

Symmetry-Seeking Models and 3D Object Reconstruction

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

Abstract

We propose models of 3D shape which may be viewed as deformable bodies composed of simulated elastic material. In contrast to traditional, purely geometric models of shape, deformable models are active—their shapes change in response to externally applied forces. We develop a deformable model for 3D shape which has a preference for axial symmetry. Symmetry is represented even though the model does not belong to a parametric shape family such as (generalized) cylinders. Rather, a symmetry-seeking property is designed into internal forces that constrain the deformations of the model. We develop a framework for 3D object reconstruction based on symmetry-seeking models. Instances of these models are formed from monocular image data through the action of external forces derived from the data. The forces proposed in this paper deform the model in space so that the shape of its projection into the image plane is consistent with the 2D silhouette of an object of interest. The effectiveness of our approach is demonstrated using natural images.

1 Introduction

The problem of reconstructing the shapes of 3D objects from their images has attracted much attention in computational vision. The development of powerful mathematical models of 3D shape is a reasonable approach toward the resolution of this problem. In this paper we propose shape models that may be viewed as deformable bodies composed of simulated elastic material. By coupling deformable tubes to deformable spines, we create models whose deformations exhibit an affinity for certain 3D symmetries. The shape constraints intrinsic to the model are formalized by strain energy functionals which dictate the elastic properties of the material [1] and by internal coupling forces which produce the desired symmetries.

In this paper we consider axial symmetry constraints inspired by the influential “generalized cylinder” model of shape [2]. However, two important features distinguish *symmetry-seeking deformable models* from generalized cylinders and from parametric shape families such as superquadrics [3, 4]. First, these other models are

purely geometric, hence *passive*. In contrast, our deformable models are *active*. They react to externally applied forces as one would expect real elastic objects to react because deformable models are governed by the variational principles of elasticity theory. A second advantage of our models is their ability to represent symmetry without sacrificing the possibility of representing detailed shape information. This stands in contrast with existing abstractions of symmetric shape, such as the parametric shape families or smoothed local symmetries [5].

The active response of deformable models suggests the following framework for object reconstruction: Model instances are formed from visual data by combining the internal shape constraints with whatever external constraints may be derived from the data. The models’ internal constraints are generic—while permitting the reconstruction of a wide variety of natural shapes, they serve to overcome the often highly underdetermined nature of the reconstruction problem. The external constraints are interpreted as external forces applied to the models. The framework is based on the dynamics of elastic

bodies and reconstruction algorithms can be obtained in a principled way by applying standard numerical solution techniques.

The symmetry-seeking model developed in this paper may be described informally as follows: Imagine a deformable sheet made of elastic material (a membrane-thin-plate hybrid). Take this sheet and roll it to form a tube. Next, pass through the tube a deformable wire spine made of similar material. At regularly spaced points along the spine, couple it to the tube with radially projecting (Hookean) spring forces so as to maintain the spine in more or less axial position within the tube. Include additional forces that coerce the tube into a quasi-symmetric shape around the spine. Finally, provide extra control over shape by introducing expansion/compression forces radiating from the spine. The rigidities of the spine and the tube can be controlled independently, and their natural rest metrics and curvatures can be either prescribed or modified dynamically. If the circumferential metric of the tube is set to zero, for instance, the tube will tend to contract around the spine, unless the other forces prevail. The model will shorten or lengthen as the longitudinal metrics of the tube and spine are modified. In short, a variety of interesting behavior (including visco-elasticity) can be obtained by adjusting the control variables designed into the model.

In this paper, the forces applied to the symmetry-seeking model are derived from monocular images. The applied forces deform a model in 3-space so that its projection into the image plane is consistent with the 2D silhouette of an object of interest. The model reconstructs 3D shapes by achieving a stable equilibrium between the externally applied forces and the internal stresses due to the model's deformation constraints. A user can guide the reconstruction process (e.g., push the model into a desired equilibrium) by applying additional forces on the model through interactive pointing devices, or by dynamically adjusting the model's material properties or the strengths of its symmetries.

Because we consider here the restricted case of computing 3D models of objects from a single monocular image, we tacitly assume that a viewpoint has been chosen wherefrom all significant object features are visible, and that the axis of the

object is not severely inclined away from the image plane. To simplify the extraction of silhouette information in the image-based force field computations, we consider images of objects with relatively subdued surface texture placed before a contrasting background. The particular image force field that we have chosen to experiment with is obtained by computing the image-intensity gradient. This is an admittedly simple force field, but our approach readily extends to any type of image forces, including those that are based on sophisticated image analysis.

Figure 1 illustrates the reconstruction of a crook-necked squash from its image. After the

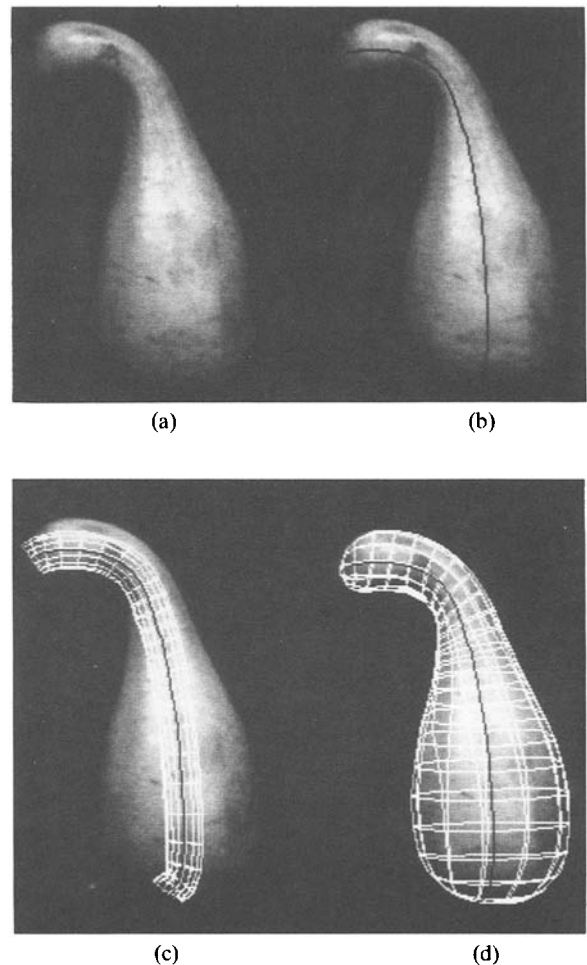


Fig. 1. Reconstruction of a 3D symmetry-seeking model. (a) Squash image. (b) User initialized spine, shown in black. (c) Initial tube. (d) Reconstructed model, displayed as wire-frames projected into the image.

user initializes the spine's projection in the image plane somewhere near the true axis of the object, the tube begins to inflate due to the internal expansion forces. As the model deforms, its configuration in space is dynamically projected onto the image plane where its boundaries are subject to image-based forces; specifically, they are attracted by and lock onto significant image-intensity gradients. Complementary computer vision and computer graphics operations are applied in synchrony during reconstruction [6]; (inverse) visual processing constrains and guides model formation, while (direct) graphical processing is responsible for the application of external forces and for providing the user with visual feedback.

In section 2 we compare our models to some common geometric models that have been employed for object representation. In section 3 we define the elastic curve and surface components of our model and give their strain energies of deformation. Section 4 defines the internal forces that couple these components together into the system of dynamic equations governing the symmetry-seeking model. In section 5 we formulate the image potential functions associated with the object reconstruction problem. In section 6 we present further results, including an example involving occluded objects. Finally, section 7 proposes extensions to our models and to our object reconstruction framework.

2 Comparison with Existing Shape Models

The symmetry-seeking model of 3D shape developed in this paper draws upon elastically deformable surface models for visual surface reconstruction [1, 7] and upon the idea of generalized cylinders first proposed by Binford [2], then implemented and further developed in several subsequent papers [8–14]. The generalized cylinder model is an intuitively appealing abstraction of elongated shapes exhibiting axial symmetry. The representation of symmetric shape has a long history in computer vision and continues to attract much interest [15–19].

Our model captures axial symmetry much like the generalized cylinder. However, we take seriously the fact that many objects of interest

possess only approximate symmetries. A major difference between our deformable model and generalized cylinders is that, while generalized cylinders *impose* symmetry constraints on objects they represent—inevitably with some loss of detailed shape information—our model expresses *preference* for certain symmetries through its internal forces, but it is not limited to shapes that possess these symmetries exactly.

The broad representational power of deformable models results by virtue of their *distributed* nature; every material point potentially contributes three spatial degrees of freedom. These degrees of freedom are mutually constrained, as tightly or as loosely as is desired, by the elastic strains internal to the model. As the internal forces are strengthened, the symmetry-seeking model will tend to impose the exact symmetries of a generalized cylinder. Conversely, by sufficiently weakening the internal forces, the model becomes capable of representing highly irregular objects including, in principle, quasi-symmetric objects with fractal-like surface structure. This assumes, of course, that sufficient constraint is maintained so that the reconstruction problem remains well-conditioned for the image data under consideration. Shape representations capable of applying constraint in this controllable manner are attractive for reconstruction and recognition; an early example is the spring-template-shape model proposed by Fischler and Elschlager [20].

The geometric coverage of deformable models can be significantly broader than the *lumped*-parameter families of shapes such as the superquadric models developed for computer graphics by Barr [3] and later advocated for use in computer vision by Pentland [4]. These lumped-parameter models are capable of accurately representing only a restricted class of objects because they “wire into the parameterization” a relatively small family of shapes, rather than place generic constraints on shape as do our deformable models. Lumped-parameter models cannot immediately accommodate most objects of interest, so objects must be hierarchically subdivided into compatible primitives that may then need to undergo shape transformations [21]. Reliably automating the subdivision and subsequent parameter-fitting steps remains a complex

and very open problem, especially when attempting to represent natural shapes for the purposes of vision [12, 22].

3 Deformable Curves and Surfaces

Let $\mathbf{x} = (x_1, \dots, x_p)$ be a point in a p -variate material coordinate system, and let (X, Y, Z) be a point in Euclidean 3-space. A 3D deformable model is given by the image of the unit p -cube in coordinate space $[0, 1]^p$ under the mapping $\mathbf{v}(\mathbf{x}) = (v_1(\mathbf{x}), v_2(\mathbf{x}), v_3(\mathbf{x}))$. We assume \mathbf{v} to be a regular mapping in the sense that the vectors $\mathbf{a}_\alpha = \partial \mathbf{v} / \partial x_\alpha$ for $\alpha = 1, \dots, p$ are linearly independent for all points $\mathbf{x} \in [0, 1]^p$.

A deformable space curve is a mapping from a univariate material coordinate space to Euclidean 3-space, while a deformable surface is a mapping from a bivariate coordinate space to 3-space. In either case, the mapping \mathbf{v} determines the configuration of the deformable model in 3-space. For instance, \mathbf{v} may represent a displacement away from a prescribed rest configuration. In this paper, however, it represents the 3-space positions of material points in the model; i.e.,

$$\begin{aligned} v_1(\mathbf{x}) &= X(\mathbf{x}) \\ v_2(\mathbf{x}) &= Y(\mathbf{x}) \\ v_3(\mathbf{x}) &= Z(\mathbf{x}) \end{aligned} \quad (1)$$

The deformable material is characterized by a functional $\mathcal{E}(\mathbf{v}(\mathbf{x}))$ which associates a non-negative strain energy with any admissible mapping. In our deformable models, \mathcal{E} is a p -variate instance of the controlled-continuity spline functionals defined in [7].

According to continuum mechanics, the non-rigid motion of a massless deformable body in a viscous ambient medium, in response to a net externally applied force $\mathbf{f}[\mathbf{v}(\mathbf{x})]$, is dictated by the first-order dynamic equation

$$\gamma \frac{\partial \mathbf{v}(\mathbf{x})}{\partial t} + \frac{\delta \mathcal{E}[\mathbf{v}(\mathbf{x})]}{\delta \mathbf{v}} = \mathbf{f}[\mathbf{v}(\mathbf{x})] \quad (2)$$

where γ is the dissipation factor of the medium [23]. The second term on the left-hand side of the equation is the variational derivative of the strain energy functional; it expresses the internal elastic

force of the model. The model achieves equilibrium when $\partial \mathbf{v} / \partial t = \mathbf{0}$.

3.1 A Deformable Spine Model

The spine model is a deformable space curve. To formulate it, we employ the scalar ($p = 1$) coordinate space $x = s \in [0, 1]$. The deformation energy associated with a configuration $\mathbf{v}(s) = [X(s), Y(s), Z(s)]$ is given by

$$\mathcal{E}^S(\mathbf{v}) = \frac{1}{2} \int_0^1 w_1(s) |\mathbf{v}_s|^2 + w_2(s) |\mathbf{v}_{ss}|^2 ds \quad (3)$$

where the subscripts of \mathbf{v} denote derivatives with respect to the independent variable s .

Ignoring boundary terms, the variational derivative of \mathcal{E}^S is given by

$$\frac{\delta \mathcal{E}^S}{\delta \mathbf{v}} = \frac{\partial^2}{\partial s^2} \left(w_2 \frac{\partial^2 \mathbf{v}}{\partial s^2} \right) - \frac{\partial}{\partial s} \left(w_1 \frac{\partial \mathbf{v}}{\partial s} \right) \quad (4)$$

The weighting functions control the material properties: $w_1(s)$ controls the tension along the spine, while $w_2(s)$ controls rigidity of the spine. In particular, setting $w_1(s_0) = w_2(s_0) = 0$ permits a position discontinuity and setting $w_2(s_0) = 0$ permits a tangent discontinuity to occur at s_0 [7].

We associate with the spine a metric function $L(s)$ which prescribes the natural arc length of the space curve as a function of s . This metric is introduced into (4) by defining $w_1(s) = \sqrt{X_s^2 + Y_s^2 + Z_s^2} - L(s)$. The natural curvature $C(s)$ is prescribed by defining $w_2(s) = \kappa(s) - C(s)$ where $\kappa(s)$ is actual curvature along the curve.

3.2 A Deformable Tube Model

The tube is composed of deformable sheet material. Using the $p = 2$ coordinate space $\mathbf{x} = (x, y) \in [0, 1]^2$, the deformation energy of the configuration $\mathbf{v}(x, y) = [X(x, y), Y(x, y), Z(x, y)]$ is given by the functional

$$\begin{aligned} \mathcal{E}^T(\mathbf{v}) &= \frac{1}{2} \int_0^1 \int_0^1 w_{1,0} |\mathbf{v}_x|^2 + w_{0,1} |\mathbf{v}_y|^2 \\ &\quad + w_{2,0} |\mathbf{v}_{xx}|^2 + 2w_{1,1} |\mathbf{v}_{xy}|^2 \\ &\quad + w_{0,2} |\mathbf{v}_{yy}|^2 dx dy \end{aligned} \quad (5)$$

where the subscripts of \mathbf{v} denote partial deriva-

tives with respect to the independent variables x and y .

Ignoring boundary terms, the variational derivative of \mathcal{E}^T is given by

$$\begin{aligned} \frac{\delta \mathcal{E}^T}{\delta \mathbf{v}} = & \frac{\partial^2}{\partial x^2} \left(w_{2,0} \frac{\partial^2 \mathbf{v}}{\partial x^2} \right) + 2 \frac{\partial^2}{\partial x \partial y} \left(w_{1,1} \frac{\partial^2 \mathbf{v}}{\partial x \partial y} \right) \\ & + \frac{\partial^2}{\partial s^2} \left(w_{0,2} \frac{\partial^2 \mathbf{v}}{\partial y^2} \right) \\ & - \frac{\partial}{\partial x} \left(w_{1,0} \frac{\partial \mathbf{v}}{\partial x} \right) - \frac{\partial}{\partial y} \left(w_{0,1} \frac{\partial \mathbf{v}}{\partial y} \right) \end{aligned} \quad (6)$$

The natural metric of the sheet along each material coordinate curve is introduced into (6) by defining $w_{1,0}(x, y) = \sqrt{X_x^2 + Y_x^2 + Z_x^2} - L_{1,0}(x, y)$ and $w_{0,1}(x, y) = \sqrt{X_y^2 + Y_y^2 + Z_y^2} - L_{0,1}(x, y)$. Analogous expressions for $w_{2,0}(x, y)$, $w_{1,1}(x, y)$, and $w_{0,2}(x, y)$ determine the sheet's natural curvatures, and discontinuities are inserted as described above (see [7] for details).

The tube is formed by prescribing boundary conditions on two opposite edges of the sheet that "seam" these edges together. We choose to seam the edge $x = 0$ to the edge $x = 1$, letting y span the length of the tube. The required periodic boundary conditions are

$$\mathbf{v}(0, y) = \mathbf{v}(1, y), \quad \mathbf{v}_x(0, y) = \mathbf{v}_x(1, y) \quad (7)$$

4 The Symmetry-Seeking Model

We introduce three forces of interaction between the tube and the spine: The first force coerces the spine into axial position within the shell, the second predisposes the tube to radial symmetry around the spine, and the third controls expansion or contraction of the tube around the spine.

To couple the two components, we first identify $y \equiv s$, bringing into correspondence the spine coordinate with the coordinate along the length of the tube (see figure 2). We then distinguish the configuration vector function of the spine $\mathbf{v}^S(s)$ from that of the tube $\mathbf{v}^T(x, s)$ with superscripts S and T .

We define $\bar{\mathbf{v}}^T(s) = l(s)^{-1} \int_0^1 \mathbf{v}^T(x, s) |\mathbf{v}_x^T(x, s)| dx$ to be the centroid of the coordinate curve ($s =$

to be the centroid of the coordinate curve ($s =$ constant) circling the tube, whose length is given by $l(s) = \int_0^1 |\mathbf{v}_x^T(x, s)| dx$. We also define the tube's radial vector function with respect to the spine as $\mathbf{r}(x, s) = \mathbf{v}^T(x, s) - \mathbf{v}^S(s)$, the unit radial vector function $\hat{\mathbf{r}}(x, s) = \mathbf{r}(x, s) / |\mathbf{r}(x, s)|$, and $\bar{r}(s) = l(s)^{-1} \int_0^1 |\mathbf{r}(x, s)| |\mathbf{v}_x^T(x, s)| dx$, the mean radius of the coordinate curve $s =$ constant.

The spine is coerced into an axial position within the tube by introducing the following forces on the spine and tube respectively:

$$\mathbf{f}_a^S(s) = a(s) [\bar{\mathbf{v}}^T(s) - \mathbf{v}^S(s)]$$

$$\mathbf{f}_a^T(x, s) = -a(s) l(s)^{-1} [\bar{\mathbf{v}}^T(s) - \mathbf{v}^S(s)] \quad (8)$$

where $a(s)$ controls the strength of the force.

To encourage the tube to be radially symmetric around the spine, we introduce the following force on the tube:

$$\mathbf{f}_b^T(x, s) = b(s) (\bar{r}(s) - |\mathbf{r}(x, s)|) \hat{\mathbf{r}}(x, s) \quad (9)$$

where $b(s)$ controls the strength of the force.

Finally, it is useful to provide control over expansion and contraction of the tube around the spine. This is accomplished by introducing the force

$$\mathbf{f}_c^T(x, s) = c(s) \hat{\mathbf{r}}(x, s) \quad (10)$$

where $c(s)$ controls the strength of the force. The tube will inflate wherever $c > 0$ and deflate wherever $c < 0$. In particular, the two open ends of the tube can be cinched shut by assigning large

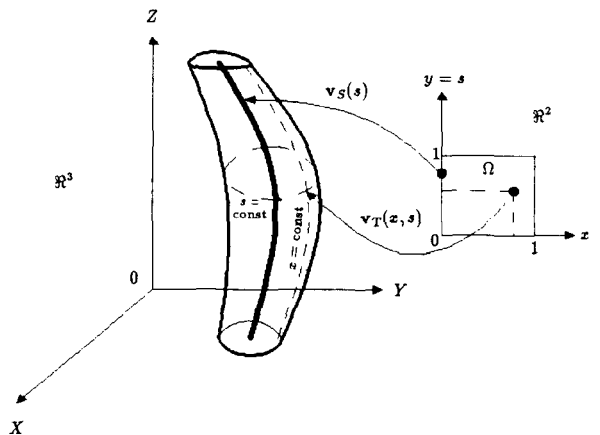


Fig. 2. The parameterization of the 3D model.

negative values to $c(0)$ and $c(1)$, thereby creating a sausage-like surface.

Summing the above coupling forces into the dynamic equations (2) associated with the spine and tube, we obtain the following coupled dynamic system describing the motion of the symmetry-seeking model:

$$\begin{aligned} \gamma \frac{\partial \mathbf{v}^S}{\partial t} + \frac{\delta \mathcal{E}^S}{\delta \mathbf{v}^S} &= \frac{\partial P^S}{\partial \mathbf{v}^S} + \mathbf{f}_a^S \\ \gamma \frac{\partial \mathbf{v}^T}{\partial t} + \frac{\delta \mathcal{E}^T}{\delta \mathbf{v}^T} &= \frac{\partial P^T}{\partial \mathbf{v}^T} + \mathbf{f}_a^T + \mathbf{f}_b^T + \mathbf{f}_c^T \end{aligned} \quad (11)$$

The elasticity terms on the left-hand side of these coupled equations are given by (4) and (6) (with $y \equiv s$ in the latter), and P^S and P^T are generalized potential functions associated with the external image-based forces applied to the spine and tube respectively.

5 Image Potentials

Given the image $I(\eta, \xi)$ of an object, our reconstruction approach applies image-based forces that deform the symmetry-seeking model to make its shape consistent with the object's silhouette (occluding contour) in the image. We formulate a force field of attraction towards significant image intensity gradients which, by assumption, characterize occluding contours in our restricted domain. The limb (occluding boundary) of the deformable tube is made sensitive to this force field. Although this simple force field which acts parallel to the image plane is our first attempt at illustrating the feasibility of our approach, we show in the next section that it nonetheless yields interesting results.

The coupling between the force field and the tube is through the external generalized potential function P^T . We define

$$P^T[\mathbf{v}^T(x, s)] = \psi(x, s) |\nabla \{G_\sigma * I(\Pi[\mathbf{v}^T(x, s)])\}| \quad (12)$$

which imparts on the tube boundary an affinity for steep image intensity changes. Here, $G_\sigma * I$ denotes the image convolved with a (Gaussian) smoothing filter whose characteristic width is σ .

The operation $\Pi[\mathbf{v}^T(x, s)]$ denotes a projection of the material point 3-space coordinates $[X(x, s), Y(x, s), Z(x, s)]$ into the image plane (η, ξ) . While a perspective projection is generally appropriate, we have obtained satisfactory results in the examples considered below using an orthographic projection $\Pi(X, Y, Z) \rightarrow (\eta, \xi): (\eta, \xi) = (X, Y)$.

The weighting function $\psi(x, s)$ is nonzero only for material points (x, s) near occluding boundaries of the tube, which are selected by setting

$$\psi(x, s) = \begin{cases} 1 - |\mathbf{i} \cdot \mathbf{n}(x, s)| & \text{if } |\mathbf{i} \cdot \mathbf{n}(x, s)| < t \\ 0, & \text{otherwise} \end{cases} \quad (13)$$

where $\mathbf{n}(x, s)$ is the unit normal of the tube at (x, s) , \mathbf{i} is the unit normal to the image plane, and t is a small threshold (nominally 0.05).

The spine of the model may be projected into the image in a similar fashion, where it can be subjected through P^S to its own set of forces, but our experiments did not exercise this feature.

6 Implementation and Results

To solve the reconstruction problem, we numerically simulate the nonlinear, first-order dynamic system (11) to equilibrium: $\mathbf{u}^S, \mathbf{u}^T$ such that $\partial \mathbf{u}^S / \partial t = 0 = \partial \mathbf{u}^T / \partial t$. The dynamic system is discretized using standard numerical methods [24]. Finite difference techniques are used to discretize the time derivatives and the variational derivatives $\delta \mathcal{E}^S[\mathbf{v}^S(s)] / \delta \mathbf{v}^S$ and $\delta \mathcal{E}^T[\mathbf{v}^T(x, s)] / \delta \mathbf{v}^T$ (see [1] and [25]). Finite differences and numerical integration are used to compute the internal coupling forces. The image force $\partial P^T[\mathbf{v}^T(x, s)] / \partial \mathbf{v}^T$ is computed numerically in the image domain (η, ξ) using bilinear interpolation and first-order central differences.

We have employed successfully both an explicit Euler time-integration scheme and an analogue to successive-overrelaxation to solve the discrete-time equations to equilibrium. These iterative solution methods require only local operations, and the former is a parallel scheme. We have observed a significantly greater range of stability with an implicit Euler time-integration scheme in conjunction with direct solution methods [26]. We have also had success with the alternating direction implicit method, a hybrid

technique that combines some of the benefits of direct and iterative schemes.

Our first example of constructing a 3D model involves the squash image shown in figure 1(a). Figure 3 shows several frames from an animation sequence of the reconstruction process, followed by a rigid rotation of the resulting 3D model in space. The 3D model is rendered as a shaded space.

The second example involves the reconstruction of two quasi-symmetric objects, a pear and a potato, from a single image. Figure 4(a) shows the grey-level image of the still life scene. Figure 4(b) shows the initial model configuration, manually specified by the user. As shown in figure 4(c, d),

our present techniques allow us to reconstruct complete 3D models of the objects. By virtue of the symmetry-seeking constraints intrinsic to the model, the invisible portions of surfaces have been smoothly extrapolated from portions in view. Figure 5 compares the 3D model reconstruction with a 2.5D visible-surface reconstruction of the still life scene. The 2.5D reconstruction of the stereo images shown was carried out using a stereo algorithm developed in an earlier paper [25].

Notice that the potato partially occludes the pear in the image. The results demonstrate that our approach can accommodate partially occluded objects (incomplete boundaries). To han-

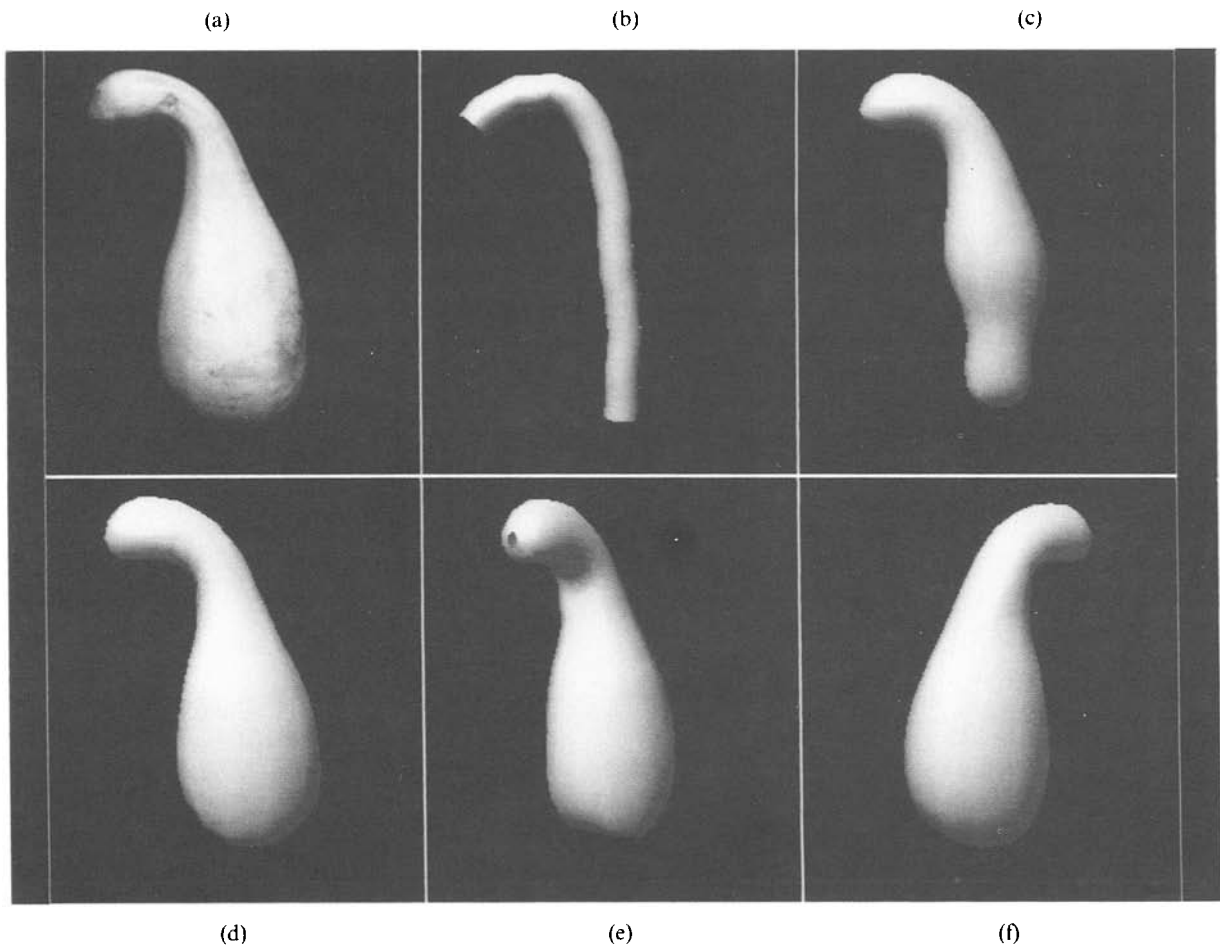


Fig. 3. Reconstruction of a squash. (a) Squash image. Selected frames from an animation sequence are shown: (b) Initial state of the 3D model. (c) Intermediate shape during reconstruction. (d) Final reconstructed model. (e, f) Model rotating rigidly in space. The 3D model is rendered as a shaded shell.

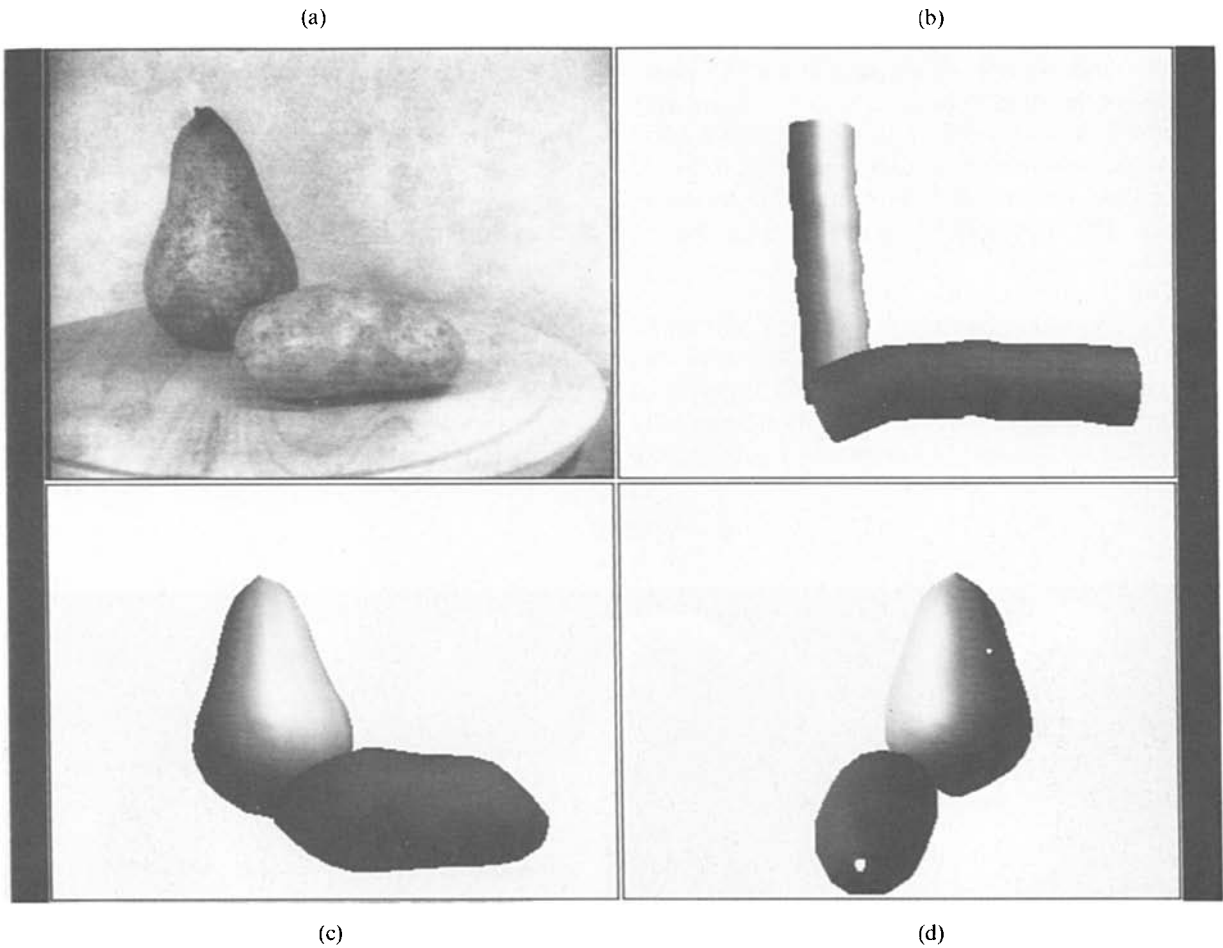


Fig. 4. Reconstruction of a still-life scene with symmetry-seeking deformable models. (a) Image of the scene. (b) Initial, user-specified configurations of the 3D models. Two graphically rendered views of the reconstructed still life: (c) frontal view of the 3D models; (d) side view.

dle occlusions, we inhibit the image forces from acting on the partially occluded model wherever its projection onto the image plane happens to be obscured by the projection of an occluding model. We use a 3D ray-casting technique in conjunction with the projection operation Π in (12) in order to determine those portions of models that are occluded by other models and must therefore be inhibited. The ray-casting operation requires knowledge of the relative depth ordering of the objects. Since our current reconstruction process is driven by only a single monocular image, this information is not available directly (but see [27]). The user presently specifies the

relative depth ordering; however, local depth ordering information can be obtained through automatic means [28].

It is our experience that the user need not initialize the spine (or shell) of the model with any great accuracy. Typically, the user draws a rough medial axis running more or less the length of the object of interest in the image. From such an initial configuration, the reconstruction process can successfully make the necessary adjustments to the spine and shell. It may be possible to automatically instantiate a reasonable spine by applying medial axis [15] or smoothed local symmetry [5] algorithms to the image.

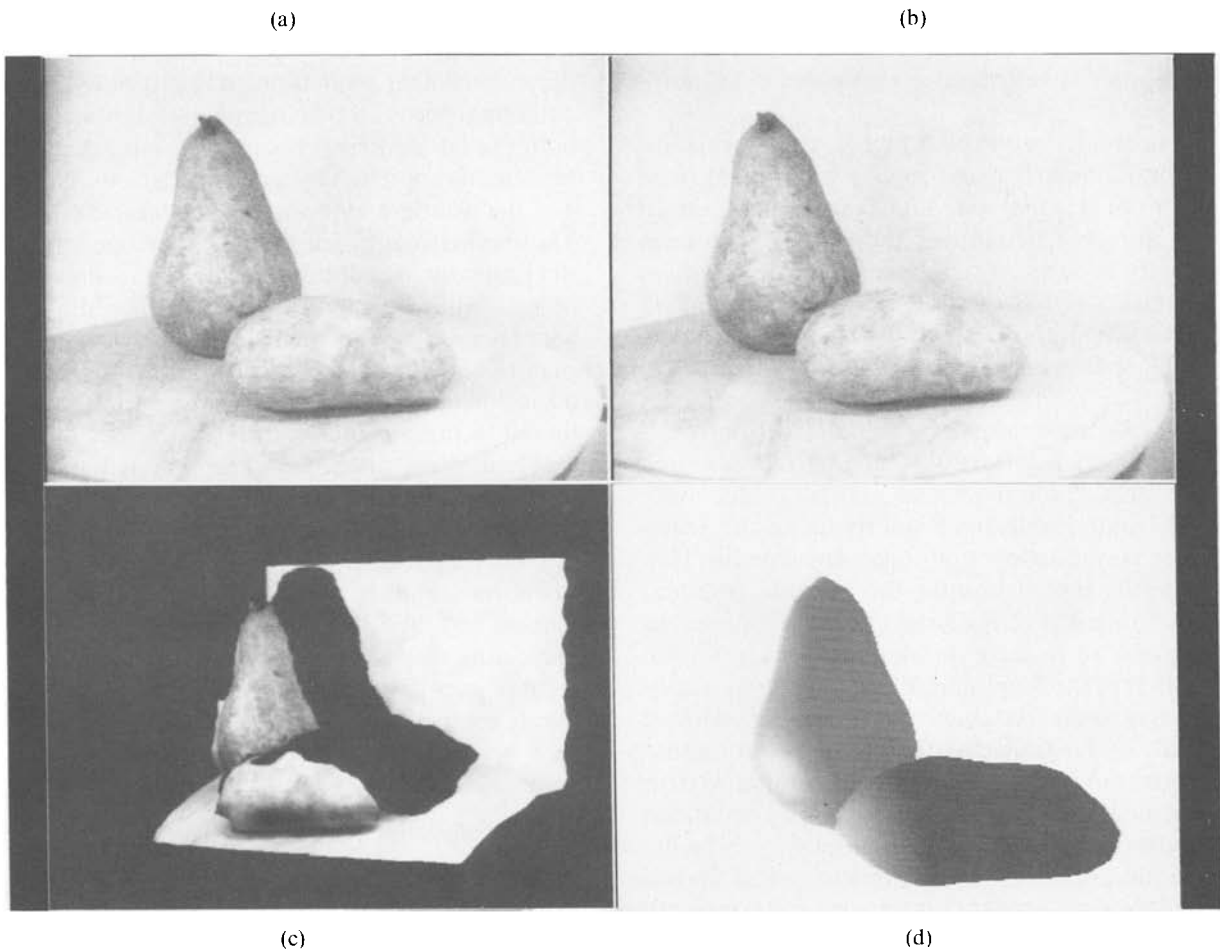


Fig. 5. Comparison of the 3D model reconstruction with a 2.5D visible-surface reconstruction of the still life scene. (a, b) Stereo images of scene. (c) 2.5D reconstruction using stereo matching techniques (from [25]). (d) 3D model reconstruction.

7 Extensions

The symmetry-seeking model developed in this paper has been designed to represent elongated objects that possess quasi-axial symmetry, and test objects have therefore been chosen appropriately. By introducing other internal forces into the model, however, preferences can be induced not only for the axial symmetries studied here, but for any sort of regularity, such as spherical or mirror symmetry. Furthermore, it appears possible to incorporate a weighted combination of forces associated with multiple symmetry states into a single deformable model. An interesting goal would be to have the *multisymmetric deformable*

model autonomously “infer” the symmetry state most appropriate for an unknown object under consideration, possibly by computing the solution to an optimization problem in the weight-parameter space.

One of the most important virtues of the deformable models developed in this paper is the possibility of using their underlying variational formulations to integrate multiple sources of visual information [29]. Physically, each information source makes a contribution to the net force field acting on the model. Information derived from higher-level sources—attentional mechanisms, top-down knowledge-based inference processes, a human operator, etc.—can

be used to instantiate the model and to guide the reconstruction process. The same is true regarding low-level information computed over multiple scales.

The model can exploit multiscale information to autonomously avoid grossly suboptimal fixed points of (11) that may hinder reconstruction. In our present technique, suboptimal equilibria usually arise from significant intensity gradient fluctuations that can intervene between the initial configuration of the shell and the silhouette of the object (e.g., due to surface markings; however, the internal expansion forces routinely drive the model over nominal intensity fluctuations arising from subdued surface texture). The reconstruction process may be guided away from undesirable fixed points using the scale-space continuation technique developed in [25]. To apply this technique, the external potential function given in equation (12) is interpreted as defining an *embedding*, where the characteristic width σ of the Gaussian filter G_σ serves as a continuous scale parameter for a parameterized family of dynamic equations (11). The continuation method yields an extended dynamic system governed by these parameterized equations coupled to an additional equation governing the evolution of σ . This deterministic system finds a good fixed point of (11) at a coarse scale (large σ), then tracks it continuously to a fine scale (small σ), avoiding suboptimal fixed points along the way.

As our examples show, 3D deformable models can be reconstructed directly from an image using meagre quantities of information. Although the computation of intervening "intrinsic image" or 2.5D surface representations may not be strictly necessary [30], there is no reason why the deformable model cannot beneficially assimilate quantitative or qualitative information from these types of representations. Similarly, stereo, motion, shading, and texture information may be integrated into a richly redundant force field which would sufficiently constrain the model over a more general class of reconstruction problem than the one considered here.

We have recently integrated stereo and motion information into the symmetry-seeking model [27]. The 3D model undergoes nonrigid motion in space, subject to forces emanating from dy-

namic stereoscopic images. The stereo forces move model boundary points laterally and in depth such that their binocular projections are consistent with the evolving object silhouettes in both the left and right images. We have demonstrated this approach by reconstructing the shape and the nonrigid motion of a 3D natural object. The explicit depth information extracted from the stereo force enables the model to reconstruct objects whose axes may be oriented obliquely with respect to the image plane. This depth information is also employed to determine interobject occlusion for the ray-casting operation mentioned in the previous section.

When mass distributions are added to our formulations, deformable models inherit realistic second-order dynamics [26]. These models not only appear promising for reconstructing free-form natural objects undergoing nonrigid motion, but they also offer a powerful approach to creating realistic graphical animation of simulated objects possessing such properties [31]. The purely geometric models in common use offer no such possibilities.

Acknowledgments

We would like to thank John Platt for amusing discussions and for contributing to our ideas and Kurt Fleischer for helping to render the figures.

References

1. D. Terzopoulos, "Multilevel computational processes for visual surface reconstruction," *COMPUTER VISION, GRAPHICS, AND IMAGE PROCESSING*, vol. 24, pp. 52-96, 1983.
2. T.O. Binford, "Visual perception by a computer," Presented at the IEEE CONF. ON SYSTEMS AND CONTROL, Miami, Fla., 1971.
3. A. Barr, "Superquadrics and angle-preserving transformations," *IEEE COMPUTER GRAPHICS AND APPLICATIONS*, vol. 18, pp. 21-30, 1981.
4. A.P. Pentland, "Parts: Structured descriptions of shape," in *PROC. NAT. CONF. ARTIF. INTEL., AAAI-86*, Philadelphia, Pa., 1986, pp. 695-701.
5. J.M. Brady and H. Asada, "Smoothed local symmetries and their implementation," *INT. J. ROBOTICS RES.*, vol. 3, pp. 33-61, 1984.
6. A. Witkin, M. Kass, D. Terzopoulos, and A. Barr, "Linking perception and graphics: Modeling with dynamic constraints," in *IMAGES AND UNDERSTANDING*, H.

- Barlow, C. Blakemore, and W. Weston-Smith (eds.), Cambridge University Press, in press.
7. D. Terzopoulos, "Regularization of inverse visual problems involving discontinuities," *IEEE TRANS. PAMI*, vol. 8, pp. 413-424, 1986.
 8. G.A. Agin and T.O. Binford, "Computer description of curved objects," in *PROC. THIRD INT. JOINT CONF. ARTIF. INTELL.*, Stanford, Calif., 1973, pp. 629-635.
 9. J.M. Hollerbach, Hierarchical shape description of objects by selection and modification of prototypes, M.I.T. Artif. Intel. Lab., Cambridge, Mass., AI-TR-346, 1975.
 10. R. Nevatia and T.O. Binford, "Description and recognition of curved objects," *ARTIFICIAL INTELLIGENCE*, vol. 8, pp. 77-98, 1977.
 11. D. Marr, "Analysis of occluding contour," *PROC. R. SOC. (London)*, B, vol. 197, pp. 441-475, 1977.
 12. D. Marr and H.K. Nishihara, "Representation and recognition of the spatial organization of three-dimensional shapes," *PROC. R. SOC. (London)*, B, vol. 200, pp. 269-294, 1978.
 13. R.A. Brooks, "Symbolic reasoning among 3-D models and 2-D images," *ARTIFICIAL INTELLIGENCE*, vol. 17, pp. 285-348, 1981.
 14. G.K. Rao and R. Nevatia, "Generalized cone descriptions from sparse 3-D data," *PROC. DARPA IMAGE UNDERSTANDING WORKSHOP*, Miami Beach, Fla., L.S. Baumann (ed.), 1985, pp. 497-505.
 15. H.A. Blum, "A transformation for extracting new descriptions of shape," in *PROC. SYMP. MODELS PERCEPT. SPEECH VISUAL FORM*, Boston, Mass., 1964, pp. 362-380.
 16. T. Kanade, "Recovery of the three dimensional shape of an object from a single view," *ARTIFICIAL INTELLIGENCE*, vol. 17, pp. 409-460, 1981.
 17. J.M. Brady, "Criteria for representations of shape," in *HUMAN AND MACHINE VISION*, A. Rosenfeld and J. Beck (eds.), Academic Press, New York, NY, 1983, pp. 39-84.
 18. W.A. Richards and D.D. Hoffman, "Codon constraints on closed 2D shapes," *COMPUTER VISION, GRAPHICS, AND IMAGE PROCESSING*, vol. 31, pp. 265-281, 1985.
 19. M. Leyton, "Constraint theorems on the prototypification of shape," *PROC. NAT. CONF. ARTIFI. INTEL.*, AAAI-86, Philadelphia, Pa., 1986, pp. 702-706.
 20. M.A. Fischler and R.A. Elschlager, "The representation and matching of pictorial structures," *IEEE TRANS. COMPUTERS*, C-22, pp. 67-92, 1973.
 21. A.H. Barr, "Global and local deformations of solid primitives," *COMPUTER GRAPHICS*, vol. 18(3), pp. 21-29, 1984 (Proc. SIGGRAPH'84).
 22. A.P. Pentland, "Recognition by parts," *PROC. FIRST INT. CONF. COMPUTER VISION*, London, UK, 1987, pp. 612-620.
 23. Y.C. Fung, *FOUNDATIONS OF SOLID MECHANICS*, Prentice-Hall: Englewood Cliffs, NJ, 1965.
 24. G. Dahlquist and A. Björck, *NUMERICAL METHODS*, N. Anderson (trans.), Prentice-Hall: Englewood Cliffs, NJ, 1974.
 25. A. Witkin, D. Terzopoulos, and M. Kass, "Signal matching through scale space," *INT. J. OF COMPUTER VISION*, vol. 1(2), pp. 133-144, 1987.
 26. D. Terzopoulos, "On matching deformable models to images: Direct and iterative solutions," *TOPICAL MEETING ON MACHINE VISION, TECHNICAL DIGEST SERIES*, vol. 12, Optical Society of America: Washington, D.C., March, 1987, 160-167.
 27. D. Terzopoulos, A. Witkin, and M. Kass, "Energy constraints on deformable models: recovering shape and non-rigid motion," *PROC. SIXTH NAT. CONF. ARTIFI. INTEL.*, AAAI-87, Seattle, Wash., July, 1987, pp. 755-760.
 28. D. Rosenberg, M.D. Levine, and S.W. Zucker, "Computing relative depth from occlusion cues," *PROC. FOURTH INT. J. CONF. PATTERN RECOGNITION*, Tokyo, 1978.
 29. D. Terzopoulos, "Integrating visual information from multiple sources," in *FROM PIXELS TO PREDICATES: Recent Advances in Computational and Robotic Vision*, A.P. Pentland (ed.), Ablex: Norwood, NJ, 1986, pp. 111-142.
 30. D.G. Lowe, *PERCEPTUAL ORGANIZATION AND VISUAL RECOGNITION*, Kluwer: Boston, Mass., 1985.
 31. D. Terzopoulos, J. Platt, A. Barr, and K. Fleischer, "Elastically deformable models," *COMPUTER GRAPHICS*, 3 vol. 21, 1987, (Proc. SIGGRAPH'87), pp. 205-214.