteratively solving a coupled hierarchy of discrete surface reconstruction problems.[3] For a detailed description of the algorithm see [Terzopoulos, 1982, 1983a].

6. Experimental Analysis of the Algorithm

The multiresolution visible-surface reconstruction algorithm was tested on a variety of data sets including synthetic data, structured light (laser) range data, automated stereopsis and photometric stereo data from natural images, and digital terrain model data. Some results are presented in this section (for further details and examples, see [Terzopoulos, 1984]). In all the examples

[2] It is possible to accelerate the basic relaxation methods so that fewer iterations are required. However, practical accelerated methods such as the conjugate gradient method, successive overrelaxation, and Chebyshev semi-iteration use global procedures to determine the acceleration parameters. In parallel implementation, the greater complexity of the globally accelerated methods and, even more importantly, the communications costs of performing the global operations nullifies any potential gains.

[3] A recursive multilevel coordination strategy was employed in the experiments described next. The recursive strategy activates only a single level at any one time. We have recently developed a concurrent strategy based on a multilevel variational principle [Terzopoulos, 1985b]. Concurrent coordination maintains all levels active simultaneously, thus achieving full parallelism.

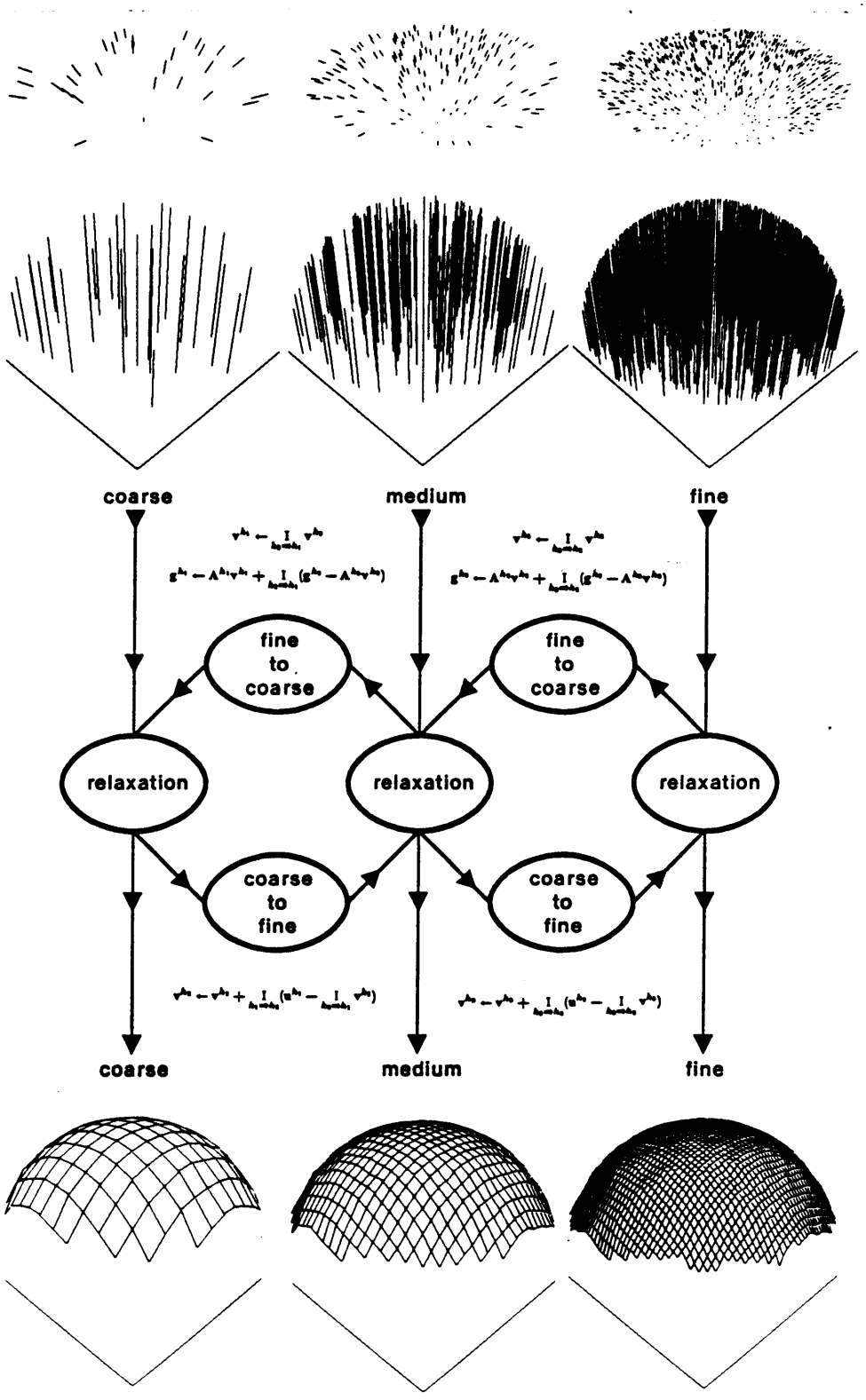

Figure 13. The structure of the multiresolution surface reconstruction algorithm. Iterative relaxation processes operate at each level. Fine-to-coarse and coarse-to-fine processes transfer information between levels. Synthetic orientation and depth constraints input to the algorithm are shown at the top. The dense multiscale surface representation is output at the bottom.

presented, the intralevel process was Gauss-Seidel relaxation and the algorithm was started from zero initial approximations on all the levels. The border nodes on each grid were preset as depth discontinuity nodes to introduce natural boundary conditions which free the reconstructed surface on the boundary of Ω.

6.1. Synthetic Data

The first two examples involve randomly placed depth constraints. The left half of Fig. 14 shows 15%-density constraints at three resolutions. These constraints were obtained by sampling a hemisphere whose z values were multiplied by a radial sinusoid. The nodes outside the circular region occupied by the constraints were specified as depth discontinuities. The reconstructed surface representation is shown on the right half of Fig. 14. In Fig. 15, the 15%-density depth constraints shown on the left are samples of a stacked set of planar surfaces at three resolutions. In this example, depth discontinuities were placed along the circular arcs bounding the planes, and along the outer edges of the grids. The reconstructed surface representation is shown on the right half of Fig. 15. This example indicates that discontinuities can be placed along arbitrary contours within Ω to prevent surface shape from being degraded by unwanted smoothing over sharp depth changes.

The next examples involve reconstructions from orientation constraints. The left half of Fig. 16 shows in perspective a set of orientation constraints over a square region. On each of three scales, the region is divided into four quadrants each containing constant orientation constraints, and the nodes along their boundaries are preset as orientation discontinuities. The surfaces reconstructed by the three-level algorithm are shown on the right half of the figure. Since absolute depth cannot be determined solely from orientation constraints, a relative depth reconstruction results, with the center of gravity of the resulting pyramidal surface resting near the x–y plane.

The left half of Fig. 17 shows 30%-density scattered orientation constraints consistent with a hemispherical surface at three resolutions. The reconstructed surface representation is shown on the right. All nodes outside the hemispherical surface patch were specified as depth discontinuities. Again, the center of gravity of the surface rests near the x–y plane.

The next examples demonstrate the integration of both depth and orientation constraints. The left half of Fig. 18 shows 15%-density depth constraints consistent with a hemispherical surface at three resolutions. On the right are 15%-density orientation constraints consistent with the same surface. Nodes outside the surface have been specified as depth discontinuities. The reconstructed surface is shown in Fig. 19. Whereas in the previous example (Fig. 17) only relative depth can be determined for lack of any depth constraints, in the present example the additional depth constraints enable the absolute depth of the surface to be determined at all points, hence the surface is "raised" to the correct height above the base plane. In addition, note that (10%) uniformly distributed noise has been added to the constraint values. With the given constraint parameters, the surface is slightly bumpy on the finest level. This can be reduced by decreasing the constraint parameters, in effect, loosening the springs of the physical model.

6.2. Structured Light Data

The multiresolution algorithm was applied to the reconstruction of several objects from raw range data supplied by a laser rangefinder constructed by P. Brou at MIT. The scan resolution in the y direction is half that in the x direction. A four-level surface reconstruction algorithm was employed in the examples. The data was introduced as depth constraints at the finest level and transferred to the coarser levels by successive 2×2 averaging between levels. To expediently segment the objects from the background, values smaller than a threshold were treated as depth

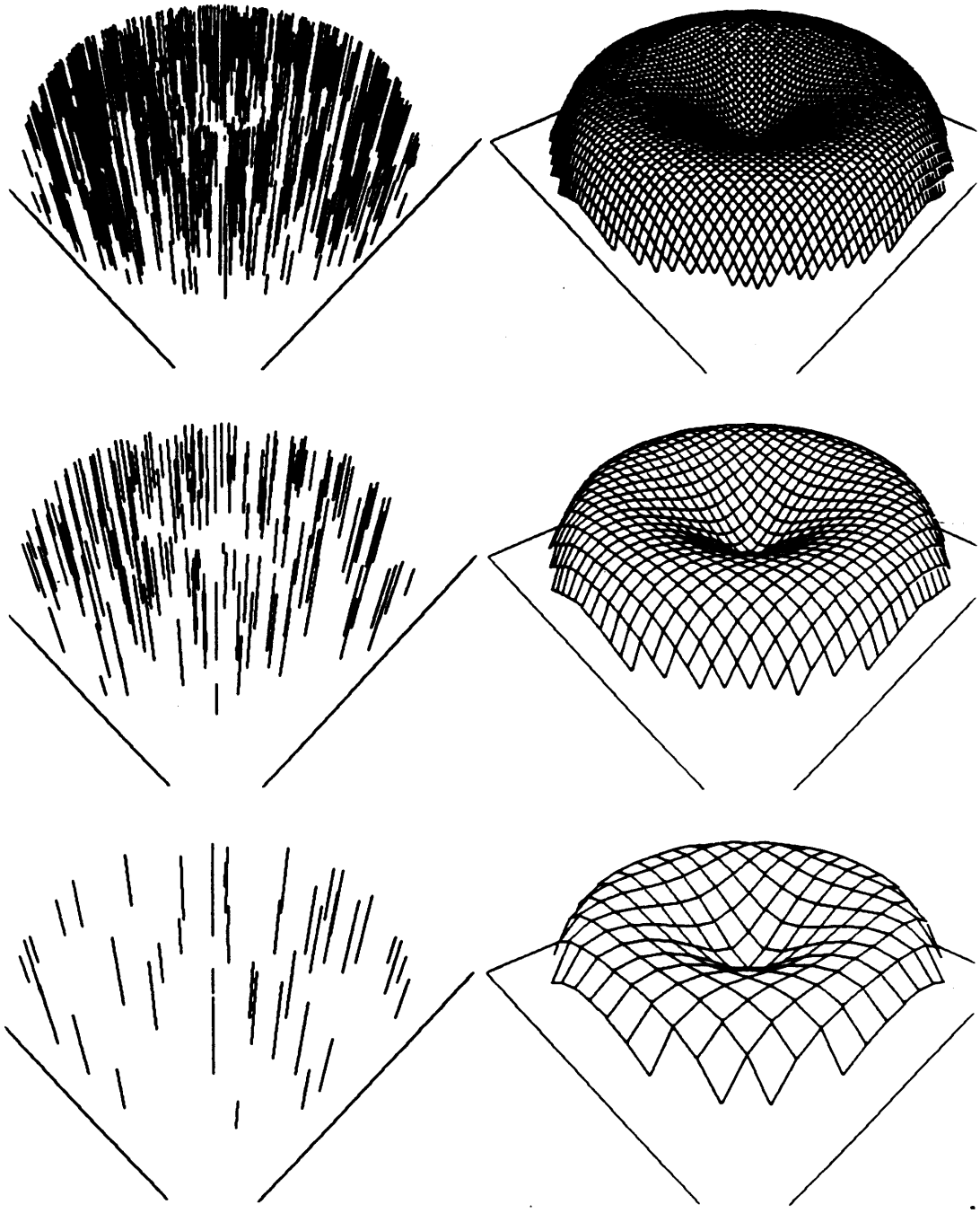

Figure 14. Reconstruction of a surface from depth constraints. (Grid dimensions: $N_x^{h_1} = N_y^{h_1} = 17$, $N_x^{h_2} = N_y^{h_2} = 33$, $N_x^{h_3} = N_y^{h_3} = 65$. Grid spacings: $h_1 = 0.4$, $h_2 = 0.2$, $h_3 = 0.1$. Constraint parameters: ${\alpha_d}^{h_j} = 2.0/h_j$. Computation: 24.25 work units.)

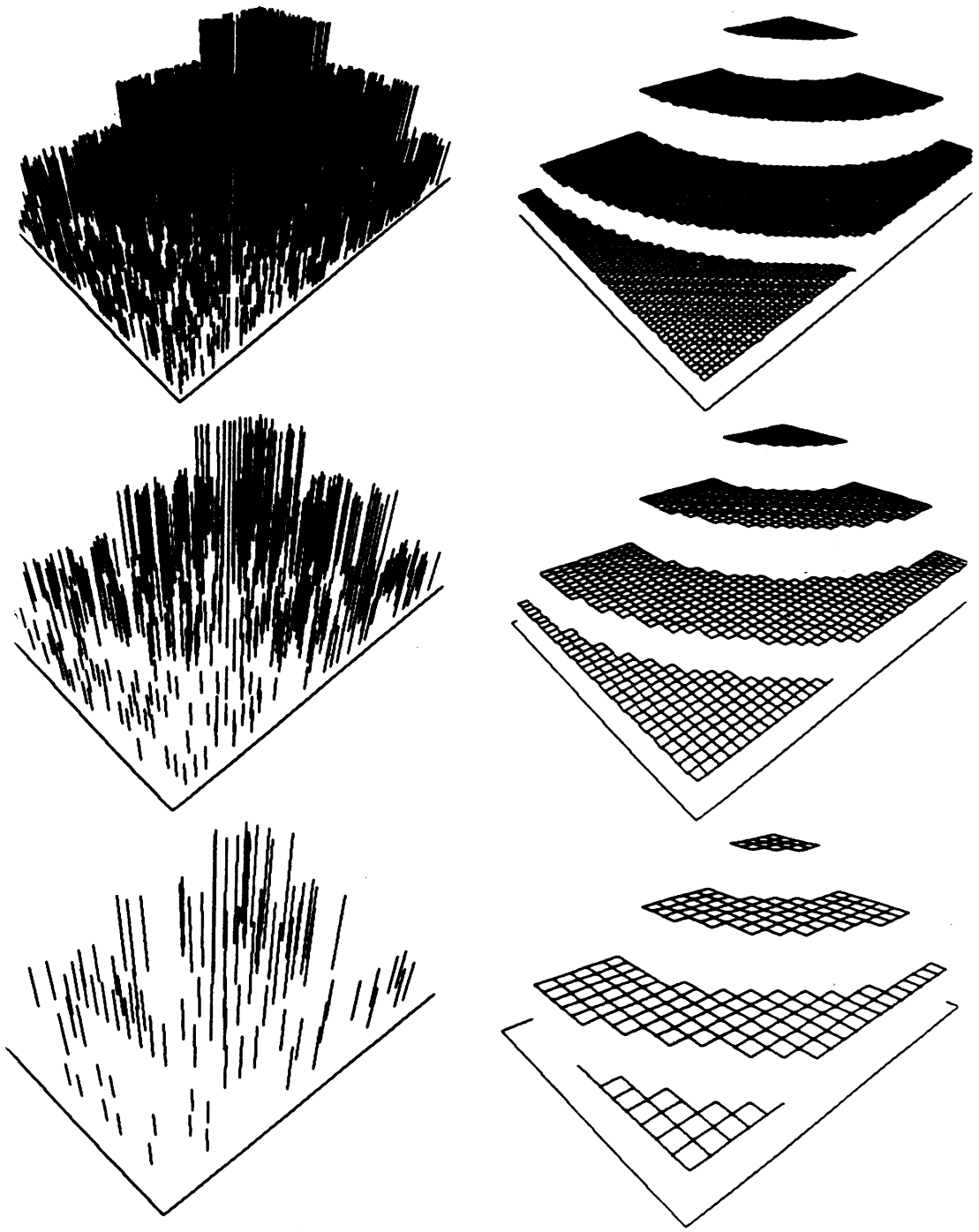

Figure 15. Reconstruction of planes with circular depth discontinuities from depth constraints. (Grid dimensions: $N_x^{h_1} \times N_y^{h_1} = 22 \times 17$, $N_x^{h_2} \times N_y^{h_2} = 43 \times 33$, $N_x^{h_3} \times N_y^{h_3} = 85 \times 65$. Grid spacings: $h_1 = 0.4$, $h_2 = 0.2$, $h_3 = 0.1$. Constraint parameters: ${\alpha_d}^{h_j} = 2.0/h_j$. Computation: 20.375 work units.)

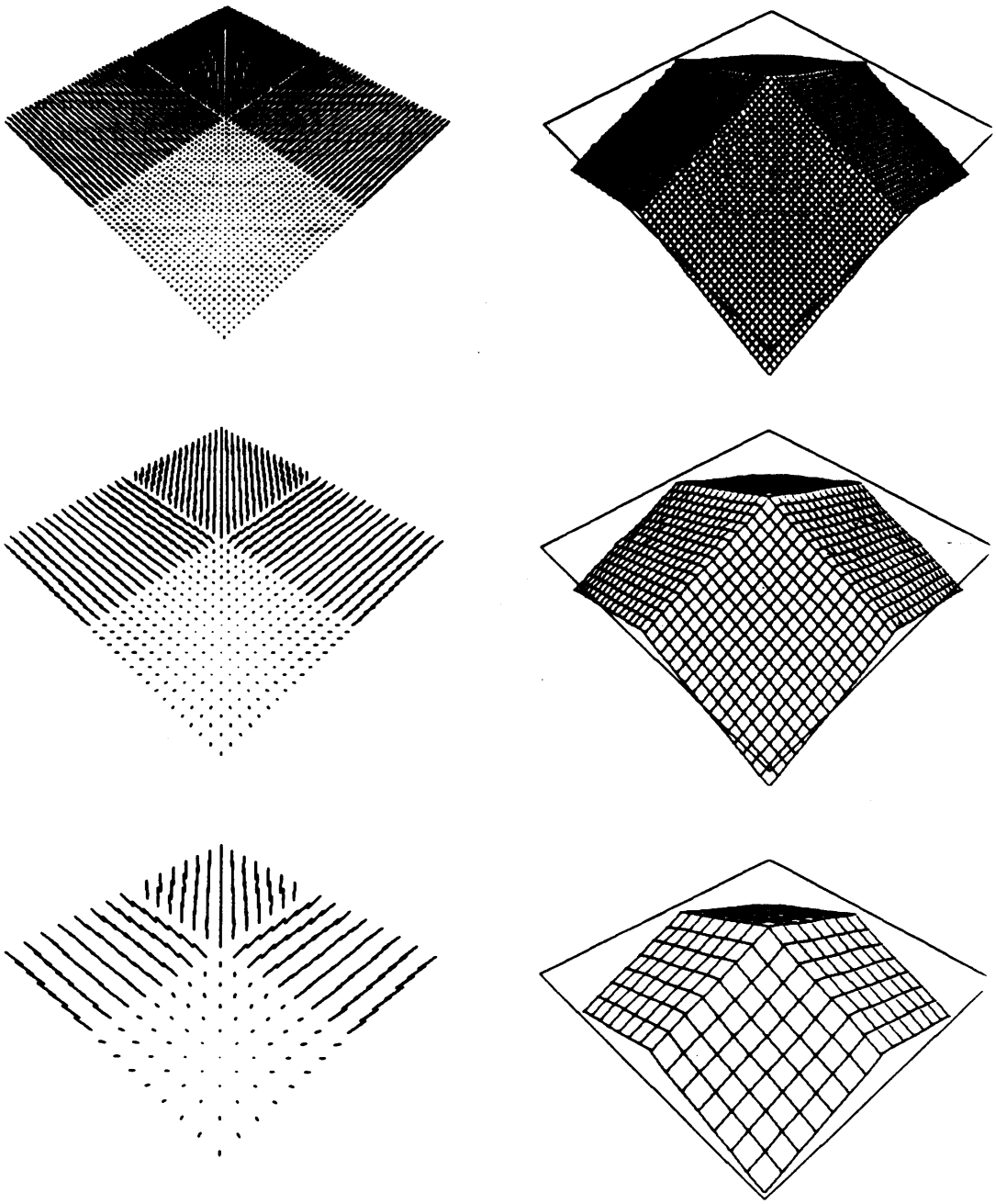

Figure 16. Reconstruction of a pyramidal surface with orientation discontinuities from orientation constraints. (Grid dimensions: $N_x^{h_1} = N_y^{h_1} = 17$, $N_x^{h_2} = N_y^{h_2} = 33$, $N_x^{h_3} = N_y^{h_3} = 65$. Grid spacings: $h_1 = 0.4$, $h_2 = 0.2$, $h_3 = 0.1$. Constraint parameters: $\alpha_p^{h_j} = \alpha_q^{h_j} = 4.0/h_j$. Computation: 19.5 work units.)

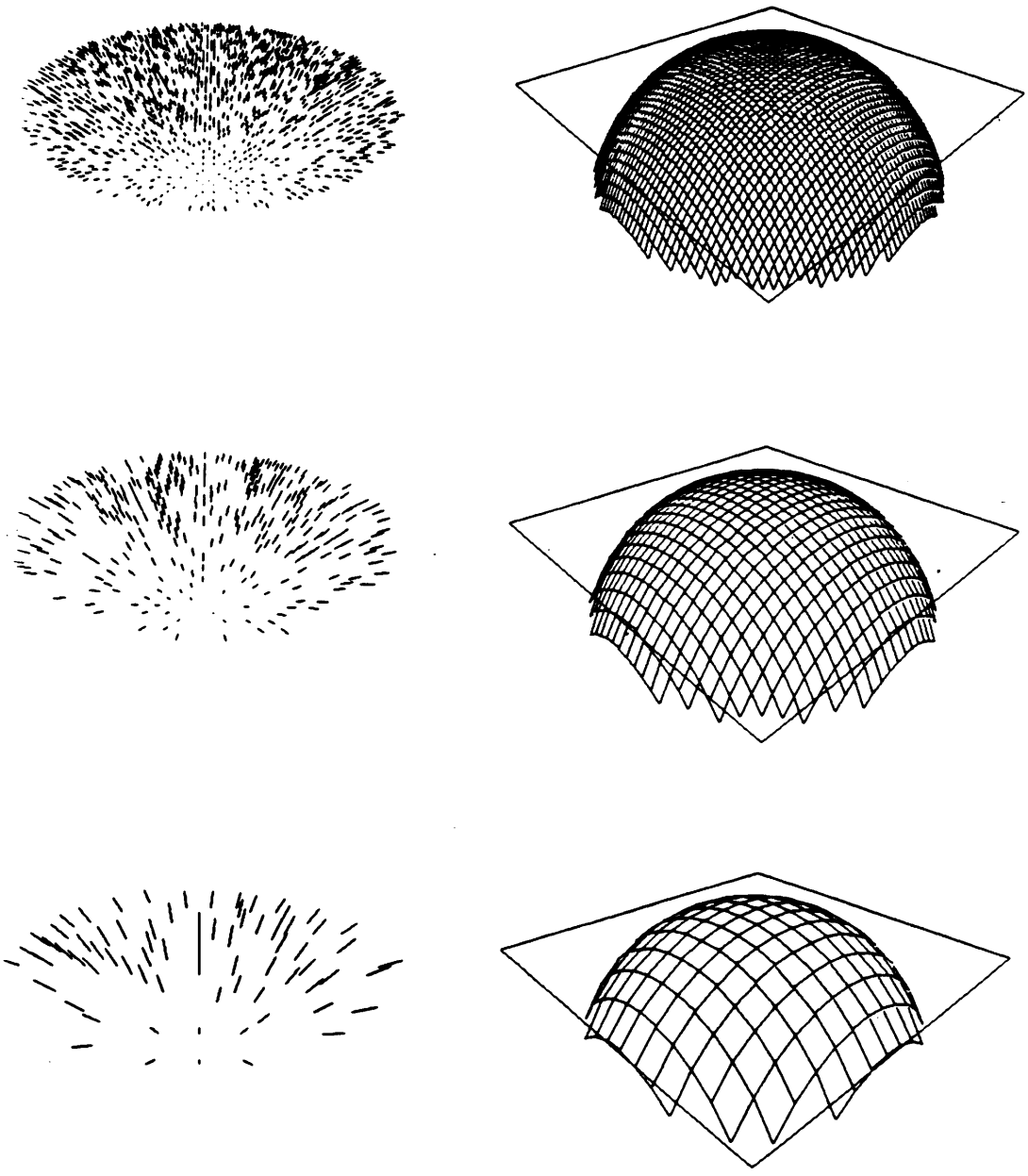

Figure 17. Reconstruction of a hemispherical surface from scattered orientation constraints. (Grid dimensions: $N_x^{h_1} = N_y^{h_1} = 17$, $N_x^{h_2} = N_y^{h_2} = 33$, $N_x^{h_3} = N_y^{h_3} = 65$. Grid spacings: $h_1 = 0.4$, $h_2 = 0.2$, $h_3 = 0.1$. Constraint parameters: $\alpha_p^{h_j} = \alpha_q^{h_j} = 4.0/h_j$. Computation: 22.125 work units.)

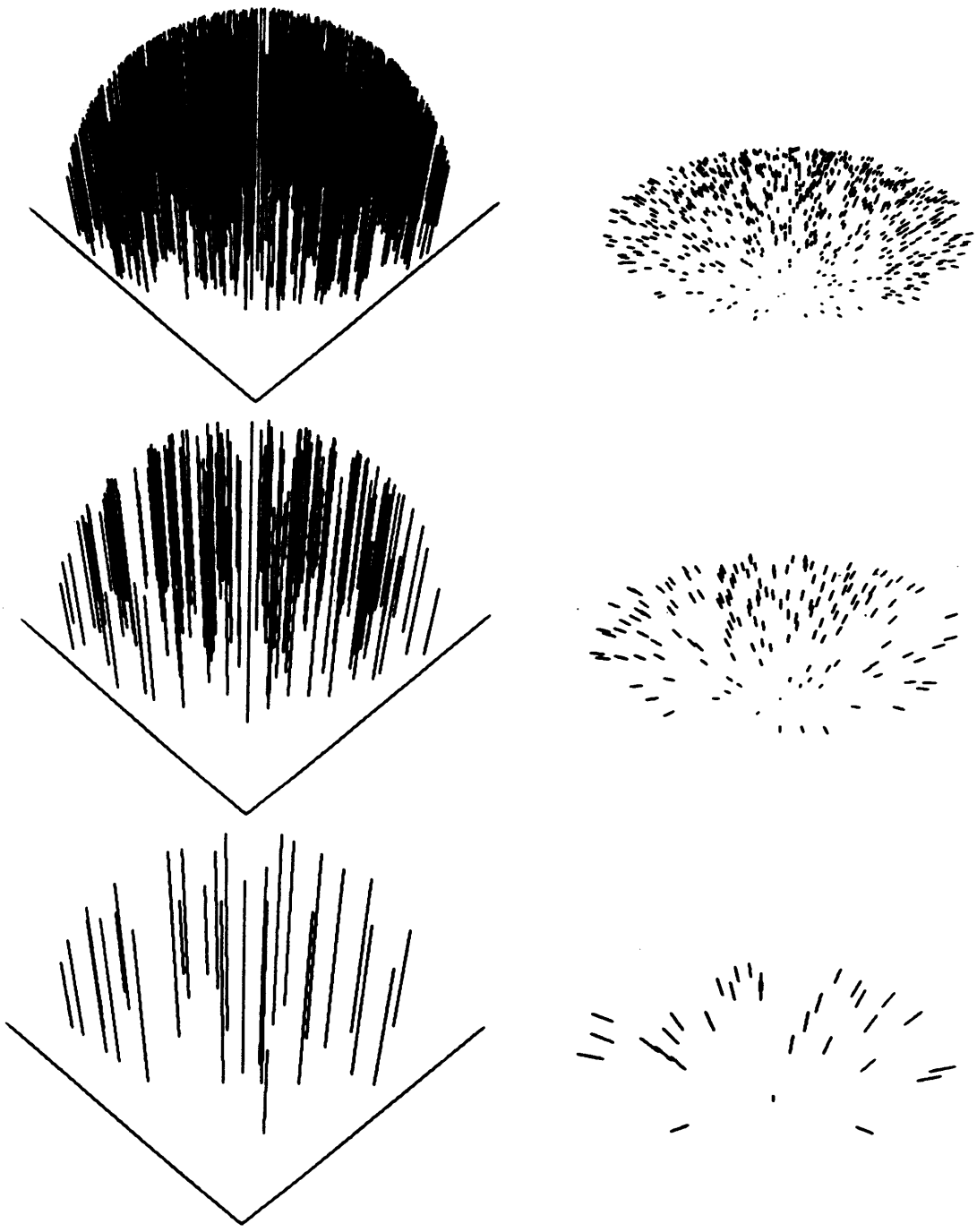

Figure 18. Depth constraints (left) and orientation constraints (right) consistent with a hemisphere at three resolutions.

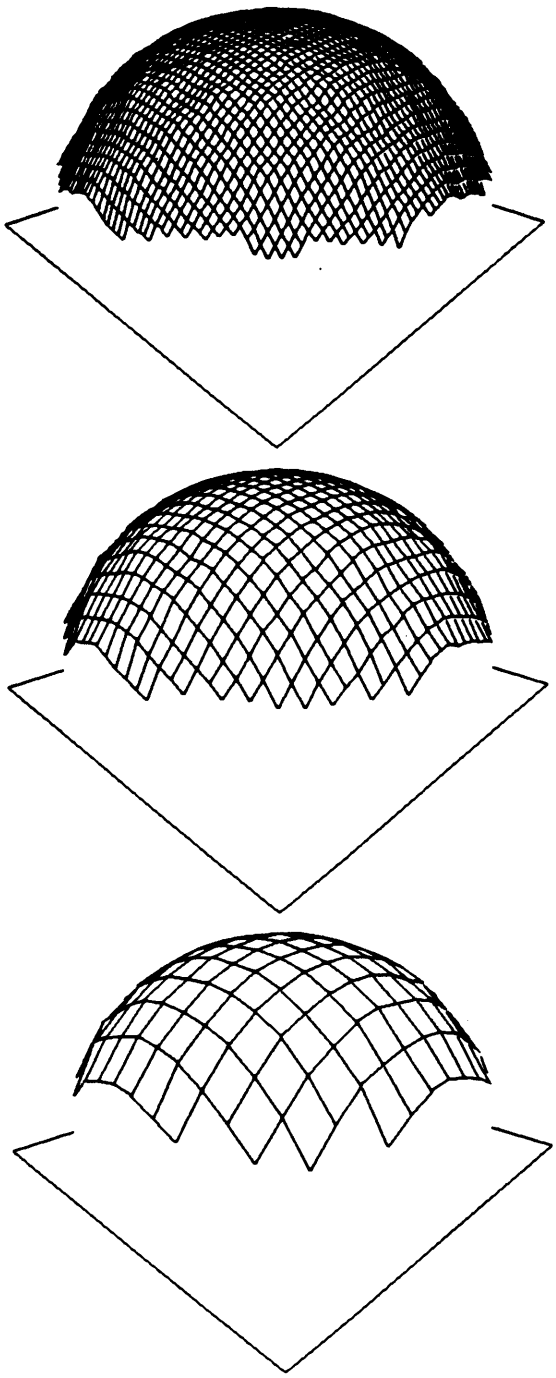

Figure 19. Reconstruction from depth and orientation constraints in Fig. 18. (Grid dimensions: $N_x^{h_1} = N_y^{h_1} = 17$, $N_x^{h_2} = N_y^{h_2} = 33$, $N_x^{h_3} = N_y^{h_3} = 65$. Grid spacings: $h_1 = 0.4$, $h_2 = 0.2$, $h_3 = 0.1$. Constraint parameters ${\alpha_d}^{h_j} = 2.0/h_j$, $\alpha_p^{h_j} = \alpha_q^{h_j} = 4.0/h_j$. Computation: 17.75 work units.)

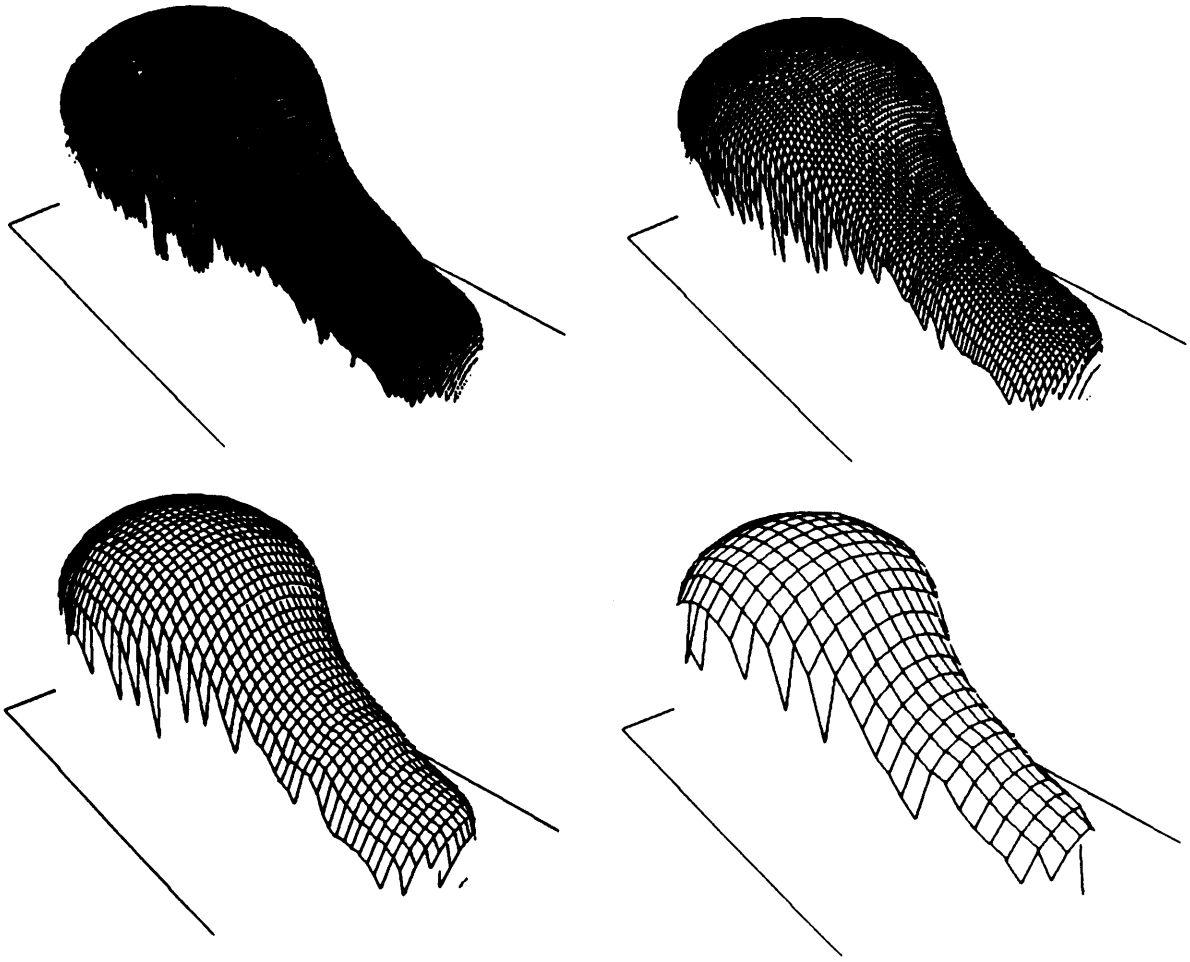

Figure 20. Reconstruction of a lightbulb from range data. (Finest grid dimensions: $N_x^{h_1} \times N_y^{h_1} = 257 \times 281$. Grid spacings: $h_1 = 0.8$, $h_2 = 0.4$, $h_3 = 0.2$, and $h_4 = 0.1$. Constraint parameters: ${\alpha_d}^{h_j} = 0.2/h_j$. Computation: 9.78 work units.)

discontinuities. Fig. 20 shows the reconstructed surface of a lightbulb. The algorithm smoothes the noise in the data and reconstructs the missing points.

6.3. Natural Image Data

In this section, we apply the multiresolution surface reconstruction algorithm to depth data originating from natural images. The examples involve photometric stereo, and two binocular stereo algorithms applied to terrain stereopairs.

6.3.1. Photometric Stereo Data

Photometric stereo is a technique that uses multiple (usually 3) images of a scene from the same viewpoint, but with differing illumination [Woodham, 1981]. Assuming that the surface material

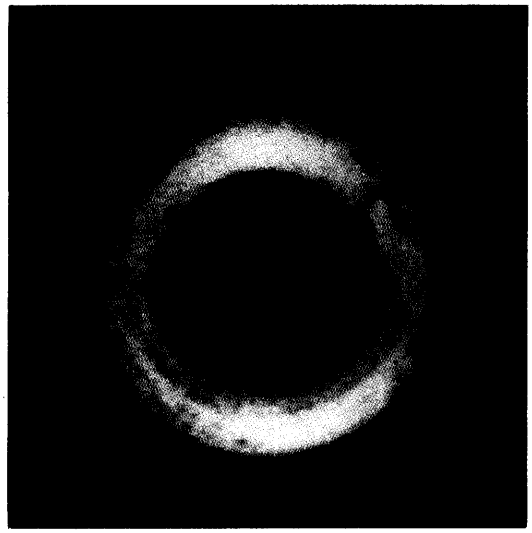

Figure 21. Image of a matte white torus.

is known and that the viewer and lightsources are far from the object, the method determines the surface orientation from the image irradiance. Our surface reconstruction algorithm provides a noise resistant technique for computing depth from the surface orientation data provided by photometric stereo. We demonstrate this with the image of torus in Fig. 21. The photometric stereo data (generated by a system implemented at MIT by K. Ikeuchi) was introduced as orientation constraints on a two-level algorithm. Aside from sporadic missing data, the constraints on the coarse level are dense, whereas only every other node on the fine level is a constraint. Fig. 22 shows the orientation data and the reconstructed torus.

Our method for reconstructing surfaces from scattered orientation constraints can be compared to a variational scheme for obtaining relative depth from dense surface gradient information reported by Horn and Brooks [1985]. Their proposed least squares integral $\int\int (v_x - p)^2 + (v_y - q)^2 \, dx \, dy$ will be recognized as being a continuous version of the orientation constraint penalty functional. By virtue of the additional smoothness functional $S_{\rho\tau}(v)$, however, our surface reconstruction algorithm can deal with orientation constraints that are scattered. It also can integrate depth constraints from other sources to arrive at absolute surface depth.

6.3.2. Correlation Based Stereo Data

At the top of Fig. 23 is a stereopair on which Kass's [1983] correlation based stereo algorithm was run. The output of the stereo algorithm is shown on the lower left, with brightness proportional to disparity. The algorithm has failed to produce a match in the neutral grey patches, so disparity is unknown in these areas. To apply the multiresolution algorithm, the disparity data on the finest level were reduced by factors of two, through averaging, to three coarser levels. Relatively small constraint parameter values were chosen in order to counteract the potentially detrimental effects of false matches and noise in the disparity data. The reconstructions on the three coarsest levels are shown as 3D plots in Fig. 24 (the finest level was too dense to represent this way). Fig. 25 shows isoelevation contour maps of the solution on all levels.

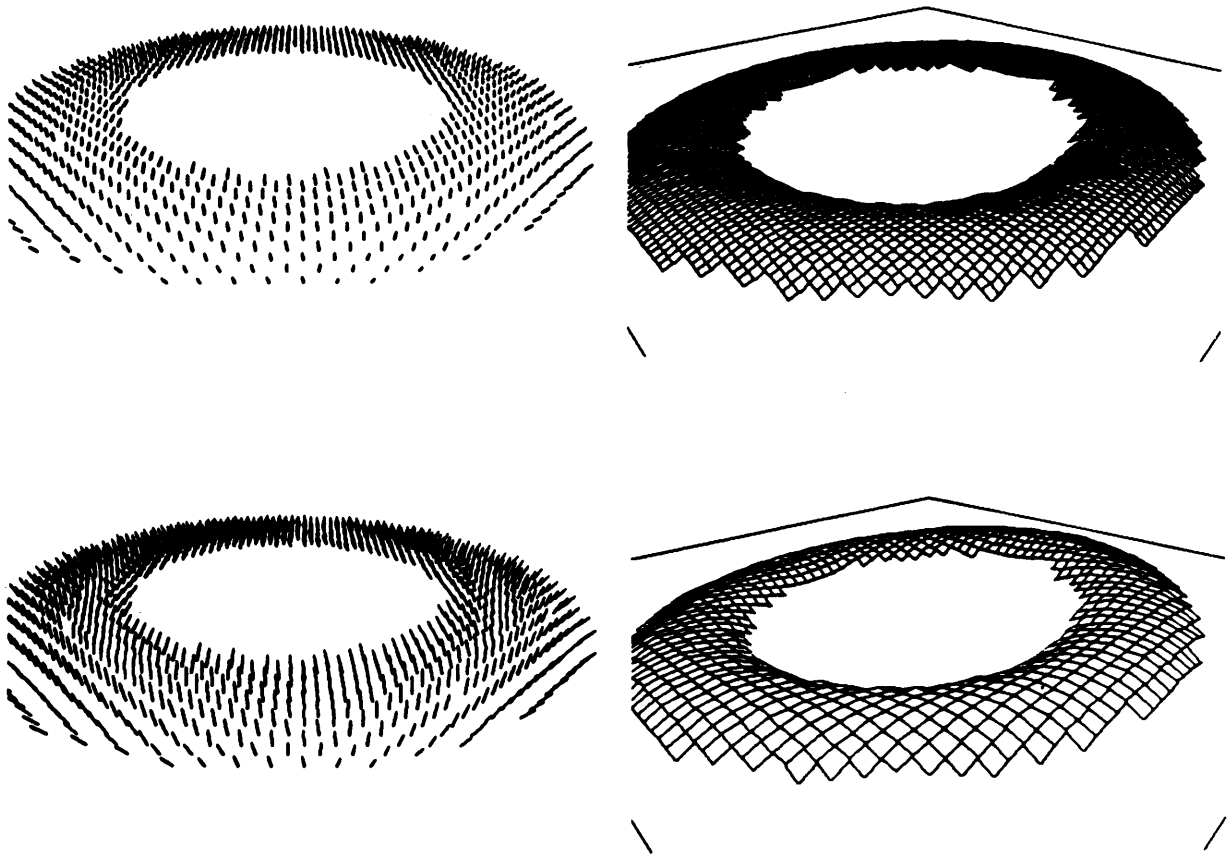

Figure 22. Reconstruction of the torus (right) from the orientation constraints provided by photometric stereo (left). (Grid dimensions $N_x^{h_1} = N_y^{h_1} = 51$ and $N_x^{h_2} = N_y^{h_2} = 101$. Constraint parameters: $\alpha^{h_j} = 4.0/h_j$. Computation: 52.0 work units.)

6.3.3. Feature Based Stereo Data

The next example involves disparity constraints generated by the MPG stereo algorithm [Grimson, 1985]. A three-channel version of the stereo algorithm was run on the stereopair at the top of Fig. 26. The output of the stereo algorithm is shown on the lower part of the figure. Disparity information is provided only along zero crossing contours at the three finest scales. In the figure, the darkness along contours is proportional to disparity. This disparity data was input to a four-level surface reconstruction algorithm. The constraints on the coarsest level were derived by averaging the constraints from the next finer level. The reconstructions on the three coarsest levels are shown as 3D plots in Fig. 27. Fig. 28 shows isoelevation contour maps of the solution on all levels.

6.4. Digital Terrain Map Data

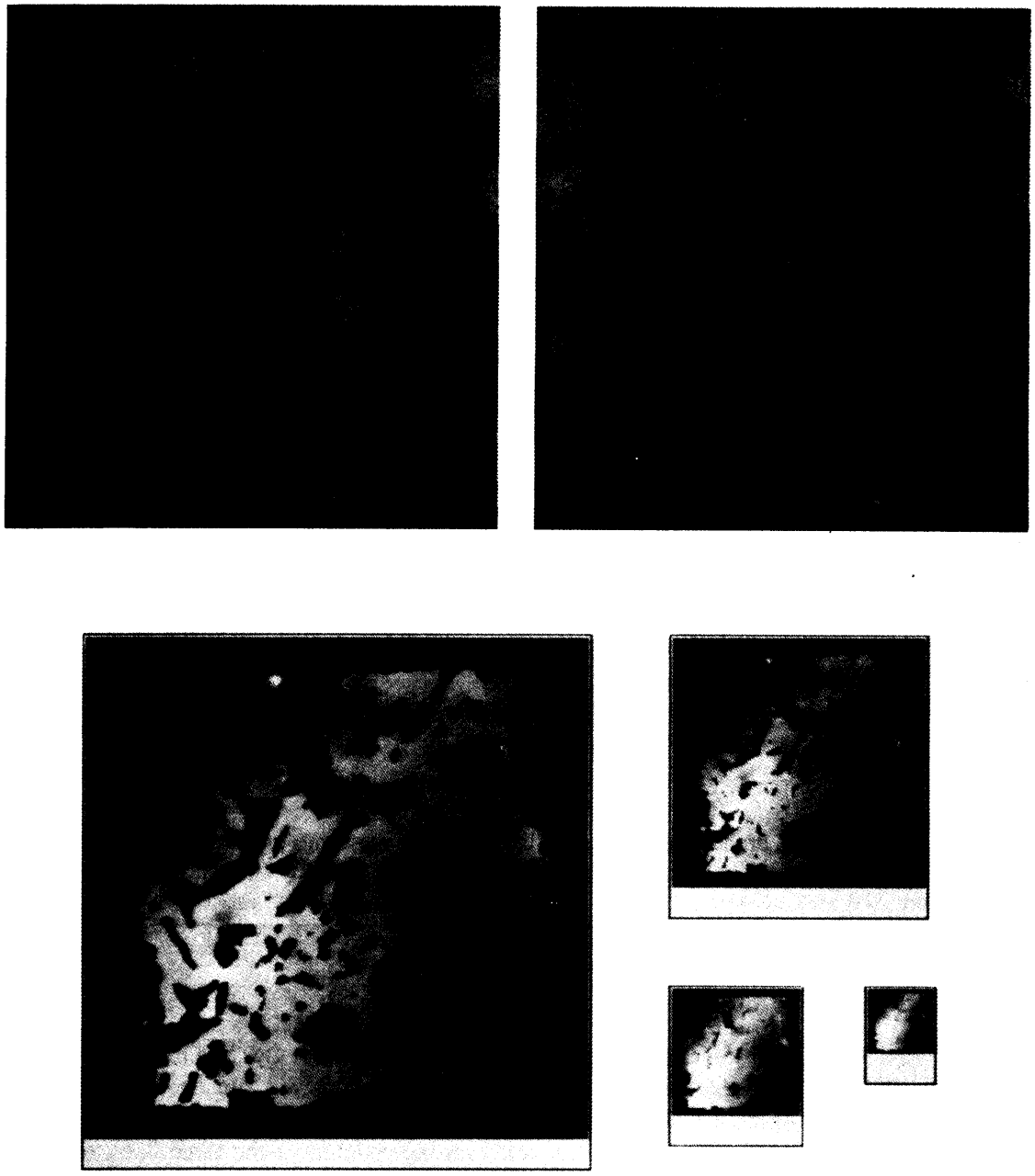

Figure 23. Natural terrain stereopair (top) and output of Kass' stereo algorithm (bottom). The images were 256 × 256 pixels, quantized to 256 levels (provided by the US Defense Mapping Agency).

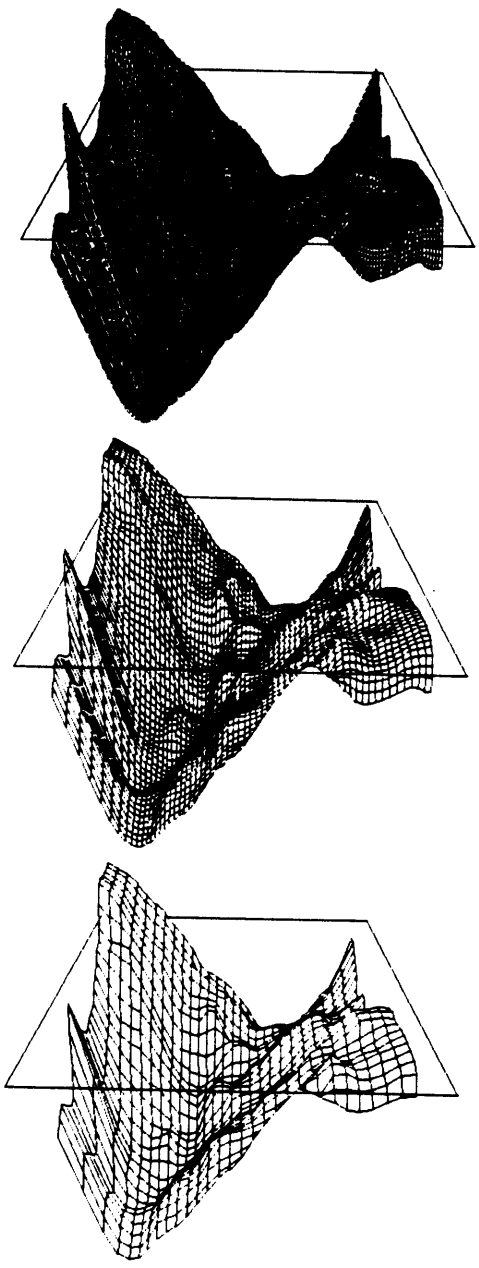

Figure 24. Reconstruction of terrain in Fig. 23. (Grid dimensions: $N_x^{h_1} = N_y^{h_1} = 33$, $N_x^{h_2} = N_y^{h_2} = 65$, $N_x^{h_3} = N_y^{h_3} = 129$, $N_x^{h_4} = N_y^{h_4} = 257$. Grid spacings: $h_1 = 0.8$, $h_2 = 0.4$, $h_3 = 0.2$, $h_4 = 0.1$. Constraint parameters: ${\alpha_d}^{h_j} = 0.01/h_j^2$. Computation: 29.0 work units.)

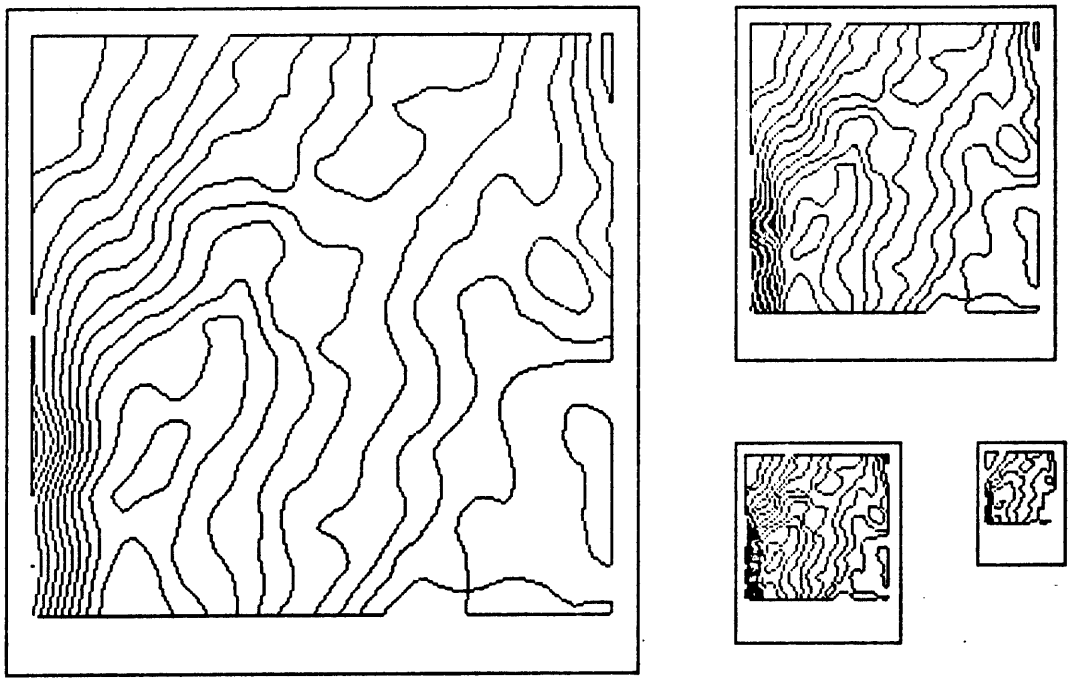

Figure 25. Isoelevation contour maps of the reconstructed terrain in Fig. 24.

A four-level surface reconstruction algorithm was applied to contoured terrain elevation data. A contour map of the Black River Gorges (published by the UK Ministry of Defense) was digitized manually on a digitizing tablet by J. Mahoney. The 256 × 256 digital contour array is shown at the top of Fig. 29. The constraints input to the algorithm are shown at the bottom of the figure. The elevation of the contours is proportional to brightness. Local averaging was used to derive the constraints on the coarser grids from those on the finest grid. The terrain reconstructions on the three coarsest levels are shown as 3D plots in Fig. 30 (the finest level is too dense to represent this way). Fig. 31 shows isoelevation contour plots of the reconstructed terrain on all levels. The reconstructed contours on the finest level can be compared subjectively with the digitized contours in Fig. 29, but note the reconstructed contours depict elevations half way between the original constraint contours for an unbiased comparison. The reconstructed contours are somewhat smoother than the (predigitized) contours in the original map — the jaggedness introduced by manual digitization has been reduced. The extent of the smoothing can be regulated by adjusting the constraint parameters. Shaded image renditions of the reconstructed terrain using reflectance map techniques for hill shading [Horn, 1981] are shown at the bottom of Fig. 31. Terrain reconstructions using the thin plate surface under tension model were compared to reconstructions using the simpler membrane spline model (Laplacian smoothing). The former gives good results, whereas the latter generally suffers from insufficient smoothness and produces flat spots across terrain peaks [Terzopoulos, 1984] (see also [Bolondi *et al.*, 1976]).

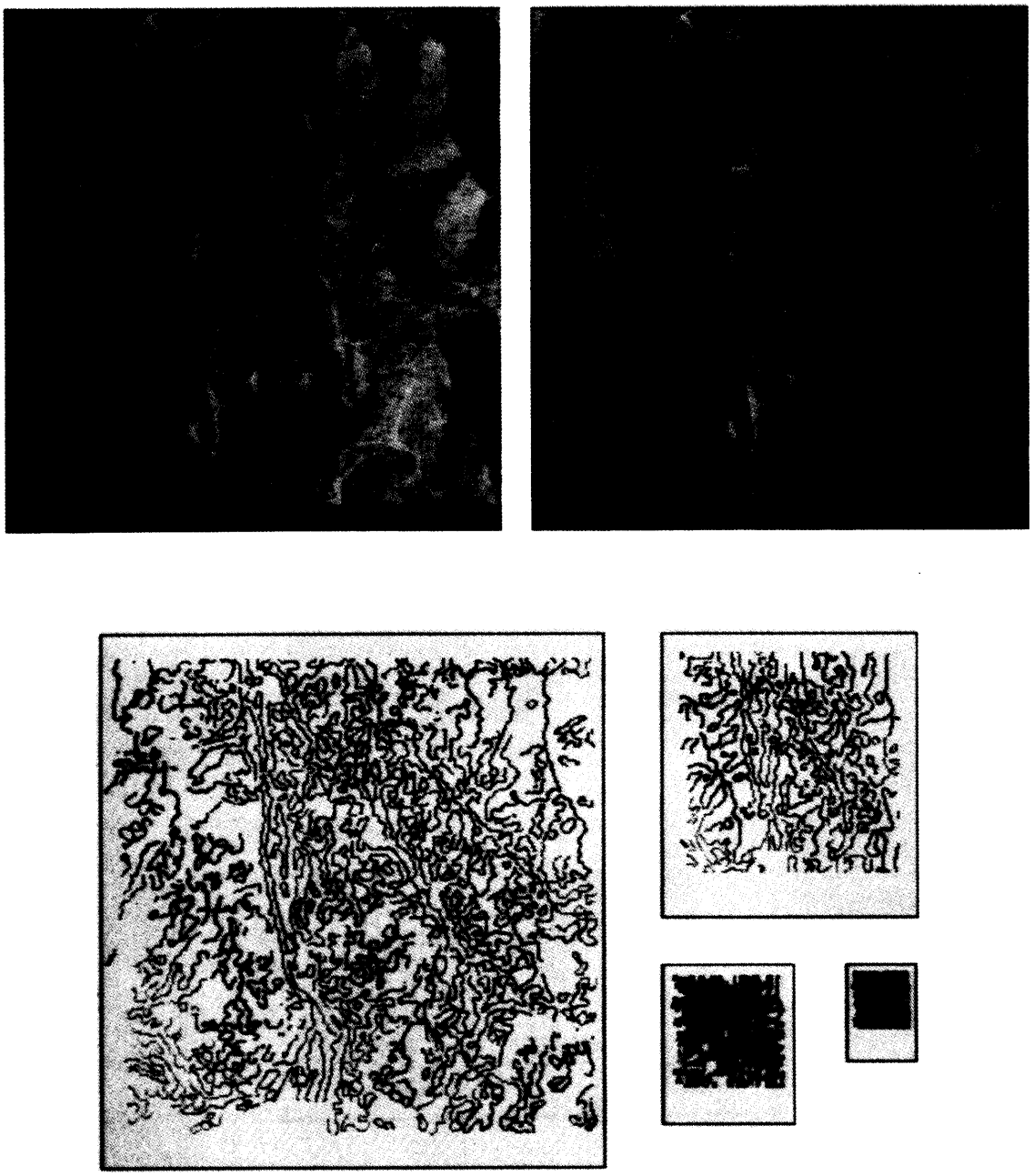

Figure 26. Natural terrain stereopair (top) and output of the MPG stereo algorithm (bottom). The images were 512 × 512 pixels, quantized to 256 levels (provided by the US Army Engineer Topographic Labs).

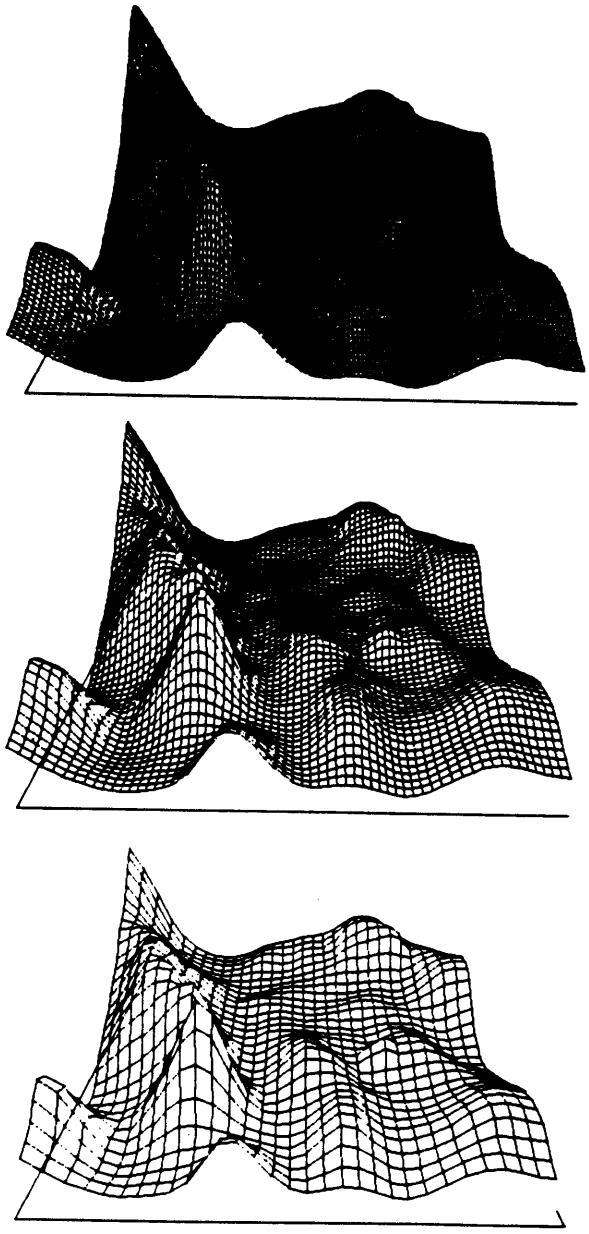

Figure 27. Reconstruction of data in Fig. 26 on the three coarsest levels. (Grid dimensions: $N_x^{h_1} = N_y^{h_1} = 33$, $N_x^{h_2} = N_y^{h_2} = 65$, $N_x^{h_3} = N_y^{h_3} = 129$, $N_x^{h_4} = N_y^{h_4} = 257$. Grid spacings: $h_1 = 0.8$, $h_2 = 0.4$, $h_3 = 0.2$, $h_4 = 0.1$. Constraint parameters: $\alpha_d{}^{h_j} = 0.01/h_j^2$. Computation: 31.0 work units.)

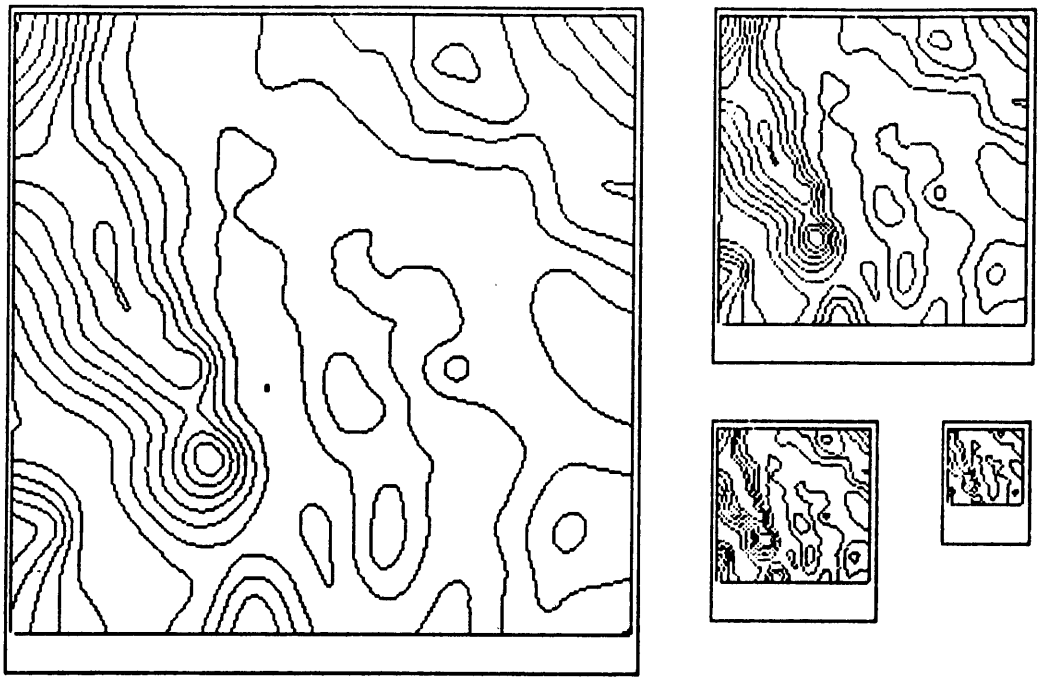

Figure 28. Isoelevation contour maps of the reconstructed terrain in Fig. 27.

6.5. Discontinuity Detection Experiments

The foregoing examples have shown that the surface reconstruction algorithm can handle discontinuities that are prespecified. In the next two sections, we present examples involving the automatic detection of discontinuities.

6.5.1. The Regularization Approach

The aerial view stereopair of Fig. 32 was input to the MPG stereo algorithm which generated the sparse disparity map shown on the top of Fig. 33. The finest level dense disparity map generated by a four-level surface reconstruction algorithm is shown at the lower left of the figure. Darkness is proportional to disparity. The discontinuities found from this disparity map, using a disparity limit $G_{ij} \geq t_d = 1$ are shown at the lower right as white contours After the detected points are added to the discontinuity map, the surface reconstruction algorithm continues iterating from the tentative approximation on the left. The amount of additional computation required is relatively small, since the tentative surface is a fairly good approximation in most places. At convergence, the reconstructed surface has fractured along the contours to give the solution on the right. Portions of the main discontinuities around the buildings have been found, but contours are broken and shifted.

The next example involves the synthesized random dot stereogram in Fig. 7. The depth constraints generated by a three-channel version of the MPG stereo algorithm are shown in Fig. 34

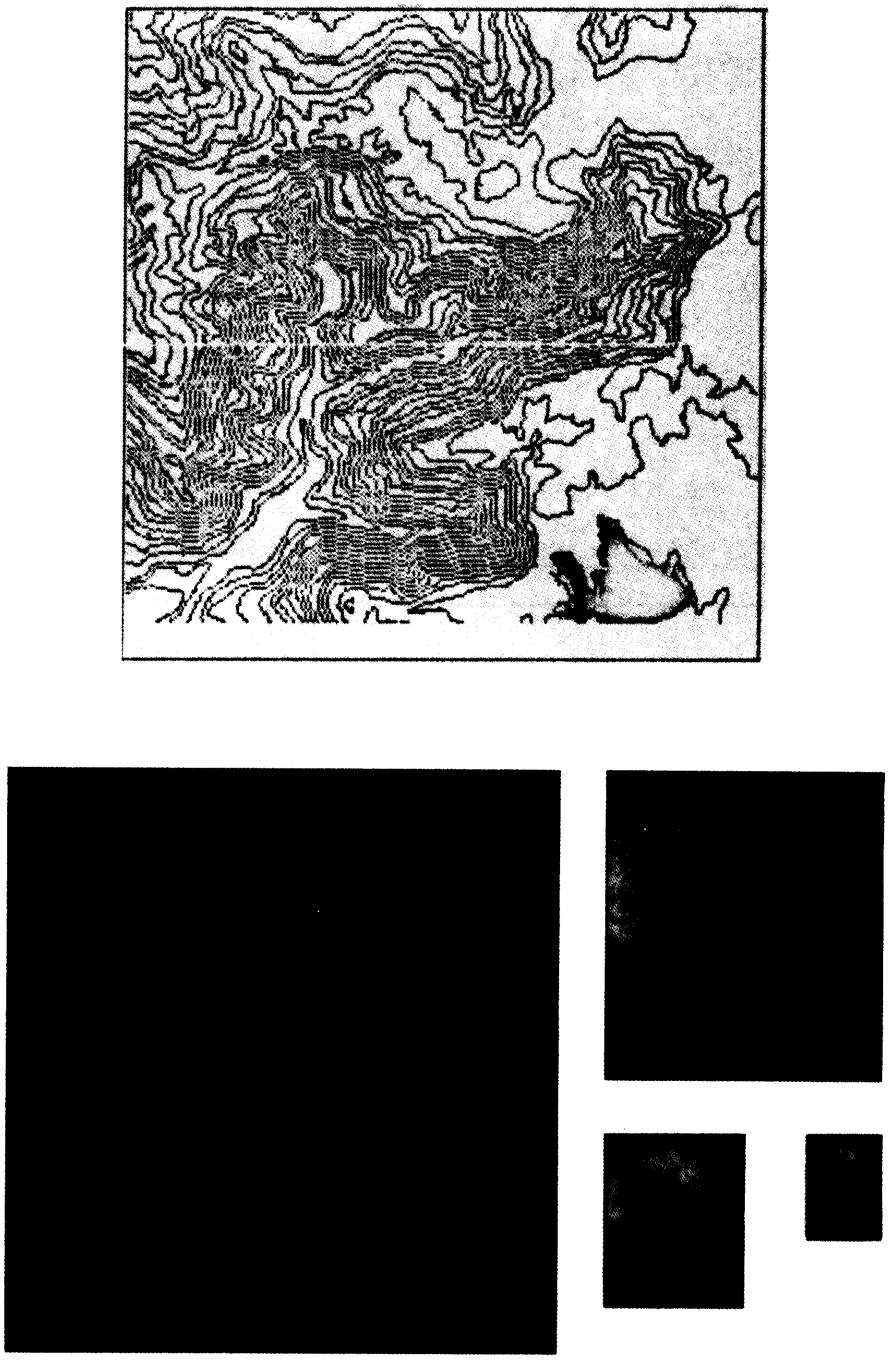

Figure 29. Digitized contour data (top) and constraints (bottom). The patch to the lower right represents a lake.

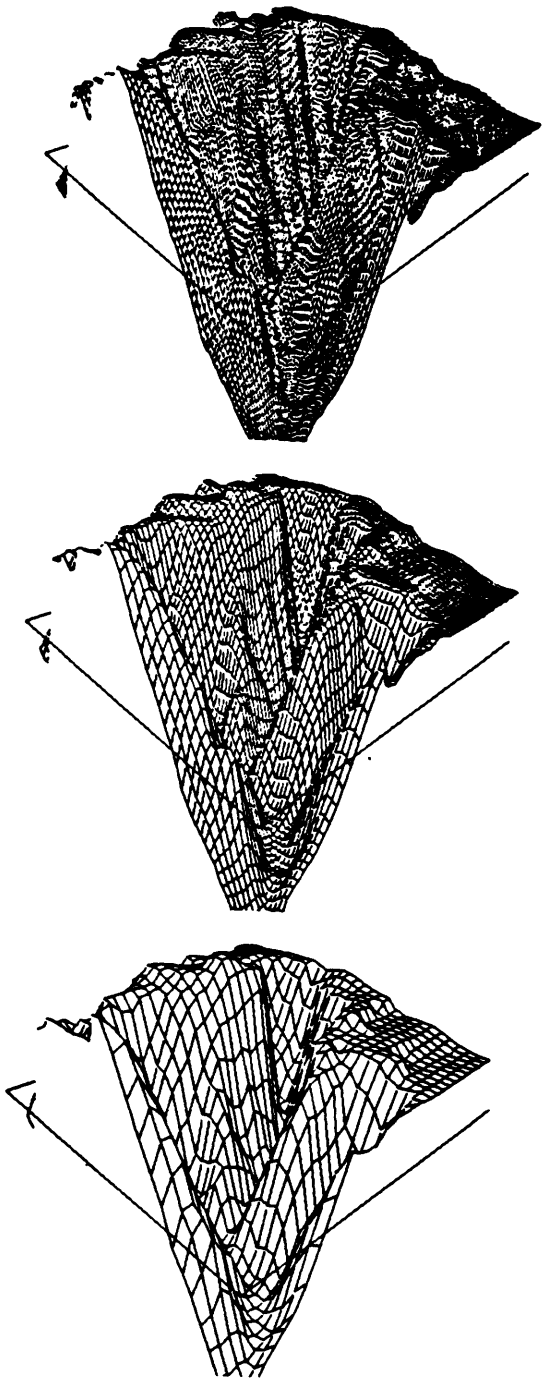

Figure 30. Reconstruction of data in Fig. 29. The terrain reconstruction on the three coarsest levels is represented as 3D surface plots. (Grid Dimensions: $N_x^{h_1} = N_y^{h_1} = 33$, $N_x^{h_2} = N_y^{h_2} = 65$, $N_x^{h_3} = N_y^{h_3} = 129$, $N_x^{h_4} = N_y^{h_4} = 257$. Grid spacings: $h_1 = 0.8$, $h_2 = 0.4$, $h_3 = 0.2$, $h_4 = 0.1$. Constraint parameters: ${\alpha_d}^{h_j} = 0.5/h_j^2$.)

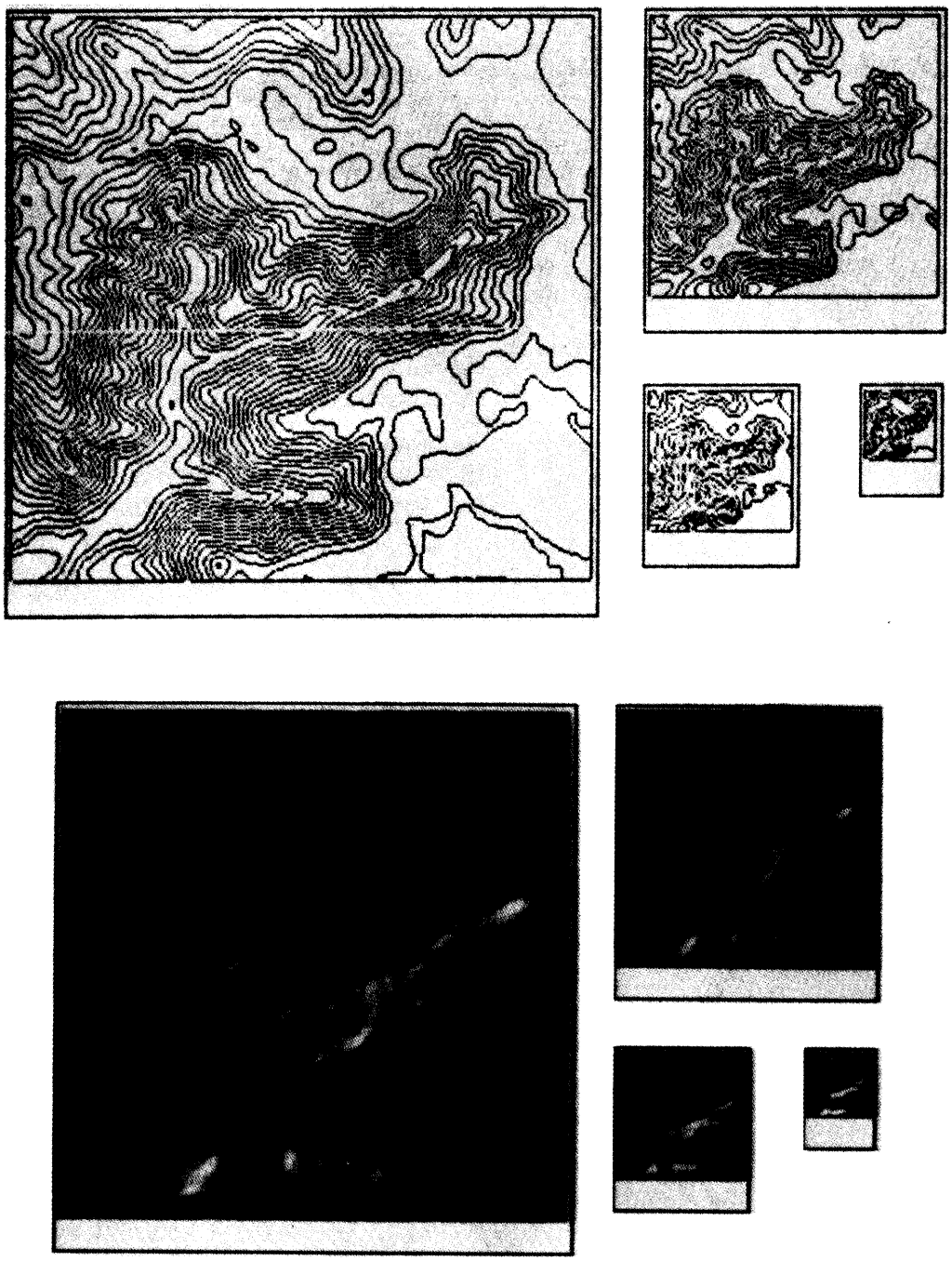

Figure 31. Isoelevation contour map (top) of the data in Fig. 30 and shaded representations of the reconstructed terrain (bottom).

Figure 32. Aerial view of a hospital complex. The stereopair was provided by the UBC Faculty of Forestry. Images are 320 × 320 pixels.

(the finest level dimensions are 320 × 320). The constraints on the coarsest level were obtained by averaging those on the next finer level. Fig. 35 shows the smooth disparity maps initially computed by a four-level surface reconstruction algorithm. Fig. 36 shows the discontinuities detected from these maps with $t_d = 1$. The discontinuities have been superimposed onto the final disparity maps in Fig. 37. Better performance is observed in this case due to the simpler surface structure, but the contours, while mostly intact, are quite ragged.

In general, not detecting true discontinuities affects surface shape more adversely over larger regions than introducing some spurious ones within a continuous surface. Discontinuity points are missed by the thresholding operation, and no adjustment of the global limit can be expected to produce perfect results. Note, however, that the surface reconstruction algorithm does not break down. Rather, the reconstructed surface degrades as it "leaks" through the gaps. The discontinuity detection procedure may be improved by allowing the disparity limit to vary spatially, or by modifying it during multiple passes. On the first pass, surface shape is poorest, so a fairly conservative limit should be set to reduce the number of false detections. Conservative limits fail to detect many discontinuities, but as more discontinuities are identified, surface shape improves and limits can be lowered in subsequent passes to find the less prominent discontinuities.

6.5.2. The Variational Continuity Control Approach

A multipass scheme is also employed in the variational continuity control approach to efficiently obtain good solutions. An example will be used to explain the strategy. Fig. 38 shows depth constraints randomly sampled from a set of sloping planes that form discontinuities along their

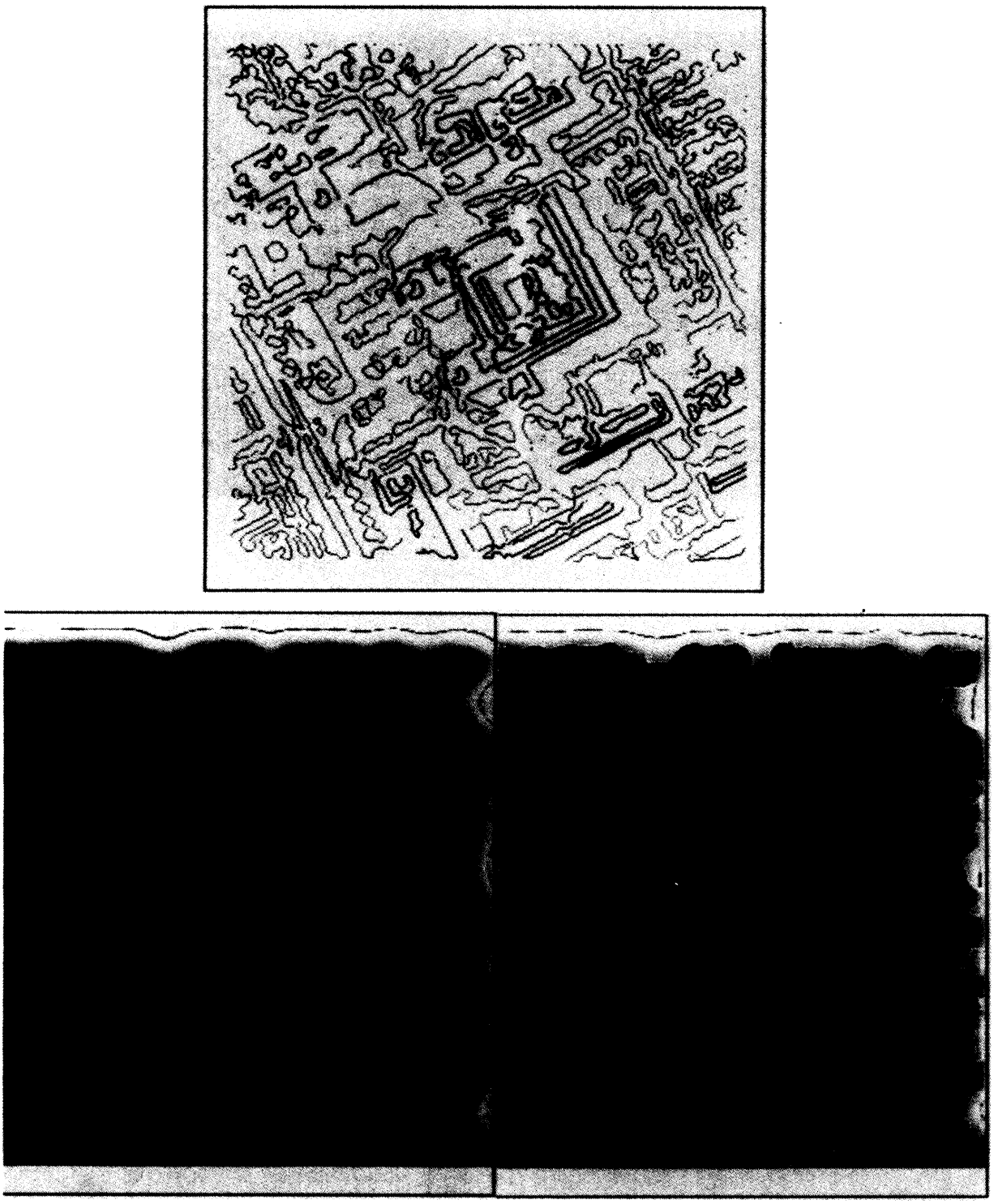

Figure 33. Discontinuities in the aerial stereogram. Disparity contours generated by stereo algorithm (top), full disparity map generated by the surface reconstruction algorithm at the finest level (lower left), and detected discontinuities superimposed on the disparity map (lower right).

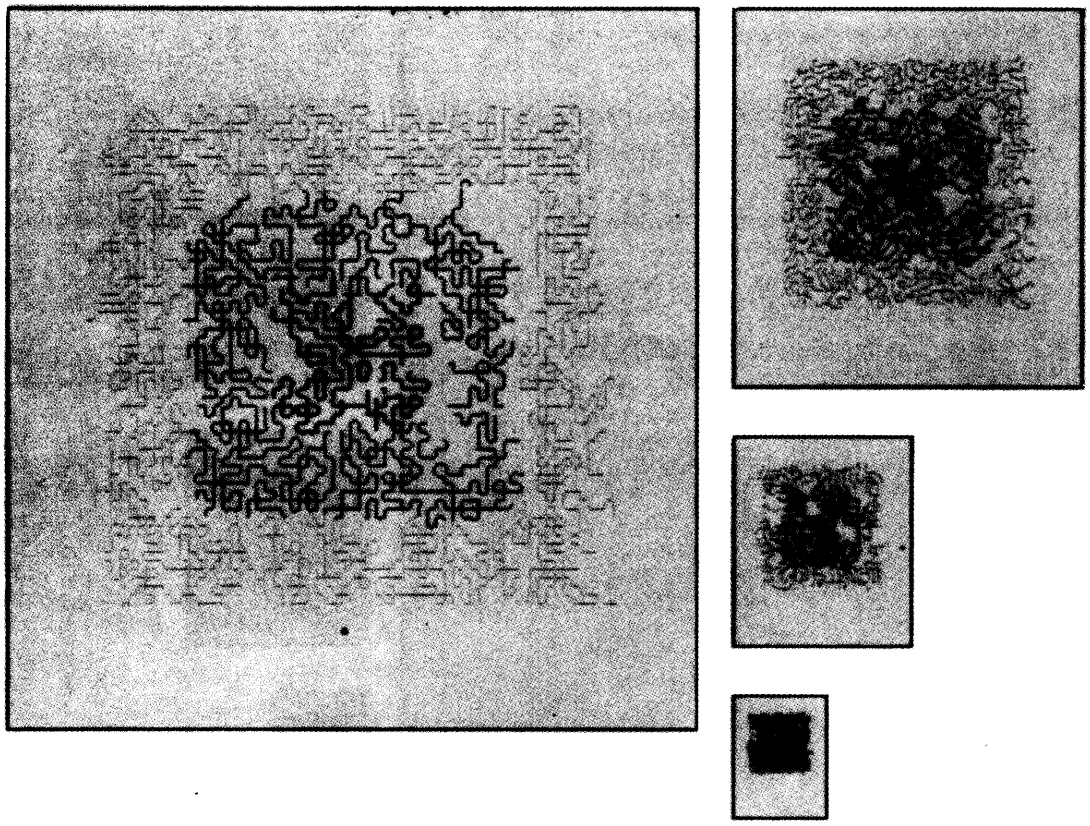

Figure 34. Depth constraints for the random dot stereogram.

extremities. A single continuous surface can be reconstructed from these constraints as shown, but it smooths over the depth discontinuities and rounds out the orientation discontinuities.

The algorithm finds both kinds of discontinuities and reconstructs a surface which preserves them. When the surface smooths through a depth discontinuity, two spurious regions of high curvature border the discontinuity. These spurious regions can easily be mistaken for orientation discontinuities. To avoid this unwanted interaction which can substantially slow down the optimization process, the algorithm postpones the orientation discontinuity detection phase until all depth discontinuities have been found. The surface evolves in several steps over which the parameters β_d^h and β_o^h in (29) are modified. Each step consists of first flipping the value of the continuity control parameter ($\rho_{i,j}^h$ or $\tau_{i,j}^h$) from 0 to 1 or conversely, if this lowers the energy (27), and then running the reconstruction algorithm to convergence (which always results in equilibrium, since the variational principle is convex for fixed $\rho_{i,j}^h$ and $\tau_{i,j}^h$).

For depth discontinuities, β_d^h is initially set to a high value that heavily penalizes their insertion, then lowered in steps. This strategy of least commitment finds the prominent discontinuities earliest, improving the surface as it does so, and leaves the more subtle ones for later. It results in the flipping of relatively few variables in each stage, hence the solution is obtained efficiently. Beginning with the continuous surface Fig. 39(a), Fig. 39(b–d) illustrates the steps of the evolving discontinuity detection process, during which discontinuities are determined with increasing accuracy as β_d^h is

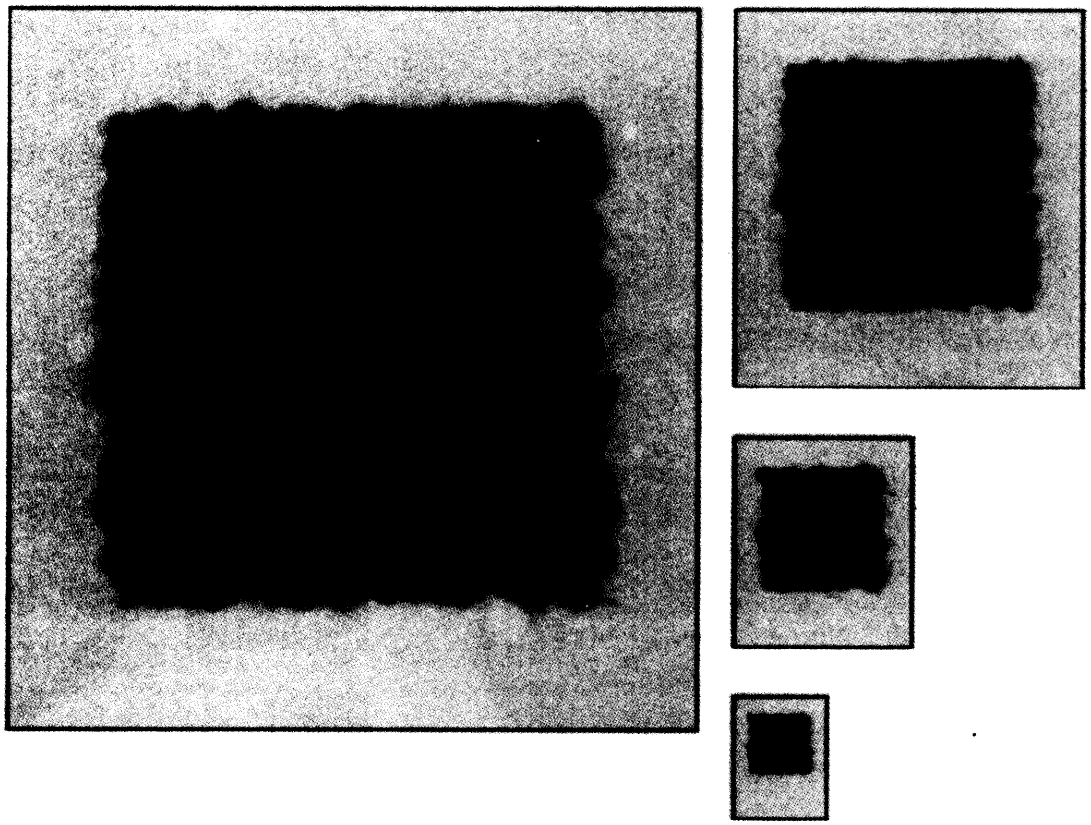

Figure 35. Full disparity maps without discontinuities.

lowered. The energy can be lowered further still if β_d^h is then increased slightly to eliminate spurious discontinuities in Fig. 39(d). Note that since the surfaces have now separated, a very large increase would be needed to flip a true discontinuity point (a hysteresis effect). The improved surface in Fig. 39(e) results. Next, the orientation discontinuity detection phase is activated and it runs in the same way, but modifies β_o^h. In this example, the orientation discontinuities are found in only one step.

Fig. 39(f) shows the final solution. The depth and orientation discontinuities have been made explicit and are preserved by the reconstructed surface. Incidentally, the global optimum of the variational principle has been found in this example; however, this procedure can generally be expected to yield good, though not necessarily optimal approximations. Its main attractions are that it is deterministic and efficient.

7. Discussion and Research Directions

Several issues concerning the framework for computing visible-surface representations are discussed in this section, and directions for future research are suggested. The discussion focuses on

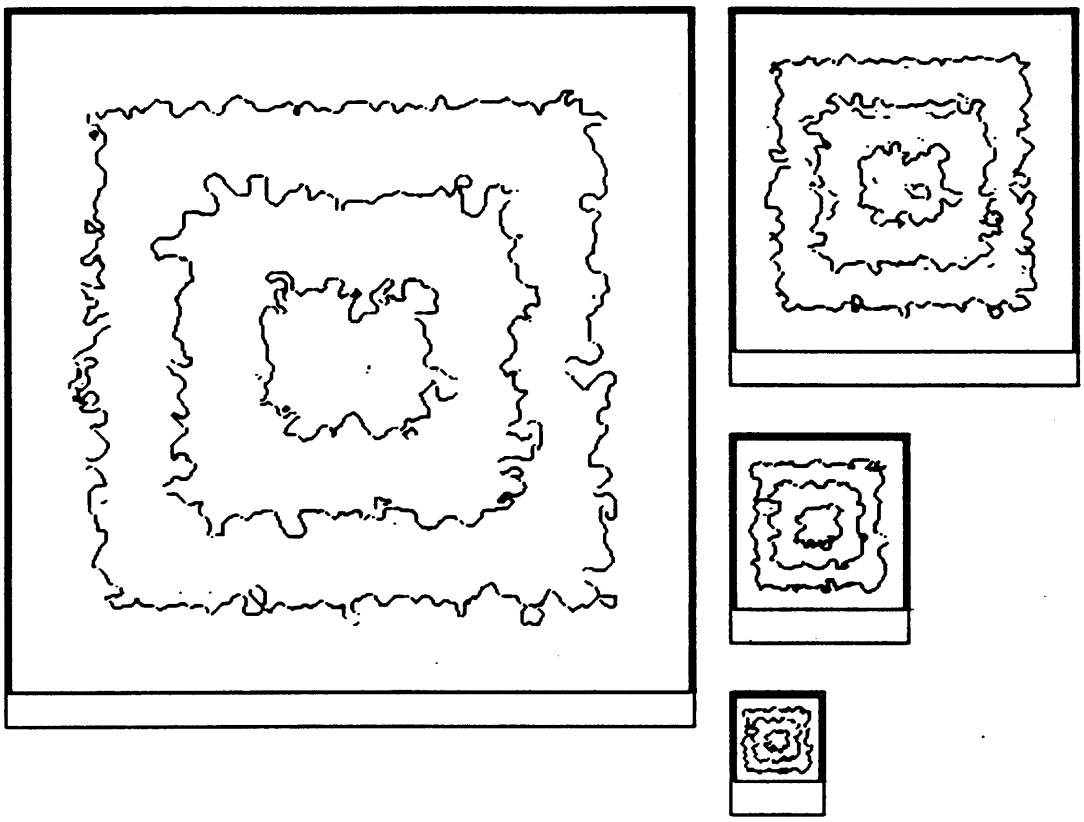

Figure 36. Detected discontinuities.

discontinuity detection, choosing constraint parameters, handling rivalries in constraints, grouping constraints, invariance properties of the surface reconstruction model, and visible-surface analysis. [Terzopoulos, 1984, Ch. 11] contains a more extensive treatment of these and other issues, including multiresolution relative depth representations of surfaces, and the possibility of computing visible-surface representations "instantaneously" by analog networks.

7.1. On Discontinuity Detection

Some recent work in image restoration is of relevance to the problem of piecewise continuous surface reconstruction. A piecewise constant image model employed by Blake [1983] for image reconstruction is interesting in that it incorporates "weak constraints" which can be broken at a cost. The resulting optimization problem is related to our variational continuity control approach, but more restricted. Blake used an adaptive method, which he referred to as "graduated nonconvexity," to obtain good solutions to the nonconvex problem. It has not been established however whether this interesting method applies to the sparse data case as well.

Geman and Geman [1985] used Markov random field models with associated Gibbs distributions to restore piecewise constant images corrupted by additive Gaussian noise. The restoration seeks a maximum *a posteriori* estimate of the original image, given the degraded image, and includes

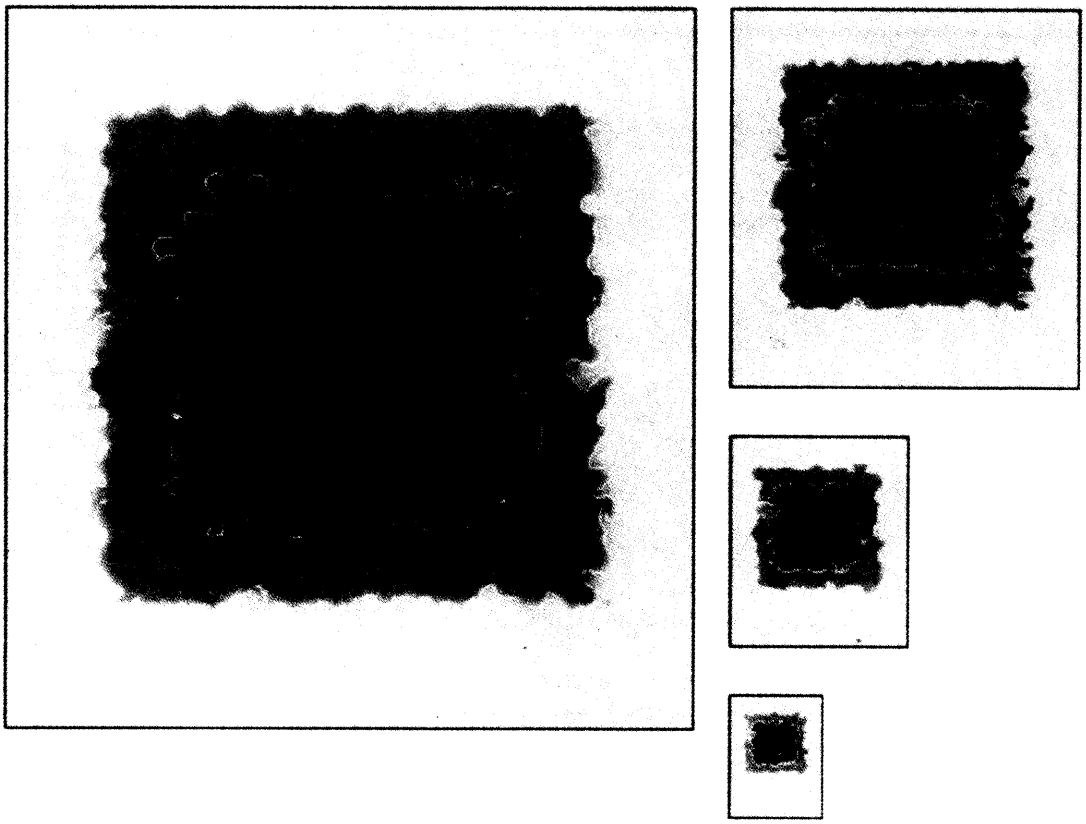

Figure 37. Discontinuities and final disparity maps.

an explicit "line process" that estimates the locations of step edges in intensity. This work was restricted to dense image data. The Gemans' approach was adopted with encouraging results to surface reconstruction from sparse depth constraints by Marroquin [1984]. His markov random field model, while less restrictive than the Gemans' piecewise constant one, in fact models a membrane spline whose smoothness is insufficient for computing visible-surface representations. A line process essentially equivalent to the Gemans' was incorporated to estimate depth discontinuities. The numerical solution strategy in both of the above studies was stochastic optimization using the Metropolis algorithm and simulated annealing to optimize the nonconvex functional [Kirkpatrick *et al.*, 1983]. This strategy can find optimal solutions, but for such large reconstruction problems it has been observed to converge notoriously slowly. Based on our experience, we believe that it can be accelerated, perhaps enough to make it practical, through the use of multiresolution processing.

Obviously, the line processes used in the above work as well as our own encoding of discontinuity contour configurations is unpleasingly heuristic and in need of refinement. The discontinuity map can be augmented by nodal variables to encode the local orientations of the curvilinear elements to a higher degree of accuracy. Such an encoding is employed by Zucker and Parent [1984] in an optimization (relaxation labeling) approach to finding contours in images. It appears that ideas from their work can also be applied to finding surface discontinuities within our framework.

A promising possibility is to employ 1D controlled-continuity stabilizers as formal models of smoothness constraints along surface discontinuity contours in the x–y plane. A functional

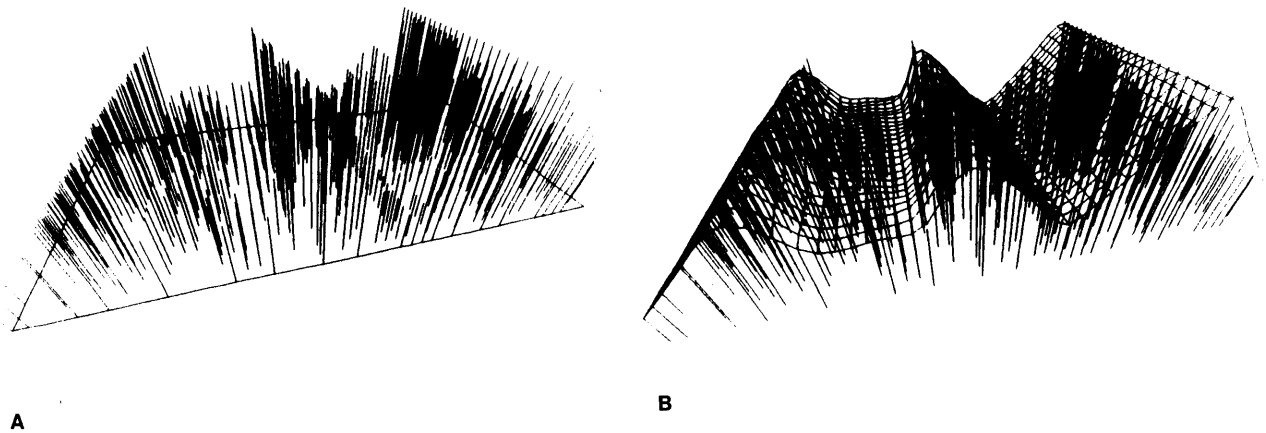

Figure 38. Scattered depth constraints consistent with sloping planes meeting discontinuously (top) and the smooth reconstructed surface (bottom).

that naturally comes to mind is the curvilinear analog of the thin plate surface under tension: $\oint_C \beta_b(s)\{\beta_a(s)(\partial^2 c/\partial s^2)+[1-\beta_a(s)](\partial c/\partial s)\}\,ds$, where s denotes arc length along discontinuity contours $c \in C$. Here, β_b allows breaks, while β_a allows angles (tanjent discontinuities) to form in the discontinuity contours. Again, additional energy penalties must be associated with these occurrences. Given our finite element representation of surfaces, curvilinear finite elements are the natural local representation for discontinuity contours. The combined variational principle has both a surface component and an analogous contour component. Although technically nontrivial, a formulation of surface reconstruction generalized along these lines has very strong appeal.

7.2. On Constraints — Parameters, Rivalries, and Grouping

The constraint (spring) parameters offer the flexibility to individually tune the coerciveness of each constraint on the reconstructed surface. In the special case of Gaussian error distributions, the parameters should be inversely proportional to the expected variances ($\alpha_i = 1/\lambda\sigma_i^2$). It ought to be possible for the low-level visual processes to associate a variance estimate or confidence with each constraint that they provide. In general, however, it's not obvious how to choose the constraint parameters optimally.

The constant of proportionality λ^{-1} can also be used to tune the overall smoothness of the reconstructed surface. Cross validation techniques may be used to set λ optimally (e.g., [Wahba and Wendelberger, 1980]). The basic criterion is to choose λ so as to minimize over all constraints the (weighted) discrepancy between each constraint and its value as estimated from the surface reconstructed using the remaining constraints. Unfortunately, this involves computationally expensive sequential algorithms. Interestingly, the continuous tuning of surface smoothness is analogous to the scale space filtering technique proposed by Witkin [1983] with the added attraction that it can be applied to scattered data.

Although the variational principle was designed to account for measurement errors in the constraints, the possibility of massive rivalries between constraints from different sources, such as stereopsis and analysis of motion processes, was disregarded. Massive rivalries are unnatural visual

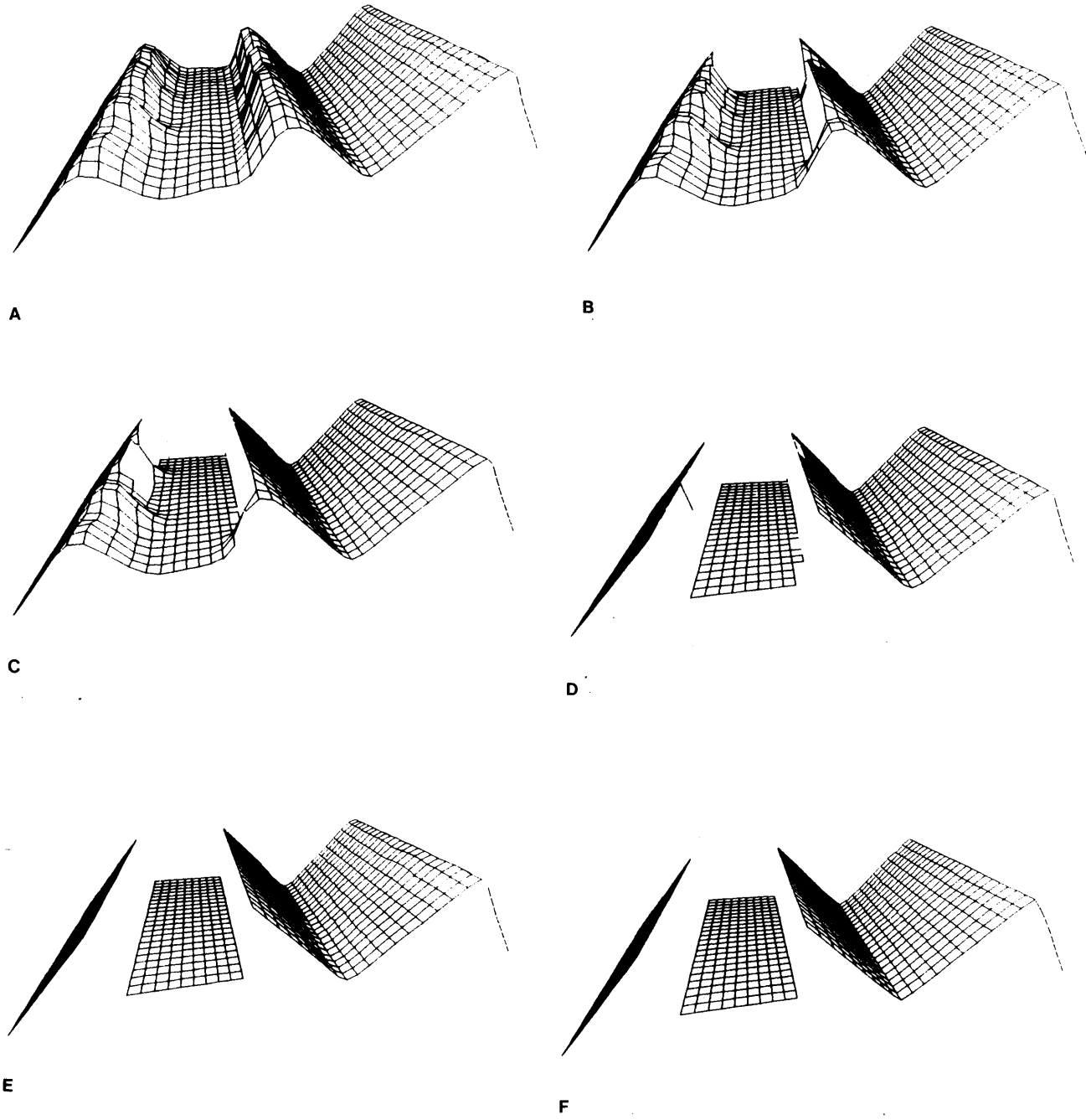

Figure 39. Evolution of the discontinuity detection process.

phenomena that can nevertheless occur, especially under contrived conditions, and they often lead to multistable percepts [Attneave, 1971]. The framework can potentially accommodate rivalries with a mechanism that inhibits individual or entire sets of constraints by nullifying selected constraint parameters. This mechanism can be activated by a global arbitrator which monitors the contents of visible-surface representations to detect rivalries may also have access to higher level knowledge

about the scene. The arbitrator's influence can account for multistability.

A particular type of rivalry arises from transparent surfaces. For instance, a surface such as a dirty window in front of a background scene would lead to two well defined populations of depth (and orientation) constraints over the same visual angle, one from the window, the other from the background. A transparency interpretation can be arrived at by an arbitrator which monitors the surface reconstruction process looking for high approximation error between surface and constraints over a significant area. Under these conditions the arbitrator can trigger a constraint grouping process which clusters the constraints into two populations, based on depth values, say. Multiple surfaces can then be reconstructed over the same visual area for each resulting constraint population. This scheme has been applied on a transparent surface random dot stereogram [Terzopoulos, 1984, Ch. 11].

7.3. On Invariance Properties of the Surface Model

As a transformation from sparse constraints to dense surfaces, the thin plate under tension model can be shown to be invariant under (i.e. commutes with) certain image plane transformations applied to the constraints; namely, translations, rotations, and similarity transformations. This implies that surface shapes will be preserved through rigid motions of the scene or viewpoint parallel to the image plane or along the view direction. These are essential invariance properties for visible-surface reconstruction [Terzopoulos, 1982].

Note, however, that the thin plate spline, characterized by the small deflection approximation $\int\int v_{xx}^2 + 2v_{xy}^2 + v_{yy}^2 \, dx\, dy$ to the bending energy density of a thin plate, is not invariant under arbitrary 3D transformations of the constraints. Thus, surface interpolation using this expression is not invariant under changes in the view direction, as Blake [1984] points out. He shows that rotating the view direction induces the 1D analog of the thin plate spline to "wobble," and he demonstrates that this effect is most pronounced as the (continuous) spline is inclined sharply with respect to the viewer or is forced to bend sharply. Blake views this as a problem that should be eliminated by employing the large deflection bending energy of the thin plate, a convex combination of the mean and Gaussian curvatures of the surface $v(x, y)$, which is view direction invariant.

Although $\mathcal{E}_{pr}(v)$ can also be made view direction invariant by employing the large deflection counterparts for the thin plate and membrane bending energies, this approach has a serious technical drawback [Terzopoulos, 1984]: The large deflection formulas lead to an extremely difficult nonlinear problem (e.g., the large deflection equations for the thin plate are two coupled nonlinear fourth-order partial differential equations known as Von Karmann's equations [Szilard, 1974]).

Fortunately, the surface reconstruction model, as it stands, is not hampered by the lack of view direction invariance because the available constraints are usually sufficiently dense in practice to tightly determine surface shape; as the view direction is varied, the reconstructed surface would vary negligibly (note that Blake's experiments reveal a significant wobble effect just in the case of extremely sparse constraints). Furthermore, the explicit introduction of depth and orientation discontinuities alleviates much of the wobble precisely at those places where Blake's experiments show it to be most pronounced on a globally continuous surface. An interesting psychophysical experiment would be to determine whether there might be some slight variance in the surfaces perceived by humans viewing sparse random dot stereograms while the dots undergo simulated rigid 3D transformations and, if so, whether the variations are consistent with the reconstruction model (J. Mayhew, personal communication).

7.4. On Visible-Surface Analysis

The visible-surface representation is an intermediate and volatile description of the 3D surfaces in scenes. It drives ensuing processes which generate stable higher-level representations of shape that are better tuned to object recognition. The processing begins with *visible-surface analysis* whose goal is to abstract from the numeric, viewer-centered representation a rich set of more symbolic, object-centered features that are stable through viewpoint changes. The extraction of geometric surface features is facilitated by the dense shape information provided by visible-surface representations.

A promising approach to visible-surface analysis is to apply concepts from differential geometry [do Carmo, 1976]. For instance, a surface's intrinsic geometry (including Gaussian curvature, geodesics, etc.) is determined completely by the first fundamental form, which defines arc length over the surface. It's extrinsic geometry (including normal curvature, principal curvatures, etc.) are determined by the second fundamental form, which describes the deviation of the surface from the local tangent plane. The fundamental theorem of the local theory of surfaces (usually attributed to Bonnet) states that the analytic study of surface properties consists of the study of the two fundamental forms; i.e., the six fundamental tensor coefficients (which are not all independent) as functions of the two independent parameters of the surface. The fundamental forms are invariant under changes in the parameterization, and together they determine surface shape up to rigid body transformations. These properties make them ideal foundations for object centered symbolic surface representations.

The visible-surface representation makes it possible to estimate the first and second fundamental forms on a point-by-point basis over the entire visible surface. The finite element shape representation reduces the computation of crucial local surface features such as the Gaussian curvature, principal curvatures, and principal directions to the evaluation of simple algebraic expressions of neighboring nodal variables (see [Terzopoulos, 1984, Ch. 11] for derivations). It is then a simple step to determine the elliptic, hyperbolic, parabolic, umbilic, and planar points, as well as geodesics, asymptotes, and lines of curvature.

For example, Fig. 39 shows the reconstructed surface of a lightbulb. Fig. 40 shows the Gaussian curvature $K(x, y)$ computed for the reconstructed lightbulb surface of Fig. 20. The elliptic points ($K > 0$) are shown in white, the hyperbolic points ($K < 0$) are shown in black, and the parabolic ($K = 0$) points separate the two regions. Note the alternation in the sign of curvature at the screw mount. Fig. 41 plots the computed field of principal directions for the lightbulb at the two coarsest scales. These demonstrations illustrate the feasibility of reliably computing from these representations higher-order intrinsic and extrinsic properties of surface shape. The reliability can be attributed to the regularizing properties of the thin plate surface under tension which overcomes the potentially detrimental effects of noise in the data, while preserving discontinuities. For further analysis of the kinds of features that can be computed from dense, numeric, representations of surfaces see, e.g., [Brady *et al.*, 1985] or [Medioni and Nevatia, 1984].

8. Conclusion

Constraints on surface shape, contributed by multiple low-level visual processes, can be computed reliably at multiple resolutions, but only at scattered locations in the field of view. Subsequent visual processing can be facilitated substantially if the scattered constraints are transformed into visible-surface representations that make surface shape explicit everywhere. To accomplish this

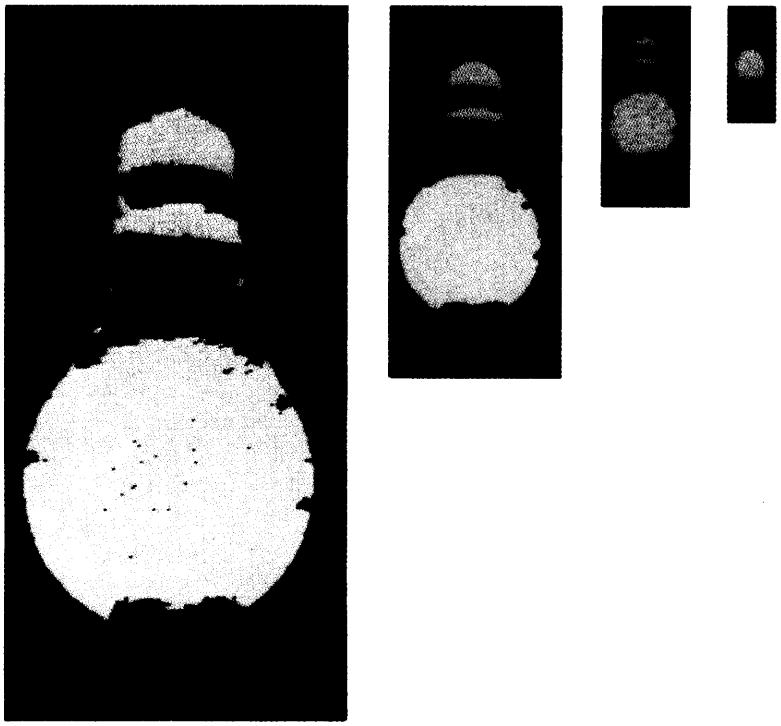

Figure 40. Elliptic (white) and hyperbolic (black) points of the reconstructed lightbulb at four scales.

effectively, information must be integrated over multiple visual modalities and fused across multiple scales of resolution.

In this paper, we have developed a computational theory of visible-surface representations. Within a unified computational framework, formal solutions were offered to fundamental problems of reconstructing visible surfaces: (i) integrating constraints on the depth and orientation of surfaces across various modalities and scales, (ii) interpolating surface shape information into (piecewise) smooth surfaces, (iii) discovering discontinuities in surface depth and orientation and enabling them to restrict interpolation, and (iv) efficiently maintaining consistency in distributed, multiresolution visible-surface representations.

A visible-surface reconstruction algorithm implements the framework. Extensive testing has shown it to be viable. The algorithm coordinates cooperative processes within a multiresolution hierarchy of surface representations to dramatically increase computational efficiency. It is well suited to implementation on massively parallel networks of simple, locally interconnected processors. Such computational networks are suggestive of biological mechanisms and are also well suited to VLSI technology.

ACKNOWLEDGEMENTS

This paper is based on a thesis supervised by Michael Brady and Shimon Ullman. Tomaso Poggio and Michael Brady provided helpful comments on a draft. Several people kindly supplied some of the test data used in the experiments: Philippe Brou provided laser rangefinder data, Eric Grimson,

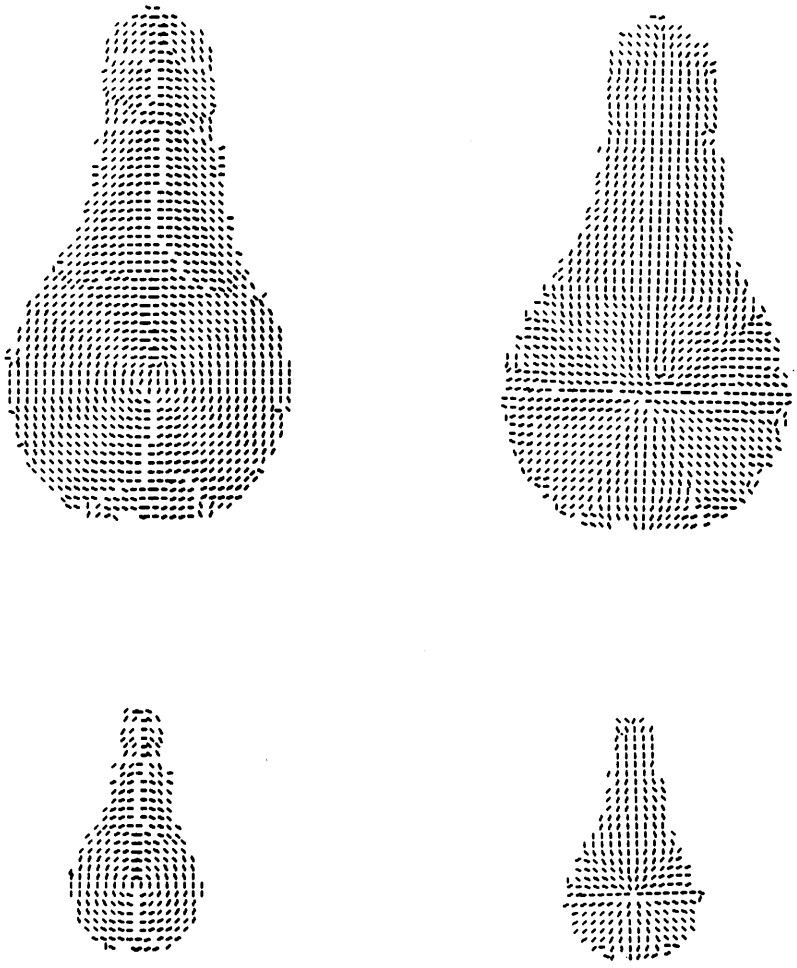

Figure 41. Principal directions for the reconstructed lightbulb at the two coarsest scales. The directions of greatest curvature are shown on the left. Those of least curvature are shown on the right.

Michael Kass, and Kieth Nishihara provided stereo data, Katsushi Ikeuchi provided photometric stereo data, and James Mahoney provided digital terrain maps.

References

Abramowitz, M., and Stegun, I.A. (ed.), [1965]. *Handbook of Mathematical Functions*, Dover, New York.

Ahlberg, J.H., Nilson, E.N., and Walsh, J.L., [1967]. *The Theory of Splines and their Applications*, Academic Press, New York.

Attneave, F., [1971]. "Multistability in perception," *Scientific American*, **225**, 6, 63–71.

Bakhvalov, N.S., [1977]. *Numerical Methods*, Mir Publishers, Moscow.

Ballard, D.H, Hinton, G.E., and Sejnowski, T.J., [1983], "Parallel visual computation," *Nature*, **306**, 5938, 21–26.

Barrow, H.G., and Tenenbaum, J.M., [1978], "Recovering intrinsic scene characteristics from images," *Computer Vision Systems* (A. Hanson and E. Riseman, Eds.), Academic Press, New York, 3–26.

Barrow, H.G., and Tenenbaum, J.M., [1979], "Reconstructing smooth surfaces from partial, noisy information," *Proc. DARPA Image Understanding Workshop*, U. Southern California, L.S. Baumann (ed.), 76–86.

Blake, A., [1983], "The least-disturbance principle and weak constraints," *Pattern Recognition Letters*, **1**, 393–399.

Blake, A., [1984], "Reconstructing a visible surface," *Proc. National Conf. on AI (AAAI-84)*, Austin, TX, 23–26.

Bolondi, G., Rocca, F., and Zanoletti, S., [1976], "Automatic contouring of faulted subsurfaces," *Geophysics*, **41**, 1377–1393.

Braddick, O.J., Campbell, F.W., and Atkinson, J., [1978], "Channels in vision: Basic aspects," *Handbook of Sensory Physiology: Perception*, Vol. 8, R. Held, H.W. Leibowitz, and H.L. Teuber (ed.), Springer, Berlin, 3–38.

Brady, J.M., and Horn, B.K.P., [1983], "Rotationally symmetric operators for surface interpolation," *Computer Vision, Graphics, and Image Processing*, **22**, 70–94.

Brady, J.M., Ponce, J., Yuille, A., and Asada, H., [1985], Describing surfaces, MIT AI Lab., Cambridge, MA, AI Memo 822.

Brady, J.M., and Yuille, A., [1984], "An extremum principle for shape from contour," *IEEE Trans. Pattern Analysis and Machine Intelligence*, PAMI-6, 288–301.

Brandt, A., [1977], "Multi-level adaptive solutions to boundary-value problems," *Math. Comp.*, **31**, 333–390.

Briggs, I.C., [1974], "Machine contouring using minimum curvature," *Geophysics*, **39**, 39–48.

Burt, P., and Julesz, B., [1980], "A disparity gradient limit for binocular fusion," *Science*, **208**, 615–617.

Collett, T.S., [1984], Extrapolating and interpolating surfaces in depth, School of Biological Sciences, University of Sussex, Brighton, UK.

Courant, R., and Hilbert, D., [1953], *Methods of Mathematical Physics*, Vol. I, Interscience, London.

do Carmo, M.P., [1976], *Differential Geometry of Curves and Surfaces*, Prentice-Hall, Englewood Cliffs, NJ.

Duchon, J., [1977], "Splines minimizing rotation-invariant semi-norms in Sobolev spaces," *Constructive Theory of Functions of Several Variables*, A. Dodd and B. Eckmann (ed.), Springer-Verlag, Berlin, 85–100.

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

Fedorenko, R.P., [1961], "A relaxation method for solving elliptic difference equations," *Zh. vych. Mat. mat. Fiz. (USSR Comp. Math. and Math. Phys.)*, **1**, 922–927.

Geman, S., and Geman, D., [1985], "Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images," *IEEE Trans. Pattern Analysis and Machine Intelligence*, PAMI-6, 721–741.

Gibson, J.J., [1950]. *The Perception of the Visual World*, Houghton Mifflin, Boston, MA.

Grimson, W.E.L., [1983], "An implementation of a computational theory of visual surface interpolation," *Computer Vision, Graphics, and Image Processing*, **22**, 39–69.

Grimson, W.E.L., [1985], "Computational experiments with a feature based stereo algorithm," *IEEE Trans. Pattern Analysis and Machine Intelligence*, PAMI-7, 17–34.

Hackbusch, W., and Trottenberg, U., (ed.), [1982], *Multigrid Methods*, Lecture Notes in Mathematics, Vol. 960, Springer-Verlag, New York.

Hageman, L.A., and Young, D.M., [1981]. *Applied Iterative Methods*, Academic Press, New York.

Harder, R.L., and Desmarais, R.N., [1972], "Interpolation using surface splines," *Jour. Aircraft*, **9**, 189–191.

Horn, B.K.P., [1981], "Hill shading and the reflectance map," *Proc. IEEE*, **69**, 14–47.

Horn, B.K.P., [1982], "Sequins and quills — A representation for surface topography," *Representation of three-dimensional objects*, R. Bajcsy (ed.), Springer-Verlag, New York.

Horn, B.K.P., and Brooks, M.J., [1985], The variational approach to shape from shading, MIT AI Lab., Cambridge, MA, AI Memo 820.

Ikeuchi, K., and Horn, B.K.P., [1981], "Numerical shape from shading and occluding boundaries," *Artificial Intelligence*, **17**, 141–184.

Julesz, B., [1971], *Foundations of Cyclopean Perception*, University of Chicago Press, Chicago, IL.

Kass, M., [1983], "A computational framework for the visual correspondence problem," *Proc. 8th Int. J. Conf. AI*, Karlsruhe, W. Germany, 1043–1045.

Kender, J.R., [1980], Shape from texture, Computer Science Dept., Carnegie-Mellon University, Pittsburgh, PA, CMU-CS-81-102.

Kimeldorf, G., and Wahba, G., [1970], "A correspondence between Bayesian estimation on stochastic processes and smoothing by splines," *Ann. Math. Stat.*, **41**, 495–502.

Kirkpatrick, S., Gelatt, C.D., Jr., Vecchi, M.P., [1983], "Optimization by simulated annealing," *Science*, **220**, 671–680.

Langridge, D.J., [1984], "Detection of discontinuities in the first derivatives of surfaces," *Computer Vision, Graphics, and Image Processing*, **27**, 291–308.

Laurent, P.J., [1972], *Approximation et Optimisation*, Hermann, Paris.

Leclerc, Y., and Zucker, S.W., [1984], The local structure of image discontinuities in one dimension, Computer Vision and Robotics Laboratory, McGill University, Montreal, Que., Canada, TR-83-19R.

Marr, D., [1982], *Vision: A Computational Investigation into the Human Representation and Processing of Visual Information*, Freeman, San Fransisco, CA.

Marr, D., and Hildreth, E.C., [1980], "Theory of edge detection," *Proc. R. Soc. Lond. B*, **207**, 187–217.

Marr, D., and Nishihara, H.K., [1978], "Representation and recognition of the spatial organization of three-dimensional shapes," *Proc. R. Soc. Lond., B*, **200**, 269–294.

Marroquin, J.L., [1984], Surface reconstruction preserving discontinuities, MIT A.I. Lab., Cambridge, MA, AI Memo No. 792.

Medioni, G., and Nevatia, R., [1984], "Description of 3D surfaces using curvature properties," *Proc. DARPA Image Understanding Workshop*, New Orleans, LA, Baumann, L.S. (ed.), 291–229.

Meinguet, J., [1979], "Multivariate interpolation at arbitrary points made simple," *Jour. Applied Math. and Physics (ZAMP)*, **30**, 292–304.

Poggio, G.F., and Poggio, T., [1984], "The analysis of stereopsis," *Ann. Rev. Neurosci.*, **7**, 379–412.

Poggio, T., and Torre, V., [1984], Ill-posed problems and regularization analysis in early vision, MIT A.I. Lab., Cambridge, MA, AI Memo No. 773, reprinted in Proc. DARPA Image Understanding Workshop, New Orleans, LA, Baumann, L.S. (Ed.), 1984, 257–263.

Rosenfeld, A. (ed.), [1984], *Multiresolution Image Processing and Analysis*, Springer–Verlag, New York.

Rosenfeld, A., and Kak, A.C., [1982], *Digital Picture Processing (Second Edition)*, Vol. 2, Academic Press, New York.

Schumaker, L.L., [1976], "Fitting surfaces to scattered data," *Approximation II*, G.G. Lorentz, C.K. Chui, L.L. Schumaker (ed.), Academic Press, New York, 203–267.

Strang, G., and Fix, G.J., [1973], *An Analysis of the Finite Element Method*, Prentice-Hall, Englewoods Cliffs, NJ.

Szilard, R., [1974], *Theory and Analysis of Plates: Classical and Numerical Methods*, Prentice–Hall, Englewood Cliffs, NJ.

Terzopoulos, D., [1982], Multilevel reconstruction of visual surfaces: variational principles and finite element representations, MIT AI Lab., Cambridge, MA, AI Memo No. 671, reprinted in *Multiresolution Image Processing and Analysis*, A. Rosenfeld (Ed.), Springer-Verlag, New York, 1984, 237–310.

Terzopoulos, D., [1983a], "Multilevel computational processes for visible surface reconstruction," *Computer Vision, Graphics, and Image Processing*, **24**, 52–96.

Terzopoulos, D., [1983b], "The role of constraints and discontinuities in visible-surface reconstruction," *Proc. 8th Int. J. Conf. AI*, Karlsruhe, W. Germany, 1073–1077.

Terzopoulos, D., [Jan., 1984], Multiresolution Computation of Visible Surface Representations, Ph.D. thesis, Department of Electrical Engineering and Computer Science, MIT, Cambridge, MA.

Terzopoulos, D., [1985a], "Regularization of ill-posed visual reconstruction problems involving discontinuities," *IEEE Trans. Pattern Analysis and Machine Intelligence*, to appear.

Terzopoulos, D., [1985b], Concurrent multigrid coordination, MIT Artificial Intelligence Lab., Cambridge, MA, in preparation.

Tikhonov, A.N., and Arsenin, V.A., [1977], *Solutions of Ill-Posed Problems*, Winston and Sons, Washington, DC.

Ullman, S., [1979], "Relaxation and constrained optimization by local processes," *Computer Graphics and Image Processing*, **10**, 115–125.

Ullman, S., [1983], "Recent computational studies in the interpretation of structure from motion," *Human and Machine Vision*, J. Beck, B. Hope, and A. Rosenfeld (ed.), Academic Press, New York, 459–480.

Wahba, G., and Wendelberger, J., [1980], "Some new mathematical methods for variational objective analysis using splines and cross validation," *Monthly Weather Review*, **108**, 1122–1143.

Witkin, A.P., [1981], "Recovering surface shape and orientation from texture," *Artificial Intelligence*, **17**, 17–45.

Witkin, A.P., [1983], "Scale-space filtering," *Proc. 8th Int. J. Conf. AI*, Karlsruhe, W. Germany, 1019–1022.

Witkin, A.P., and Tenenbaum, J.M., [1983], "On the role of structure in vision," *Human and Machine Vision*, J. Beck, B. Hope, and A. Rosenfeld (ed.), Academic Press, New York, 481–543.

Woodham, R.J., [1981], "Analyzing images of curved surfaces," *Artificial Intelligence*, **17**, 117–140.

Zucker, S.W., Hummel, R. A., and Rosenfeld, [1977], "An application of relaxation labeling to line and curve enhancement," *IEEE Trans. Computers*, **C-26**, 394–403.

Zucker, S.W., and Parent, P., [1984], "Multiple size operators and optimal curve finding," *Multiresolution Image Processing and Analysis*, A. Rosenfeld (ed.), Springer–Verlag, New York, 200–210.