Constraints on Deformable Models: Recovering 3D Shape and Nonrigid Motion*

Demetri Terzopoulos, Andrew Witkin and Michael Kass

Schlumberger Palo Alto Research, 3340 Hillview Avenue, Palo Alto, CA 94304, U.S.A.

Recommended by AAAI-87 Program Committee

ABSTRACT

Inferring the 3D structures of nonrigidly moving objects from images is a difficult yet basic problem in computational vision. Our approach makes use of dynamic, elastically deformable object models that offer the geometric flexibility to satisfy a diversity of real-world visual constraints. We specialize these models to include intrinsic forces inducing a preference for axisymmetry. Image-based constraints are applied as extrinsic forces that mold the symmetry-seeking model into shapes consistent with image data. We describe an extrinsic force that applies constraints derived from profiles of monocularly viewed objects. We generalize this constraint force to incorporate profile information from multiple views and use it to exploit binocular image data. For time-varying images, the force becomes dynamic and the model is able to infer not only depth, but nonrigid motion as well. We demonstrate the recovery of 3D shape and nonrigid motion from natural imagery.

1. Introduction

A primary goal of early vision is to recover the shapes and motions of 3D objects from their images. To achieve this goal, we must synthesize visual models that satisfy a bewildering variety of constraints. Some constraints derive from the sensory information content of images. Others reflect background knowledge about image formation and about the shapes and behaviors of real-world objects. Exploiting diverse constraints in combination has proven to be a challenge. We need models which not only integrate constraints, but which escape the confines of conventional representations that impose simplify-

---

* This paper was the most outstanding paper in the subfield "Perception-Vision" at the AAAI-87 Conference.

ing assumptions about shape and motion. Computational vision calls for general-purpose models having the capability to accurately represent the free-form shapes and nonrigid motions of natural objects—objects with which the human visual system copes routinely. Clearly, we need new models that can accommodate deformation, nonconvexity, nonplanarity, inexact symmetry, and a gamut of localized irregularities.

We propose a physically based modeling framework for shape and motion reconstruction of free-form flexible objects from their images. In this framework, objects are modeled as elastically deformable bodies subject to continuum mechanical laws. Constraints are expressed as forces applied to these bodies. The applied forces deform the elastic models and propel them through potentially complicated motions such that they satisfy the available constraints over time. We develop algorithms for inferring from natural images the structures and motions of flexible objects moving nonrigidly in three dimensions. The algorithms compute detailed 3D object models directly from image intensity data without making use of intermediate optic-flow fields or 2.5D surface representations.

To reconstruct models directly from natural images that possibly involve significant occlusions, we must exploit several powerful constraints in unison. Our physical models focus the constraints in a natural way—by summing together the associated forces. There are two types of forces: Intrinsic forces encode constraints internal to our deformable models. Extrinsic forces couple the models to the external image data and provide an avenue for user interaction.

1.1. Intrinsic constraints

The intrinsic constraints reflect generically valid assumptions about natural objects. Our deformable models apply a basic constraint that is characteristic of physical bodies: surface coherence. The constraint is inherent in the elastic forces prescribed by the physics of deformable continua—these forces elicit piecewise continuous deformations.

A second generic constraint built into our models is symmetric regularity, an attribute of many natural and synthetic objects. Rather than imposing strict symmetries through explicit parameterization, we design more liberal *symmetry-seeking* intrinsic forces. These forces constrain the deformations of the model in order to give it a preference for certain desired symmetries. Representing symmetry as constrained deformation rather than through geometric parameterization frees the model from the shackles of particular parametric shape families such as, say, the quadrics—spheres, cylinders, ellipsoids, etc.

Our work to date considers the reconstruction of the 3D shape and nonrigid motion of objects possessing approximate axial symmetry. Our axisymmetry-seeking model of shape is essentially a deformable tube surrounding and coupled to a deformable spine. The coupling results from intrinsic forces that

imbue the combined model with a preference for axisymmetry. In this regard, our model is close in spirit to the generalized cylinder representation first recommended in 1971 by Binford as a convenient description of 3D surfaces for the purposes of vision [1].

Generalized cylinders often are overly restrictive in that they can represent with accuracy only perfectly axisymmetric shapes. Since some amount of asymmetry is evident in many synthetic and most natural shapes, generalized cylinders can result in a loss of crucial information about objects. By contrast, the symmetry-seeking model accommodates deviations from symmetry by deforming. Only as the intrinsic forces are strengthened does the symmetry-seeking model tend to impose the strict symmetries of a generalized cylinder. As the intrinsic forces are weakened, however, the model will be able to faithfully represent increasingly asymmetric shapes, although axisymmetric shapes have greater stability and hence are preferred.

1.2. Extrinsic constraints

The extrinsic constraints reflect, in part, observations about the environment that can be extracted from sensory data. Although, in principle, we can exploit within our framework a variety of image-based cues, including shading and texture, the present paper makes exclusive use of information about profiles (the profile of an object, also known as its occluding contour, refers to the curve which outlines the image region covered by the projection of the object). The human visual system has a remarkable ability to infer the 3D shapes of objects from their 2D profiles in images. To demonstrate this ability, David Marr was fond of showing Picasso's "Rites of Spring," which consists entirely of silhouettes (Fig. 1).

Silhouette information in confluence with the a priori constraints intrinsic to symmetry-seeking deformable models proves sufficient to recover 3D shapes. For example, Fig. 2 illustrates the reconstruction of Picasso's "Rites" directly from the silhouettes in Fig. 1 using 43 instances of the symmetry-seeking model. The symmetry-seeking models are embedded in a force field which encodes the profile information. The ambient forces mold the deformable models to make their 3D shapes consistent with the observed 2D profiles of objects or their subparts. Perfect axisymmetries are generally absent from the reconstructed shapes, as dictated by the artistic silhouettes.

Another possible source of extrinsic constraints is a human operator. We augment the ambient force field with forces controlled by computer pointing devices, thereby providing opportunity for a user to willfully guide the reconstruction process. We can create symmetry-seeking models and pull or push them through space while we monitor their shapes and motions as wire-frame projections on the image plane(s). As we bring models near imaged objects of interest, we observe them reconstructing the detailed shapes of these objects.

Fig. 1. "'Rites of Spring' by Pablo Picasso. We immediately interpret such silhouettes in terms of particular three-dimensional surfaces—this despite the paucity of information in the image itself. In order to do this, we plainly must invoke certain a priori assumptions and constraints about the nature of the shapes." (D. Marr [2, p. 213]).

Fig. 2. 3D rendition of Picasso's "Rites." This 3D reconstruction employs 43 instances of the symmetry-seeking model. The instances were reconstructed semi-automatically from the silhouettes in Fig. 1 (see text).

On a sufficiently powerful computer, deformable models offer an interactive modeling medium of practical interest in its own right. Nonetheless, our long-range goal in the context of machine vision is to replace the user with fully automatic top-down control processes, possibly linked to knowledge bases.

1.3. The reconstruction method

In this paper, we first review the special case of shape recovery from a static monocular view [3], then we generalize our approach to shape and motion recovery from input data consisting of a temporal sequence of binocular image pairs.

Figure 3 illustrates the reconstruction of a crook-necked squash from its monocular image using a symmetry-seeking model. The user initializes the
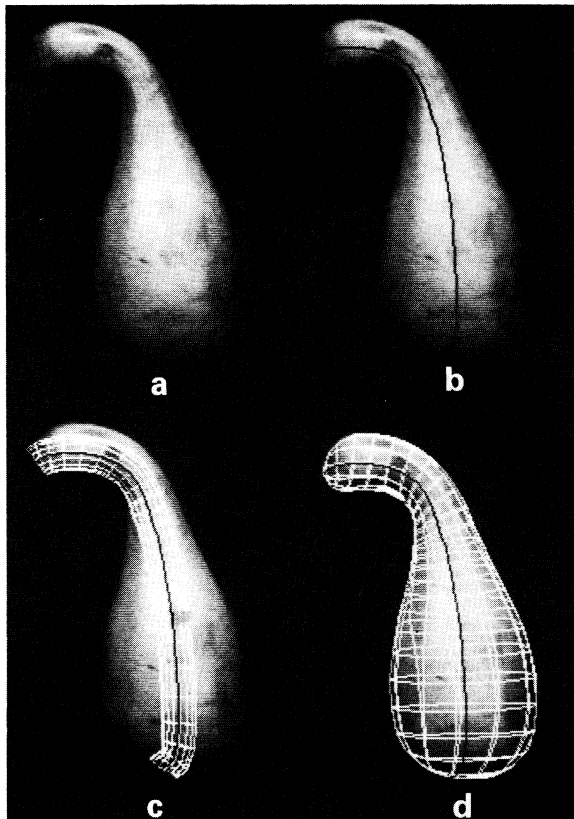

Fig. 3. Reconstruction of a 3D symmetry-seeking model. (a) Squash image. (b) User-initialized spine shown in black. (c) Initial tube. (d) Reconstructed model displayed as a wire frame projected into the image.

model by specifying the projection of the spine in the image plane near the medial axis of the object. In the monocular case the model is subject to extrinsic forces expressed as the gradient of an image potential function. The potential is a measure of the local contrast in the image after an appropriate smoothing transformation. Hence, the high contrast contour in the image (by assumption, the profile of the object) attracts the occluding boundary of the model (the occluding boundary of a 3D solid refers to the locus of points along which lines of sight graze its surface). The shape of model achieves a fixed point in the ambient force field. At equilibrium, the model's occluding boundary, relative to the viewpoint associated with the image is consistent with the shape of the object profile in the image. The model's intrinsic continuity and symmetry forces specify 3D shape over the remainder of its surface. Figure 4 sketches the monocular reconstruction scenario.

The image potential can be generalized to exploit more complete geometric information provided by profiles in multiple images acquired from different viewpoints around an object. For the particular case of stereo, the potential incorporates two images from slightly different vantage points. Deprojection of its gradient through a binocular camera model creates a stereo force field in space. Points on the model's occluding boundaries with respect to both the left and right eye are sensitive to the stereo force field. The forces position boundary points laterally and in depth such that their binocular projections coincide as much as possible with object profiles in both images. By attending
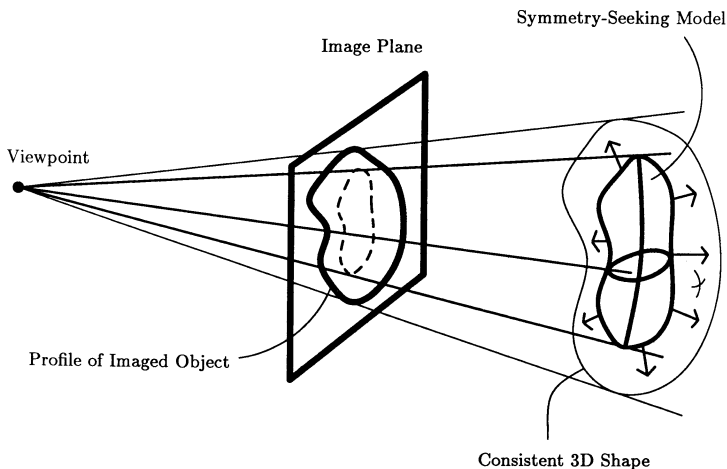

Fig. 4. Monocular reconstruction scenario. The arrows depict extrinsic forces in space which act on the symmetry-seeking model's occluding boundary as seen from the viewpoint. The forces deform the 3D model so as to make its image plane projection (dotted curve) more consistent with the 2D profile of the imaged object. The model comes to equilibrium as soon as its 3D shape achieves maximal consistency with the image data.

to the occluding boundary associated with each image individually, our method overcomes the difficulties that boundaries of smooth objects are known to present to conventional stereo matching techniques.

When the objects under consideration move, the ambient force field becomes dynamic. It carries the model through nonrigid motions, continually molding its shape to maintain maximal consistency with the evolving image data. The evolution of the model is computed by numerically integrating the partial differential equations of motion for the deformable body as it reacts to the dynamic force field.

The remainder of this paper is organized as follows: Section 2 discusses our approach relative to other work in vision and modeling. Section 3 describes the geometry and dynamics of the deformable symmetry-seeking model. Section 4 describes the image forces; first the monocular force, then the more general motion–stereo force. Section 5 briefly overviews the implementation of the model reconstruction algorithm. Section 6 presents results. Section 7 concludes the paper with a discussion.

2. Background

2.1. Constraint-based modeling

The work in this paper develops further a constraint-based modeling paradigm which has been successful on a variety of problems in computer graphics and animation as well as in computer vision: Terzopoulos et al. [4] apply physical constraints to deformable curve, surface, and solid models to construct and animate computer graphics objects made of simulated rubber, cloth, and other flexible materials. Witkin et al. [5] apply geometric constraints to parameterized shape primitives such as cylinders or spheres to automatically dimension, assemble, and animate objects constructed from such parts. Barzel and Barr [6] assemble articulated objects with dynamic constraints and simulate these objects with accurate Newtonian dynamics.

Witkin et al. [7] subject a deformable sheet in image coordinates to constraint forces derived from area correlation to perform stereo reconstruction in the style of the 2.5D sketch (Fig. 5(c)). Terzopoulos et al. [3] use symmetry-seeking models in a limited way to perform object reconstruction from static monocular profiles (Fig. 5(d)). Kass et al. [8] apply image-based constraints to deformable planar curves, dubbed "snakes," to interactively locate and track edges and other image features (Fig. 6). Platt [9] extends a deformable space curve model into a space-time surface, and uses it to recover rigid motion.

2.2. Comparison to conventional models

The elastically deformable models developed in this paper have evolved from variational models for visual surface reconstruction [10]. A number of features
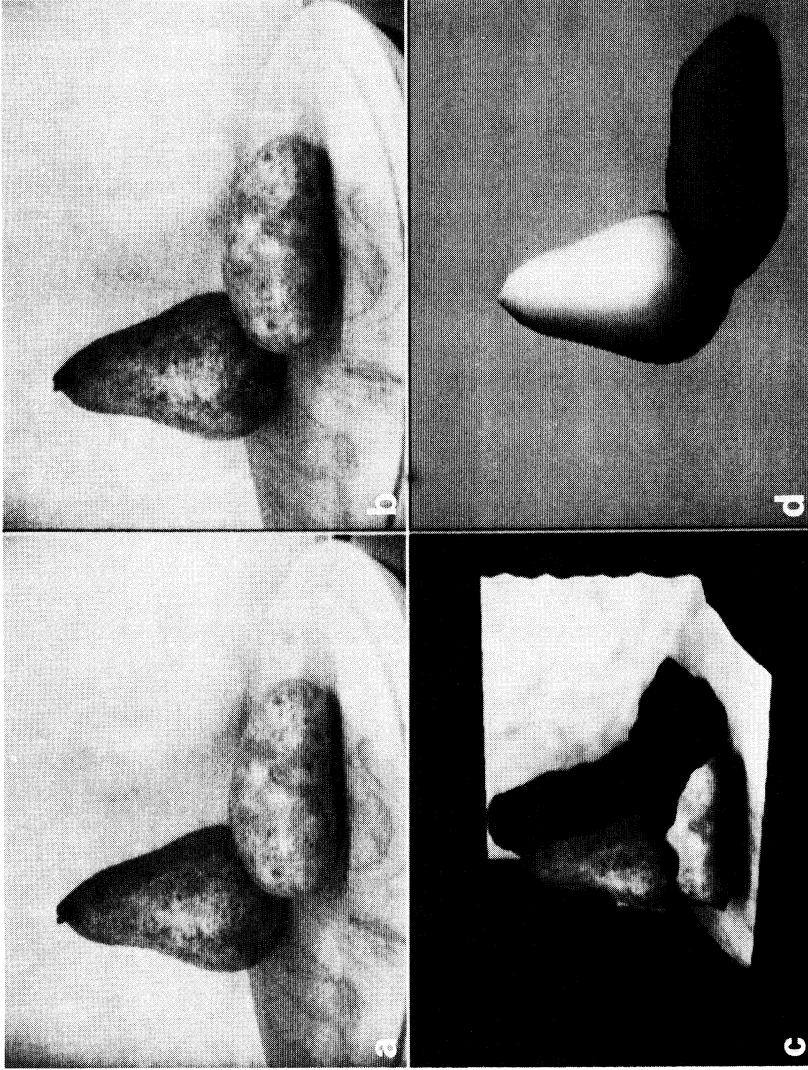
Fig. 5. Reconstructions of a still life scene. (a)–(b) Stereo images. (c) 2.5D reconstruction of stereo pair using a deformable sheet disparity model (from [7]). Left image is mapped on the reconstructed surface which is rendered from an oblique viewpoint to show depth discontinuities as gaps in surface. (d) 3D reconstruction of objects in left image using a symmetry-seeking deformable model (from [3]).
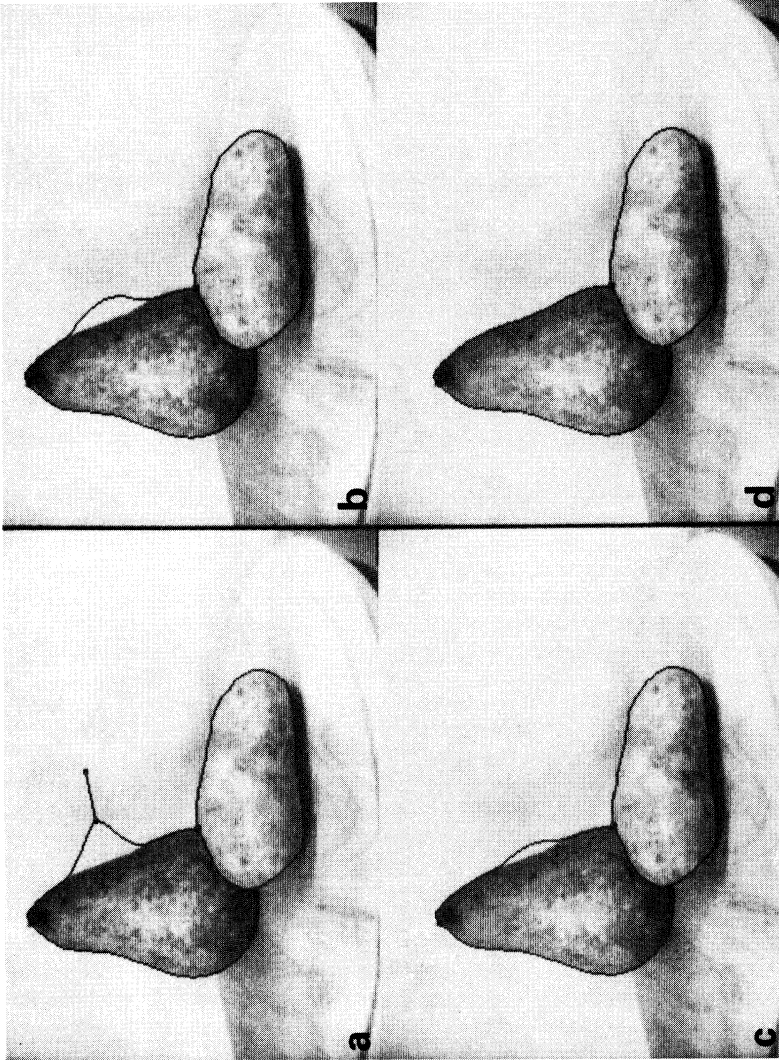
Fig. 6. Deformable planar contour models ("snakes"—see [8]). Snakes are elastic contour models subject to extrinsic forces that attract them towards image features such as lines and edges. After allowing two snakes to capture the outline of the pear and the potato, the user pulls one of the snakes away from the edge of the pear using an interactive spring force. When the user releases the snake, it is pulled back to the edge of the pear.

distinguish our modeling approach from the norm in computational vision. The comparison is summarized by Table 1.

Conventional models of 3D shape are purely geometric, hence passive. By contrast, our deformable models are active. They react to extrinsic forces as one would expect real elastic objects to react to physical forces. This is because deformable models are governed by physics; specifically, by the principles of elasticity theory as expressed through Lagrangian dynamics. Not only are these models suitable for reconstructing free-form natural objects undergoing non-rigid motion, but they also suggest a powerful approach to creating realistic graphical animation of simulated objects possessing such properties [4]. The purely geometric models in common use, being kinematic, offer no such possibilities.

The *distributed* nature of deformable models enhances their representational power. Every material point potentially contributes three spatial degrees of freedom which are mutually constrained with variable tightness by the elastic forces intrinsic to the model. Consequently, the geometric coverage of deformable models can be significantly broader than the *lumped*-parameter families of shapes such as the superquadric models employed in computer graphics by Barr [11] and advocated for use in computer vision by Pentland and others [12, 13]. These lumped-parameter models are capable of accurately representing only a restricted class of artificial 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 natural objects of interest, so precise hierarchical subdivision and parameterized deformations become practical necessities to contend with [14, 15]. By contrast, the free-form flexibility of our deformable models renders them immediately adaptable to natural shapes.

Our symmetry-seeking model is inspired by the idea of generalized cylinders first proposed by Binford [1], then implemented and further developed in several subsequent papers [16–22]. The generalized cylinder model is an intuitively appealing abstraction of elongated, axisymmetric shapes. Our model captures axisymmetry much like the generalized cylinder; however, we take

Table 1
Deformable models versus conventional models

Deformable models	Conventional models
Physics and geometry	Only geometry
Active	Passive
Dynamic	Kinematic
Distributed	Lumped parameter
Broad coverage	Narrow coverage
Controlled constraints	Strict constraints

seriously the fact that many objects of interest are only approximately symmetric.

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—the forces intrinsic to our model express *preference* for certain symmetries, but the model is not limited to strictly symmetric shapes. As the intrinsic forces are strengthened, the symmetry-seeking model will tend to impose the strict symmetries of a generalized cylinder. Conversely, by weakening the intrinsic forces, the model acquires the capacity to represent irregular objects with significant fine structure (assuming, 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 desirable for reconstruction and recognition; an early example is the spring–template shape model proposed by Fischler and Elschlager [23].

2.3. Recovery of nonrigid motion

Among the contributions of the present paper is a technique for recovering from image data the structure and motion of nonrigid objects in space. The significantly simpler case of recovering rigid motion has a long history and continues to attract much attention in computational vision. The literature is replete with analyses of the minimal sets of pointwise motion data sufficient to solve the structure-from-motion problem uniquely, subject to a rigidity assumption (see the survey [24]). However, a rigid-body assumption can be sensitive to noise, and it is clearly inappropriate when dealing with flexible objects.

In recent years, a desire to relax rigidity assumptions has begun to motivate new investigations into using motion information to invert the optical projection equations. Initial computational analyses treat restricted types of nonrigid motion—such as articulated bodies [25, 26]—while subsequent work recognizes the need for further generalization. Webb and Aggarwal [27] shed the global rigid-body assumption in favor of assuming local rigidity of planar patches. Under orthogonal projection, this implies a measurable affine distortion in the image plane. They are able to recover surface orientations from intensity and motion information. Chen [28] discusses the representation of nonrigid objects from the perspective of linear elasticity theory and generalizes the method of Webb and Aggarwal to handle perspective projection while presupposing a weaker assumption: that object surfaces deform isometrically (only bending occurs, so that distances along the surface are preserved). Koenderink and van Doorn [29] pursue a similar surface isometry assumption and they present an algorithm for recovering local aspects of shape from two views of a seven-point

polygonal model. Ullman [30] considers the problem of recovering the structure of arbitrarily deforming "objects" consisting of a set of points in space from the projected positions of these points at several successive time instants. His algorithm incrementally adjusts the 3D positions of points in a model to best account for the projected positions, while maximizing the model's rigidity across successive frames, as measured by the weighted sum of squared deviations of 3D distances between all pairs of points (global interaction).

Our approach goes beyond the above methods in significant ways, while retaining some of their best features. Our methods apply successfully to raw image sequences of extended natural objects undergoing nonrigid motion, not just to pointwise synthetic data. Our models accommodate arbitrary deformations, including bending, stretching, and shearing components. In principle, they also permit piecewise continuous deformations with explicit treatment of discontinuities (although the present paper does not pursue discontinuities). The underlying formulation is based on elasticity theory, and the elastic deformation forces are computable by local interaction in the discretized elastic body.

2.4. Multimodal integration and multiple views

The continuum mechanical equations of motion that govern our models provide a conceptually simple mechanism for integrating multiple visual modalities and information from multiple views. Each information source makes a contribution to the net force field acting on the model (cf. [31]).

Techniques have been developed for constructing volumetric descriptions of objects by intersecting projective cones associated with profiles in images taken from multiple viewpoints [32, 33]. Active sensors have been incorporated as well [34]. Although it makes use of profiles in multiple images, the volume intersection approach is fundamentally different from ours.

We demonstrate multimodal and multiview integration by combining stereo with motion in order to recover the evolving shape in proper depth robustly. Several researchers have investigated motion–stereo fusion as a means of facilitating the recovery of 3D velocity information from images [35–39]. Typically, they treat the case of observers moving through static environments.

Our approach is intended to handle flexible objects moving independently. It extracts 2D shape information directly from profiles and is able to estimate depth over each object's surface from profiles in stereo image pairs. Our current algorithms to do not involve explicit stereo or motion correspondence matching computations; each model simply tracks its associated object via the evolving force field to produce a 3-space velocity field over the entire surface of the object (including hidden surfaces) in proper depth.

3. The Symmetry-Seeking Model

Before we review the formulation of the symmetry-seeking model proposed in [3], here is an informal description. Consider a deformable sheet made of elastic material (a membrane–thin-plate hybrid). Roll this sheet to form a tube. Next, pass a deformable spine made of similar material down the length of the tube. At regularly spaced points along the spine, couple it to the tube with radially projecting forces so as to maintain the spine in approximate 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 the shape by introducing expansion/compression forces radiating from the spine. The rigidities of the spine and the tube are independently controllable, and their natural rest metrics and curvatures can either be 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 viscoelasticity and fracture) can be obtained by adjusting the control variables designed into the model.

We represent the spine and tube as geometric mappings from material coordinate domains into Euclidean 3-space $\mathbb{R}^3$. We express the mappings as vectors whose component functions denote time-varying components of position in space. The spine is a deformable space curve defined by mapping a univariate material coordinate domain $s \in [0, 1]$ into $\mathbb{R}^3$: $\boldsymbol{v}(s, t) = (X(s, t), Y(s, t), Z(s, t))$. The tube is made from a deformable space sheet defined by mapping a bivariate material coordinate domain $(x, y) \in [0, 1]^2$ into $\mathbb{R}^3$: $\boldsymbol{v}(x, y, t) = (X(x, y, t), Y(x, y, t), Z(x, y, t))$.

A functional $\mathscr{E}(\boldsymbol{v})$ characterizes the deformable material by associating a nonnegative strain energy with any admissible mapping. In our deformable models, $\mathscr{E}$ is an instance of the controlled-continuity spline functions defined in [40].

The continuum mechanical equation

$$\mu \, \frac{\partial^2 \boldsymbol{v}}{\partial t^2} + \gamma \, \frac{\partial \boldsymbol{v}}{\partial t} + \frac{\delta \mathscr{E}(\boldsymbol{v})}{\delta \boldsymbol{v}} = f(\boldsymbol{v}) \tag{1}$$

governs the nonrigid motion of a body in response to a net extrinsic force $f(\boldsymbol{v})$, where μ is the mass density function of the deformable body and γ is the viscosity function of the ambient medium [41]. The third term on the left-hand side of the equation is the variational derivative of the strain energy functional $\mathscr{E}$ [42]; it expresses the elastic force internal to the body.

The deformation energy associated with the spine mapping $\boldsymbol{v}(s, t)$ is given by

$$\mathcal{E}^{\mathrm{S}}(\boldsymbol{v}) = \int_0^1 w_1 \left| \frac{\partial \boldsymbol{v}}{\partial s} \right|^2 + w_2 \left| \frac{\partial^2 \boldsymbol{v}}{\partial s^2} \right|^2 \, \mathrm{d}s \,, \tag{2}$$

where vertical bars denote Euclidean vector norms. Here, $w_1(s, t)$ determines the local tension along the spine, while $w_2(s, t)$ determines its local rigidity.

The deformation energy associated with the sheet mapping $\boldsymbol{v}(x, y, t)$ is given by the functional

$$\mathcal{E}^{\mathrm{T}}(\boldsymbol{v}) = \int_0^1 \int_0^1 w_{10} \left| \frac{\partial \boldsymbol{v}}{\partial x} \right|^2 + w_{01} \left| \frac{\partial \boldsymbol{v}}{\partial y} \right|^2$$

$$+ w_{20} \left| \frac{\partial^2 \boldsymbol{v}}{\partial x^2} \right|^2 + 2 w_{11} \left| \frac{\partial^2 \boldsymbol{v}}{\partial x \partial y} \right|^2 + w_{02} \left| \frac{\partial^2 \boldsymbol{v}}{\partial y^2} \right|^2 \mathrm{d}x \, \mathrm{d}y \,. \tag{3}$$

The weighting functions $w_{10}(x, y, t)$ and $w_{01}(x, y, t)$ locally control the tension of the sheet along each material coordinate curve, while $w_{20}(x, y, t)$, $w_{11}(x, y, t)$, and $w_{02}(x, y, t)$ locally control its rigidities.

Aside from boundary terms, the variational derivatives of $\mathcal{E}^{\mathrm{S}}$ and $\mathcal{E}^{\mathrm{T}}$ expressing the elastic forces are given as follows:

$$\frac{\delta \mathcal{E}^{\mathrm{S}}}{\delta \boldsymbol{v}} = \frac{\partial^2}{\partial s^2} \left(w_2 \frac{\partial^2 \boldsymbol{v}}{\partial s^2} \right) - \frac{\partial}{\partial s} \left(w_1 \frac{\partial \boldsymbol{v}}{\partial s} \right), \tag{4}$$

$$\frac{\delta \mathcal{E}^{\mathrm{T}}}{\delta \boldsymbol{v}} = \frac{\partial^2}{\partial x^2} \left(w_{20} \frac{\partial^2 \boldsymbol{v}}{\partial x^2} \right) + 2 \frac{\partial^2}{\partial x \partial y} \left(w_{11} \frac{\partial^2 \boldsymbol{v}}{\partial x \partial y} \right) + \frac{\partial^2}{\partial y^2} \left(w_{02} \frac{\partial^2 \boldsymbol{v}}{\partial y^2} \right)$$

$$- \frac{\partial}{\partial x} \left(w_{10} \frac{\partial \boldsymbol{v}}{\partial x} \right) - \frac{\partial}{\partial y} \left(w_{01} \frac{\partial \boldsymbol{v}}{\partial y} \right). \tag{5}$$

The weighting functions in the above expressions control the elastic properties. When necessary, we can regulate the natural shapes of the deformable bodies—i.e., their equilibrium shapes exclusive of external forces—through suitably defined control functions. For instance we can encourage the spine to maintain a nonzero natural arc length $L_1(s)$, by defining

$$w_1 = K_1 \left(\left| \frac{\partial \boldsymbol{v}}{\partial s} \right| - L_1 \right),$$

where K_1 is a tension factor. Similarly we can define

$$w_{10} = K_{10} \left(\left| \frac{\partial \boldsymbol{v}}{\partial x} \right| - L_{10} \right), \qquad w_{01} = K_{01} \left(\left| \frac{\partial \boldsymbol{v}}{\partial y} \right| - L_{01} \right)$$

to encourage natural lengths $L_{10}(x, y)$ and $L_{01}(x, y)$ for the sheet along the x

and y material coordinate directions respectively. Analogous expressions for w_2 will encourage a natural nonzero curvature for the spine, while for w_{20}, w_{11}, and w_{02} they will encourage natural curvatures for the sheet.

To allow discontinuities to occur in the spine at any material point s_0, we set $w_1(s_0, t) = w_2(s_0, t) = 0$ which permits a position discontinuity, or $w_2(s_0, t) = 0$ which permits a tangent discontinuity. The obvious extensions hold with regard to the five control functions associated with the sheet. See [40] for further details regarding discontinuities.

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

$$v(0, y, t) = v(1, y, t) , \qquad \left.\frac{\partial v}{\partial x}\right|_{(0,y,t)} = \left.\frac{\partial v}{\partial x}\right|_{(1,y,t)} . \qquad (6)$$

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 Fig. 7). We then distinguish the mapping function of the spine $v^S(s, t)$ from that of the tube $v^T(x, s, t)$ with superscripts S and T.
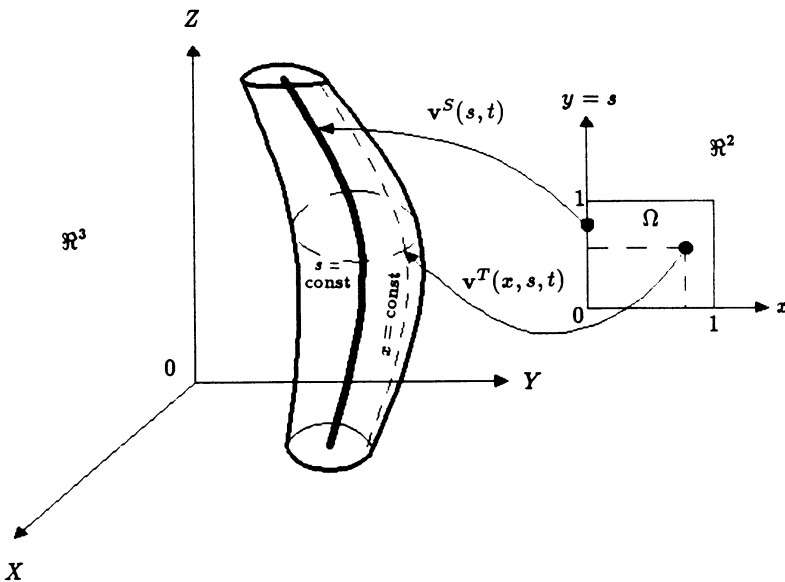

Fig. 7. Geometric representation of the symmetry-seeking model. For fixed time t, the spine and tube components of the model map material coordinates $y = s$ and (x, s), respectively, into positions in $\mathbb{R}^3$.

We define

$$\bar{\boldsymbol{v}}^{\mathrm{T}}(s) = \frac{1}{l} \int_0^1 \boldsymbol{v}^{\mathrm{T}} \left| \frac{\partial \boldsymbol{v}^{\mathrm{T}}}{\partial x} \right| \mathrm{d}x$$

to be the centroid of the coordinate curve ($s = $ constant) circling the tube, whose length is given by

$$l(s) = \int_0^1 \left| \frac{\partial \boldsymbol{v}^{\mathrm{T}}}{\partial x} \right| \mathrm{d}x \ .$$

We also define the tube's radial vector function with respect to the spine as $r(x, s) = \boldsymbol{v}^{\mathrm{T}} - \boldsymbol{v}^{\mathrm{S}}$, the unit radial vector function $\hat{r}(x, s) = r/|r|$, and

$$\bar{r}(s) = \frac{1}{l} \int_0^1 |r| \left| \frac{\partial \boldsymbol{v}^{\mathrm{T}}}{\partial x} \right| \mathrm{d}x$$

as 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:

$$\begin{aligned} f_a^{\mathrm{S}}(s, t) &= a(\bar{\boldsymbol{v}}^{\mathrm{T}} - \boldsymbol{v}^{\mathrm{S}}) \ , \\ f_a^{\mathrm{T}}(x, s, t) &= -(a/l)(\bar{\boldsymbol{v}}^{\mathrm{T}} - \boldsymbol{v}^{\mathrm{S}}) \ , \end{aligned} \tag{7}$$

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

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

$$f_b^{\mathrm{T}}(x, s, t) = b(\bar{r} - |r|)\hat{r} \tag{8}$$

on the tube, 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

$$f_c^{\mathrm{T}}(x, s, t) = c\hat{r} \ , \tag{9}$$

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 negative values to $c(0)$ and $c(1)$, thereby creating a sausage-like surface. This force is also useful for counteracting a tendency of the tube to contract around the spine due to the quadratic

curvature approximations appearing in $\mathscr{E}$. Use of the true curvature tensor in $\mathscr{E}$ will alleviate the contraction problem directly, but at the expense of non-linearities that are computationally expensive and badly behaved numerically.

Summing the above coupling forces into the equations of motion (1) associated with the spine and tube, we obtain the following dynamic system describing the motion of the symmetry-seeking model:

$$\mu \frac{\partial^2 \boldsymbol{v}^S}{\partial t^2} + \gamma \frac{\partial \boldsymbol{v}^S}{\partial t} + \frac{\delta \mathscr{E}^S}{\delta \boldsymbol{v}^S} = \frac{\delta P^S}{\delta \boldsymbol{v}^S} + \boldsymbol{f}_a^S \,,$$

$$\mu \frac{\partial^2 \boldsymbol{v}^T}{\partial t^2} + \gamma \frac{\partial \boldsymbol{v}^T}{\partial t} + \frac{\delta \mathscr{E}^T}{\delta \boldsymbol{v}^T} = \frac{\delta P^T}{\delta \boldsymbol{v}^T} + \boldsymbol{f}_a^T + \boldsymbol{f}_b^T + \boldsymbol{f}_c^T \,.$$

$$(10)$$

The elasticity terms on the left-hand side of these coupled equations are given by (4) and (5) (with $y \equiv s$ in the latter).

$P^S(\boldsymbol{v}^S)$ and $P^T(\boldsymbol{v}^T)$ are generalized potential functions associated with the spine and tube respectively. Their variational derivatives express the extrinsic forces that act on these bodies.

4. Extrinsic Force Fields

This section explains how we transform one or more input images into generalized potential functions suitable for reconstruction. The resulting force field brings instances of the symmetry-seeking model into maximal consistency with the images, and it maintains the consistency over time in the dynamic case.

More specifically, profiles in the images exert an attraction over the model such that the deformable tube, as projected into the image planes through a suitable camera model, accounts as much as possible for the observed profiles. In this paper, the spine of the model experiences no image-based forces ($P^S = 0$), although we project it into the image plane along with the tube for display purposes.

To simplify the potential functions, we consider objects with subdued texture which are imaged in front of a contrasting background. Hence, we can expect that the stronger image intensity gradients are associated with object profiles. The force fields that we propose yield interesting results, their simplicity notwithstanding.

4.1. Monocular potential

We first define a generalized potential function P^T to couple the tube to a static, monocular image. Given the image of an object, the ambient force field resulting from this potential deforms the symmetry-seeking model to make its shape consistent with the object's profile in the image. The occluding boundary

of the deformable tube with respect to the viewpoint is made sensitive to this force field which attracts it towards significant intensity gradients in the image.

Let $I(\xi, \eta)$ be the image intensity function. We define the potential

$$P^{\mathrm{T}}(\boldsymbol{v}^{\mathrm{T}}) = \beta |\nabla(G_\sigma * I(\Pi[\boldsymbol{v}^{\mathrm{T}}]))| , \tag{11}$$

which imparts on the tube boundary an affinity for large image intensity changes. The operations symbolized by the expression are as follows: The operator Π denotes an imaging projection. $\Pi[\boldsymbol{v}^{\mathrm{T}}]$ expresses the projection of the tube $\boldsymbol{v}^{\mathrm{T}}(x, s) = (X^{\mathrm{T}}(x, s), Y^{\mathrm{T}}(x, s), Z^{\mathrm{T}}(x, s))$ from 3-space into the image plane (ξ, η). The expression $G_\sigma * I$ denotes the image convolved with a (Gaussian) smoothing filter whose characteristic width is σ. The operator ∇ denotes the gradient of the smoothed image, and the vertical bars signify the magnitude of this gradient.

The weighting function $\beta(x, s)$ is nonzero only for material points (x, s) near occluding boundaries of the tube which are not obscured behind other objects. We activate points near occluding boundaries by defining

$$\beta(x, s) = \begin{cases} 1 , & \text{if } |\boldsymbol{i} \cdot \boldsymbol{n}| < \tau , \\ 0 , & \text{otherwise} , \end{cases} \tag{12}$$

where τ is a small threshold (nominally 0.05),

$$\boldsymbol{n}(x, s) = \left(\frac{\partial \boldsymbol{v}^{\mathrm{T}}}{\partial x} \times \frac{\partial \boldsymbol{v}^{\mathrm{T}}}{\partial s} \right) \bigg/ \left| \frac{\partial \boldsymbol{v}^{\mathrm{T}}}{\partial x} \times \frac{\partial \boldsymbol{v}^{\mathrm{T}}}{\partial s} \right|$$

is the unit normal over the surface of the tube, and $\boldsymbol{i}(x, s)$ is a unit vector from the imaging focal point to any point (x, s) on the tube.

While a perspective projection is generally appropriate for Π, we have obtained satisfactory results in the examples considered below using orthographic projection $\Pi : (\xi, \eta) = (X, Y)$. Everywhere over the tube in this case, $\boldsymbol{i}$ is the unit normal to the image plane, and the resulting force field acts on the tube in a direction parallel to the image plane.

Figure 8 illustrates the effect of the image operations for progressively larger values of σ (we used similar operations in [8]). The ambient force field adjusts the 3D shape of the model's occluding boundary in space so as to maximize the magnitude of the image gradient (darkness) along the projected model profile in the image. By using potential functions at several scales, we can trade off localization accuracy against long-range attraction. As is evident from the figure, broad wells surround the local minima of the image potential at the coarser scales. Such wells attract the model from a considerable distance, but the associated minima are blurred and localize the profiles in the image data rather poorly. Continuous scale space [43] provides a good medium for obtaining both long-range attraction and good localization. We can apply a
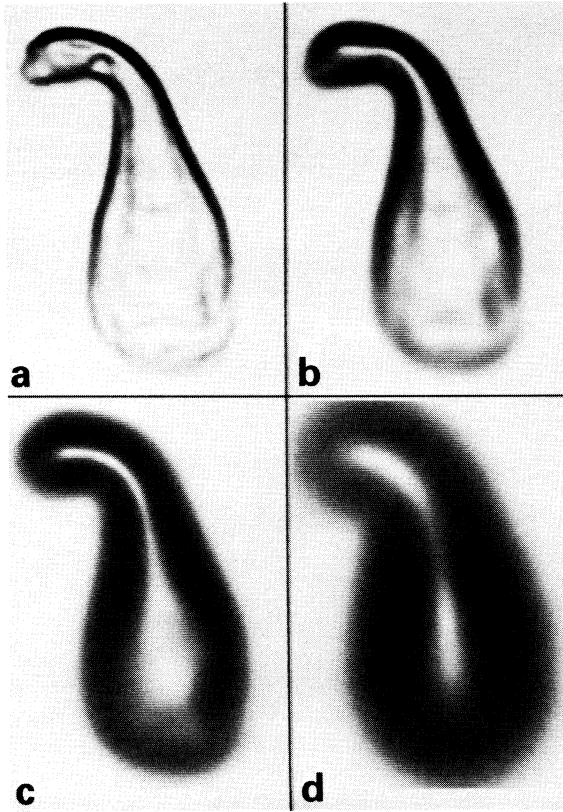
Fig. 8. Creating an image potential function. Result of applying image operations (see text) to the squash image of Fig. 3. (a)–(d) σ increases progressively. Darkness indicates magnitude of local gradient of the Gaussian blurred image. Each processed image has been rescaled to span the available intensity range.

continuation method in the image potential scale space, parameterized by σ, which allows the model to equilibrate at a coarse scale, then continuously reduces the smoothing σ to track an equilibrium trajectory from coarse to fine scale [7].

4.2. Binocular potential

The next refinement is to match the deformable model to an object's profiles in a binocular image pair. The symmetry-seeking model is free to undergo motion in 3-space while deforming such that its stereoscopic projection through the binocular camera model best accounts for the observed profiles in both images.

Let $I_L(\xi_L, \eta_L)$ represent the left and $I_R(\xi_R, \eta_R)$ the right image intensity function. The binocular potential function P^T is a straightforward extension of

the monocular case. We define

$$P^{\mathrm{T}}(\boldsymbol{v}^{\mathrm{T}}) = \beta_{\mathrm{L}}|\nabla(G_\sigma * I_{\mathrm{L}}(\Pi_{\mathrm{L}}[\boldsymbol{v}^{\mathrm{T}}]))| + \beta_{\mathrm{R}}|\nabla(G_\sigma * I_{\mathrm{R}}(\Pi_{\mathrm{R}}[\boldsymbol{v}^{\mathrm{T}}]))| , \qquad (13)$$

which imparts on the tube boundary an affinity for large intensity changes in both images. The expressions $\Pi_{\mathrm{L}}[\boldsymbol{v}^{\mathrm{T}}]$ and $\Pi_{\mathrm{R}}[\boldsymbol{v}^{\mathrm{T}}]$ denote the two-component stereoprojection of the tube $\boldsymbol{v}^{\mathrm{T}} = (X^{\mathrm{T}}(x, s), Y^{\mathrm{T}}(x, s), Z^{\mathrm{T}}(x, s))$ into the image planes $(\xi_{\mathrm{L}}, \eta_{\mathrm{L}})$ and $(\xi_{\mathrm{R}}, \eta_{\mathrm{R}})$ respectively. The image-processing operations are the same as for the monocular case.

We activate points near occluding boundaries with respect to the left image by defining

$$\beta_{\mathrm{L}}(x, s) = \begin{cases} 1, & \text{if } |i_{\mathrm{L}} \cdot \boldsymbol{n}| < \tau , \\ 0, & \text{otherwise} , \end{cases} \qquad (14)$$

where τ is a small threshold, $\boldsymbol{n}(x, s)$ is the unit normal over the surface of the tube, and $i_{\mathrm{L}}(x, s)$ is a unit vector from the focal point of the left image to any point (x, s) on the tube. We define the weighting function $\beta_{\mathrm{R}}(x, s)$ for the right image similarly.

Although it is possible to use a general binocular camera model (see, e.g., [44, Section 10.6]), its parameters need not be known with great accuracy for our approach to work. Consequently, we have found it convenient to employ the following simplified perspective stereoprojection with eye vergence at infinity:

$$\begin{aligned} \Pi_{\mathrm{L}} &: (\xi_{\mathrm{L}}, \eta_{\mathrm{L}}) = (X^{\mathrm{T}} + \alpha Z^{\mathrm{T}}, Y^{\mathrm{T}}) , \\ \Pi_{\mathrm{R}} &: (\xi_{\mathrm{R}}, \eta_{\mathrm{R}}) = (X^{\mathrm{T}} - \alpha Z^{\mathrm{T}}, Y^{\mathrm{T}}) , \end{aligned} \qquad (15)$$

where α is a constant. In this case, i_{L} and i_{R} become unit normals to the left and right image planes respectively.

4.3. Dynamic potential

When the images are time-varying the ambient force field becomes dynamic. It carries the model through motions, continually molding its shape to maintain maximal consistency with the evolving image data. Our method for allowing the symmetry-seeking model to track an object undergoing nonrigid motion is as follows. The first frame of the image sequence is presented to the model as if it were a static scene. The model achieves the best possible reconstruction using this initial data. The projected boundary points equilibrate at a fixed point in the ambient force field and the model locks on to the consistent state.

Whereas the equilibrium persists indefinitely in the static case, in the dynamic case we immediately present the model with the next frame of the

image sequence. Now the ambient force field is perturbed because of the motions of objects in the scene. The model actively seeks a new consistent state by moving towards the nearest fixed point. If the motion of the object is sufficiently slow and continuous, the model will track the dynamic equilibrium point, thus updating its state in accordance with the new image information available to it. By repeating this procedure with each successive incoming frame, the symmetry-seeking model integrates the incoming information over time.

One possible variant to our standard nonrigid tracking procedure is a two-step rigid-deformable technique. With each incoming frame, we first restrict the model to rigid motion in order to obtain a suboptimal result quickly. Only then do we allow it to deform in order to account for the residual nonrigid motion. This method accelerates the numerical solution when objects under analysis move nearly rigidly.

5. Numerical Solution

The continuous differential equations of motion (10) for the symmetry-seeking model pose a nonlinear initial boundary value problem. To obtain a numerical solution, we first perform a semi-discretization in the material coordinates of the model. We then integrate the resulting coupled system of second-order ordinary differential equations through time using standard techniques. In the static field case, we integrate until the viscous damping (γ-term) dissipates all kinetic energy, thus bringing the model to static equilibrium. In the dynamic field case, we integrate through time, computing the motion of the model in reaction to the evolving force field. Numerically, this amounts to solving a sequence of dynamic equilibrium problems, each solution providing initial conditions to the subsequent problem.

Semi-discretization in material coordinates is carried out using standard finite difference approximations on regular grids of nodes [45]. We approximate the three components of the elastic force $\delta\mathscr{E}^S/\delta v^S$ by discretizing the partial derivatives in (4) on a linear N_s point grid. We use a rectangular $N_x \times N_s$ grid to similarly discretize the partial derivatives in the components of $\delta\mathscr{E}^T/\delta v^T$ in (5). The extrinsic force components $\delta P^T(v^T)/\delta v^T$ are computed numerically in the image domains (ξ, η) using bilinear interpolation between centrally differenced pixel values.

Finally, we approximate the time derivatives in (10) at discrete times using finite differences over regular intervals. We have successfully employed both an explicit Euler time integration scheme and an analogue to successive overrelaxation (SOR) to "solve" the dynamic equilibrium problems in sequence, each a large sparse system of algebraic equations. These iterative solution methods require only local operations, and the former is a parallel scheme. Explicit solvers are relatively inexpensive per time step (linear in the number of nodes),

but they quickly become unstable as the intrinsic or extrinsic forces are increased, thus necessitating tiny time steps. As expected, we observe stability over larger time steps when using an implicit Euler time integration scheme in conjunction with matrix factorization methods [45]. The drawback of these robust, direct solution methods is that they become expensive for finely discretized models (cubic in the number of nodes).

A good compromise is to use a hybrid technique that combines some of the benefits of direct and iterative schemes. We have had success with an operator splitting approach used in alternating direction implicit (ADI) methods [45, 46]. This efficient procedure exploits the fact that we are dealing with a rectangular grid of nodes. Each time step of the procedure involves a sweep in the x direction solving N_x independent systems of algebraic equations in N_s unknowns, followed by a sweep in the s direction solving N_s independent systems in N_x unknowns. Our strategic linearizations of the elastic forces permit us to apply the method independently to each of the three tube position components (X^T, Y^T, Z^T). The spine gives rise to an additional system of equations in N_s unknowns for each of its position components (X^S, Y^S, Z^S).

As a consequence of the controlled-continuity deformation model, each of the unidimensional systems of equations has a pentadiagonal matrix of coefficients, and it is efficiently solvable (linear time in the number of unknowns) using direct solution methods. We employ a normalized Cholesky decomposition step followed by a forward–reverse resolution step. See [47] for a derivation of the pentadiagonal matrix and for a discussion of the direct method and [8] for its application to image contour models.

Resolution, a relatively inexpensive step, must be performed at every iteration as the applied forces change. Matrix decomposition is somewhat more expensive, but it is required only when the material properties of the model are altered (e.g., to increase rigidity or to introduce discontinuities). Currently, we perform only an initial decomposition because we have not yet experimented with the variation of material properties during solution.

We find that for larger grid sizes and increasingly rigid material the alternating direction approach evolves solutions faster than the SOR type method that we employed in [3]. This is attributable to the fact that the direct solution of each unidimensional system obtained through operator splitting "immediately" distributes to all nodes along two perpendicular parametric grid lines the effects of forces acting on their common node.

6. Results

Figure 9 shows selected frames from an animation sequence showing the reconstruction of a single symmetry-seeking model from a monocular image. Given an image of the squash (Fig. 3), the user initializes the spine's projection in the image plane somewhere near the medial axis of the squash. The tube
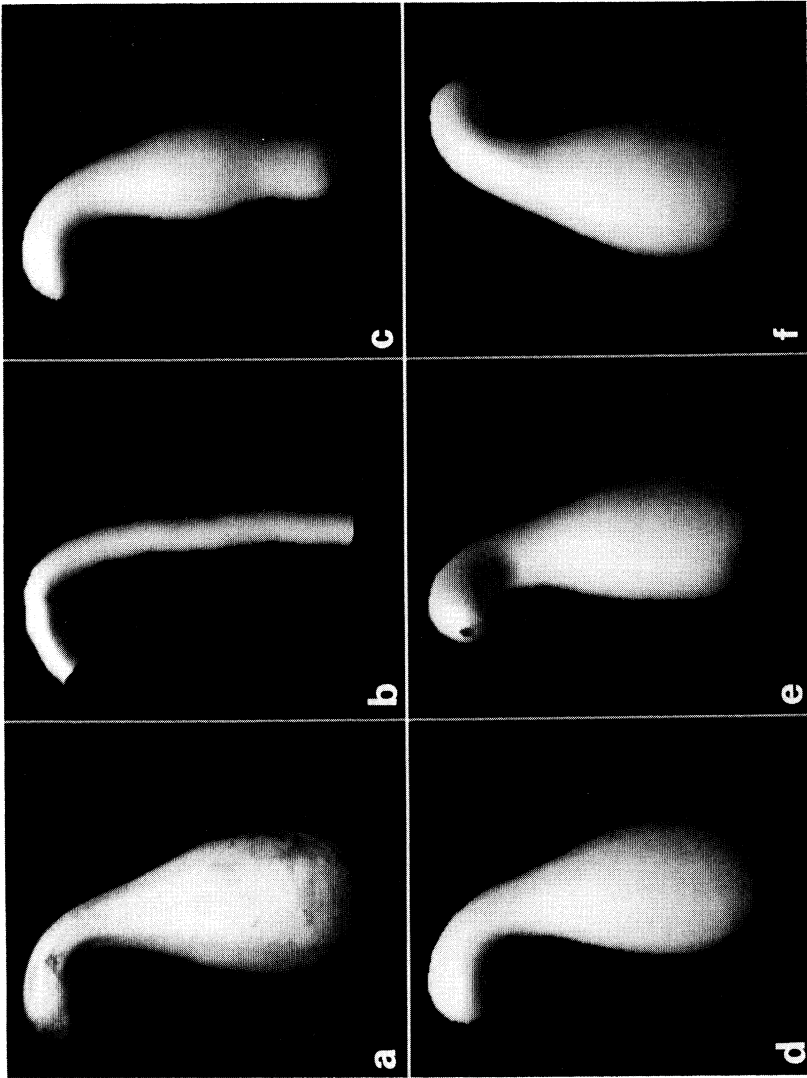
Fig. 9. Reconstruction of a squash. (a) Squash image. Selected frames from an animation sequence are shown: (b) Initial state of the 3D model. (c) Intermediate shape during reconstruction. (d) Final reconstructed model. (e)–(f) Model rotating rigidly in space. The 3D model is rendered as a shaded shell.

begins as a generalized cylinder with uniformly circular cross-section. It inflates due to the action of the internal expansion forces. As the tube approaches a profile, the force field due to the significant intensity edge attracts it strongly. It equilibrates over the profile to reconstruct the nonconvex, quasi-symmetric 3D shape of the squash, as shown from different angles in the figure. Notice how the model begins as a generalized cylinder yet recovers the inexact axisymmetry of the squash in the course of reconstruction. By virtue of the symmetry-seeking constraints intrinsic to the model, hidden surface portions are smoothly extrapolated from portions in view. Essentially the same reconstruction procedure applies to each of 43 symmetry-seeking model instances in the reconstruction of Picasso's "Rites" in Fig. 2.

The next example involves the reconstruction of two quasi-symmetric objects, a pear and a potato, from a single image. Figure 10(a) shows the grey-level image of the still life scene. Notice that the potato partially occludes the pear in the image. Figure 10(b) shows the initial model configurations, manually specified by the user. Figures 10(c) and (d), show the reconstructed 3D models from two points of view. To handle the partial occlusions (incomplete boundaries), we nullify the weighting function $\beta(x, s)$ over portions of an occluded model's surface which are obscured from the viewpoint by another model. Appropriately, the obscured parts of the model feel no image forces, because they are invisible in the image. We use a standard 3D ray-casting technique in conjunction with the projection operation Π in (11) in order to test surface patches for visibility using a depth buffer. The ray-casting operation requires knowledge of the relative depth ordering of the objects. In the monocular case, this information is not available directly, so the user presently specifies the relative depth ordering (although it may be possible to obtain local depth ordering information automatically, through analysis of occlusion cues in the image [48]).

The most direct way of obtaining true depth information is through stereo. The reconstruction method was applied to a stereo–motion sequence consisting of 40 video fields portraying the 3D motion of a human finger. The imaging apparatus was a beam-splitting stereo adaptor mounted on a CCD camera. The user specifies an initial spine on the first stereo pair shown in Fig. 11(a). The initial tube is a cylinder around the spine (Fig. 11(b)). The model's differential equations (using $\mu = 0$) are solved on a $N_x \times N_s = 25 \times 25$ grid for the initial frame (requiring about 40 alternating direction iterations), thus reconstructing the shape of the object in proper depth. Figures 11(c)–(e) show the reconstructed shape rendered from several viewpoints. Using this equilibrium shape as initial condition, the equations of motion are then integrated through time over the remaining frames of the stereo sequence (using 20 alternating direction iterations per frame). This produces a dynamic 3D reconstruction of the finger's shape and motion. Figure 12 shows six representative frames of the sequence along with the corresponding reconstructed shapes in motion.
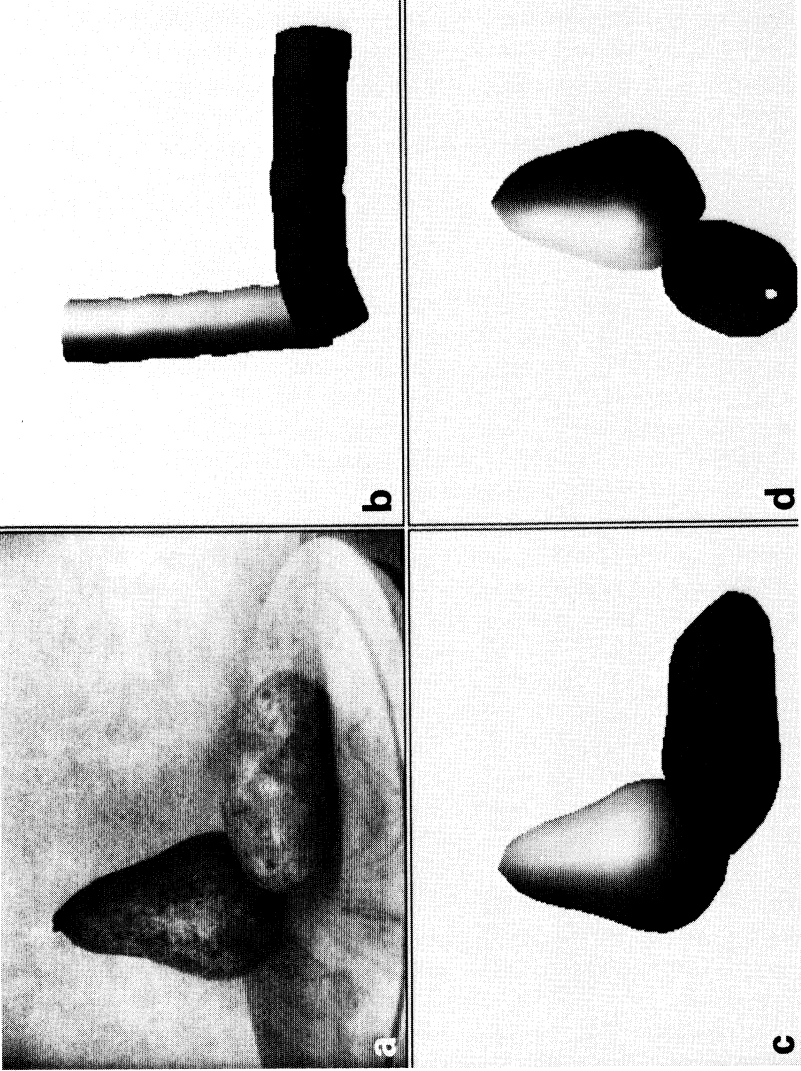
Fig. 10. Reconstruction of a still life scene with symmetry-seeking deformable models. (a) Image of the scene. (b) Initial user-specified configurations of the 3D models. Two rendered views of the reconstructed still life: (c) Frontal view of the 3D models; (d) side view.
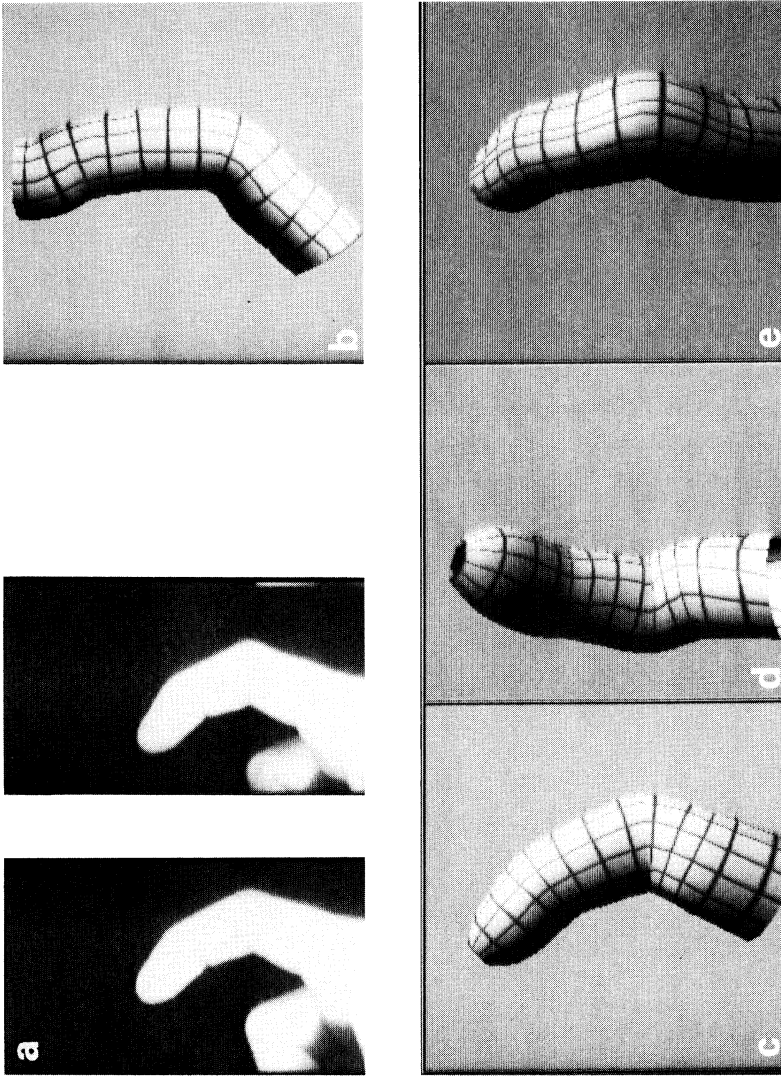
Fig. 11. Initial 3D reconstruction of a finger. (a) Finger stereo pair for first time instant. (b) User-initialized cylinder. (c)–(e) Initial reconstructed shape from three viewpoints. Every other grid line has been drawn on the surface.
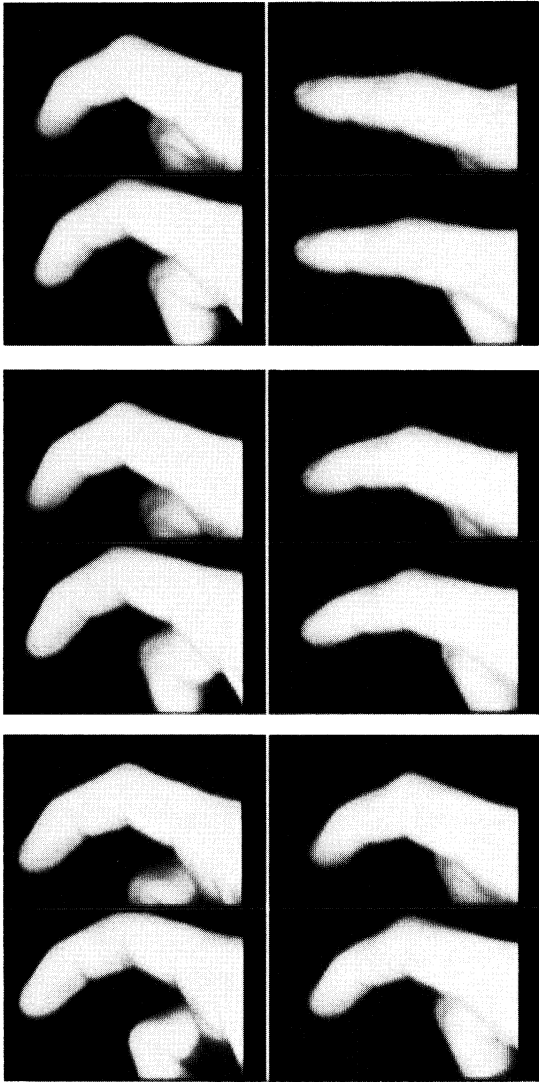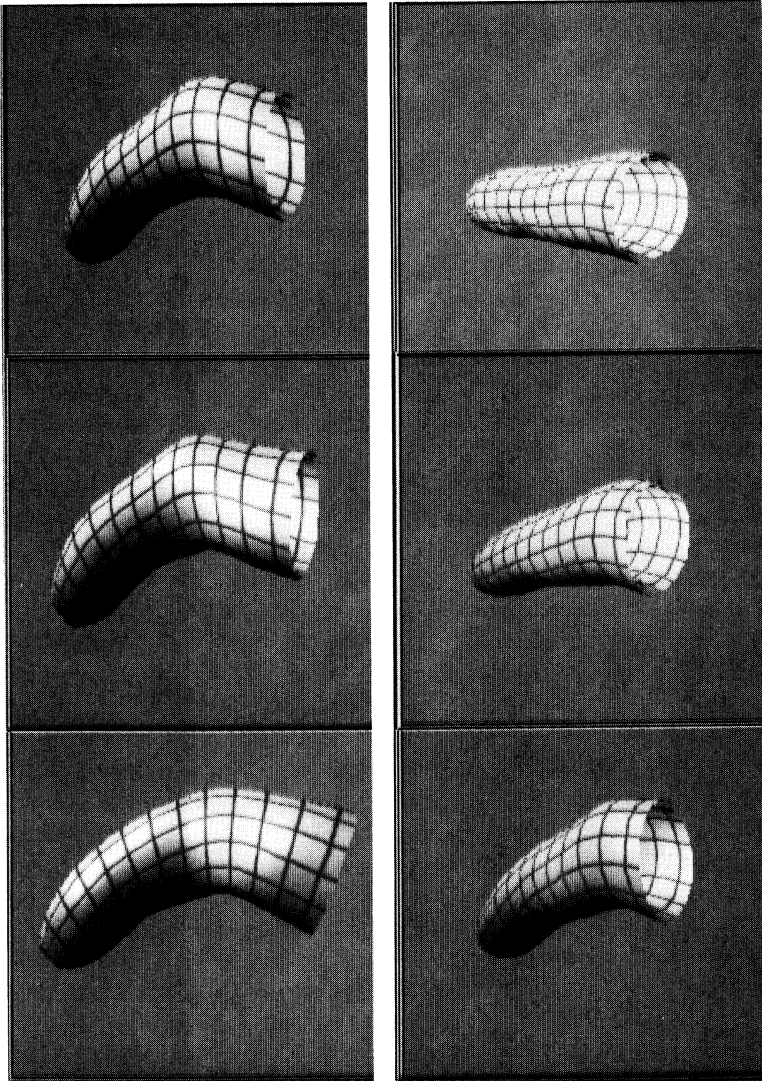
Fig. 12(a). Six frames of the stereo sequence.

Fig. 12(b). Evolving shape and motion of the model.

7. Discussion

Referring back to Fig. 5, we can compare 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 is generated by a stereo algorithm that we developed in [7]. The algorithm generates depth discontinuities along the occluding boundaries of the objects. Profiles of smooth objects and the occluding boundaries which generate them are known to present difficulties to conventional stereo matching techniques. This is mainly because profiles observed in the left and right image map to different occluding boundaries on smooth objects. Our method for reconstructing models directly from images overcomes this problem. The use of a full 3D model in tandem with separate left and right projection operators simplifies the association of each image profile with the suitable occluding contour on the model.

In our experience, the user need not initialize the spine (or tube) of the model with any great accuracy. Typically, the user draws rough medial axes running more or less the length of the objects of interest in the image. Figure 13(a) shows a typical set of initial spines. With the tubes starting out as cylinders extruded along these spines, the reconstruction process makes the necessary adjustments to both the spine and tube (Fig. 13(c)). Figure 13(b) shows the projections of the final spines. Compared to the initial axes, the final axes are improved; they are smoother and have suitable lengths. It appears possible, therefore, to instantiate reasonable spines automatically using medial axis transform [49] or smoothed local symmetry [50] algorithms, and perhaps our reconstruction method, where applicable, could improve the axes generated by such algorithms.

In the restricted case of computing 3D models of objects from a single monocular image, we tacitly assume that a suitable viewpoint has been chosen such that all significant object features are visible, and that the axis of the object is not severely inclined away from the image plane. However, in the more general case where we make use of stereo information, we experience little difficulty in tracking the shape and motion of an object even if its axis tilts away from the image plane significantly.

A shortcoming of our current algorithms stems from the fact that profile information alone, even moving stereo profiles, provide incomplete information about objects. When image data is not utilized over substantial portions of the object's surface, the symmetry-seeking model may yield a reconstructed shape more symmetric than the actual one. Also, it is difficult to detect rotations around the object's axis exclusively from moving profile information. Region-based measures over the surface would be helpful in this regard.

Despite the limitations of our current implementation, a crucial advantage of our approach is the ease of integrating additional constraints into the solution. For instance, we can straightforwardly generalize the binocular potential

Fig. 13. Spines for the symmetry-seeking models. (a) Projections of the initial spines drawn by the user. (b) Projections of the final spines.

Fig. 13(c). Full resolution wire-frame grids of the reconstructed models.

function (13) as a sum over any number of views taken from known viewpoints around the object (cf. [33]). A focus of our current work is the formulation and implementation of extrinsic constraints that exploit shading and texture information over the entire visible surface, as well as stereo–motion constraints based on local area correlation which promise to effectively supplement our current edge-based information. By applying more sophisticated image-processing methods, we expect to obtain extrinsic forces that can deal with textured objects and more general imaging conditions. We also look forward to incorporating more sophisticated analytic camera models into the generalized potential functionals and to automatically solve for the camera parameters as an integral part of the reconstruction procedure.

We are investigating the use of scale space continuation methods [7] to partially automate the initialization of models. For the time being, however, our algorithms remain interactive, and it is up to the user to supply reasonable initial conditions. Nonetheless, our approach suggests force constraint mechanisms for automatically bringing higher-level knowledge to bear on the reconstruction process. The development of such mechanisms is an interesting topic for future research.

REFERENCES

1. Binford, T.O., Visual perception by computer, Invited Talk, *IEEE Systems and Control Conference*, Miami, FL (1971).
2. Marr, D., Visual information processing: The structure and creation of visual representations, *Proc. Roy. Soc. London B* **290** (1980) 199–218.
3. Terzopoulos, D., Witkin, A. and Kass, M., Symmetry-seeking models and 3D object reconstruction, *Int. J. Comput. Vision* **1** (1987) 211–221.
4. Terzopoulos, D., Platt, J., Barr, A. and Fleischer, K., Elastically deformable models, *Comput. Graph.* **21** (4) (1987) (*Proceedings SIGGRAPH-87*) 205–214.
5. Witkin, A., Fleischer, K. and Barr, A., Energy constraints on parameterized models, *Comput. Graph.* **21** (4) (1987) (*Proceedings SIGGRAPH-87*) 225–232.
6. Barzel, R. and Barr, A., Modeling with dynamic constraints, in: *Topics in Physically-Based Modeling*, *ACM SIGGRAPH '87 Course Notes* **17**, Anaheim, CA (1987).
7. Witkin, A., Terzopoulos, D. and Kass, M., Signal matching through scale space, *Int. J. Comput. Vision* **1** (1987) 133–144.
8. Kass, M., Witkin, A. and Terzopoulos, D., Snakes: Active contour models, *Int. J. Comput. Vision* **1** (1987) 321–331.
9. Platt, J., An elastic model for interpreting 3D structure from motion of a curve, Unpublished Manuscript (1987).
10. Terzopoulos, D., Multilevel computational processes for visual surface reconstruction, *Comput. Vision Graph. Image Process.* **24** (1983) 52–96.
11. Barr, A., Superquadrics and angle-preserving transformations, *IEEE Comput. Graph. Appl.* **18** (1981) 21–30.
12. Pentland, A.P., Parts: Structured descriptions of shape, in: *Proceedings AAAI-86*, Philadelphia, PA (1986) 695–701.
13. Bajcsy, R. and Solina, F., Three dimensional object representation revisited, in: *Proceedings First International Conference on Computer Vision*, London, England (1987) 231–240.
14. Barr, A.H., Global and local deformations of solid primitives, *Comput. Graph.* **18** (3) (1984) (*Proceedings SIGGRAPH-84*) 21–29.
15. Pentland, A.P., Recognition by parts, in: *Proceedings First International Conference on Computer Vision*, London, England (1987) 612–620.
16. Agin, G.A. and Binford, T.O., Computer description of curved objects, *IEEE Trans. Comput.* **25** (1976) 439–449.
17. Hollerbach, J.M., Hierarchical shape description of objects by selection and modification of prototypes, AI-TR-346, MIT AI Lab., Cambridge, MA (1975).
18. Nevatia, R. and Binford, T.O., Description and recognition of curved objects, *Artificial Intelligence* **8** (1977) 77–98.
19. Marr, D., Analysis of occluding contour, *Proc. Roy. Soc. London B* **197** (1977) 441–475.
20. Marr, D. and Nishihara, H.K., Representation and recognition of the spatial organization of three-dimensional shapes, *Proc. Roy. Soc. London B* **200** (1978) 269–294.
21. Brooks, R.A., Symbolic reasoning among 3-D models and 2-D images, *Artificial Intelligence* **17** (1981) 285–348.
22. Shafer, S.A., *Shadows and Silhouettes in Computer Vision* (Kluwer Academic, Boston, MA, 1985).
23. Fischler, M.A. and Elschlager, R.A., The representation and matching of pictorial structures, *IEEE Trans. Comput.* **22** (1973) 67–92.
24. Ullman, S., Recent computational studies in the interpretation of structure from motion, in: J. Beck, B. Hope and A. Rosenfeld (Eds.), *Human and Machine Vision* (Academic Press, New York, 1983) 459–480.
25. Webb, J.A. and Aggarwal, J.K., Visually interpreting the motions of objects in space, *Computer* **14** (1981) 40–46.
26. Hoffman, D.D. and Flinchbaugh, B.E., The interpretation of biological motion, *Biol. Cybern.* **42** (1982) 195–204.

27. Webb, J.A. and Aggarwal, J.K., Shape and correspondence, *Comput. Vision Graph. Image Process.* **21** (1983) 145–160.
28. Chen, S., Structure-from-motion without the rigidity assumption, in: *Proceedings IEEE 3rd Workshop on Computer Vision: Representation and Control*, Bellaire, MI (1985) 105–112.
29. Koenderink, J.J. and van Doorn, A.J., Depth and shape from differential perspective in the presence of bending deformations, *J. Opt. Soc. Am. A* **3** (1986) 242–249.
30. Ullman, S., Maximizing rigidity: The incremental recovery of 3-D structure from rigid and nonrigid motion, *Perception* **13** (1984) 255–274.
31. Terzopoulos, D., Integrating visual information from multiple sources, in: A.P. Pentland (Ed.), *From Pixels to Predicates: Recent Advances in Computational and Robotic Vision* (Ablex, Norwood, NJ, 1986) 111–142.
32. Baumgart, B.G., Geometric modeling for computer vision, AI Memo 249, Stanford Artificial Intelligence Laboratory, Stanford, CA (1974).
33. Martin, W.N. and Aggarwal, J.K., Volumetric description of objects from multiple views, *IEEE Trans. Pattern Anal. Mach. Intell.* **5** (1983) 150–158.
34. Wang, Y.F. and Aggarwal, J.K., On modelling 3-D objects using multiple sensory data, in: *Proceedings IEEE Conf. Robotics and Automation*, Raleigh, NC (1987) 1098–1103.
35. Nevatia, R., Depth measurement from motion stereo, *Comput. Vision Graph. Image Process.* **9** (1976) 203–214.
36. Regan, D. and Beverley, K.I., Binocular and monocular stimuli for motion in depth: Changing disparity and changing size feed the same motion in depth stage, *Vision Res.* **19** (1979) 1331–1342.
37. Ballard, D.H. and Kimball, O.A., Rigid body motion from depth and optical flow, *Comput. Vision Graph. Image Process.* **22** (1983) 95–115.
38. Richards, W., Structure from stereo and motion, *J. Opt. Soc. Am. A* **2** (1985) 343–349.
39. Waxman, A.M. and Sinha, S.S., Dynamic stereo: Passive ranging to moving objects from relative image flows, *IEEE Trans. Pattern Anal. Mach. Intell.* **8** (1986) 406–412.
40. Terzopoulos, D., Regularization of inverse visual problems involving discontinuities, *IEEE Trans. Pattern Anal. Mach. Intell.* **8** (1986) 413–424.
41. Hunter, S.C., *Mechanics of Continuous Media* (Ellis Horwood, Chichester, England, 2nd ed., 1983).
42. Courant, R. and Hilbert, D., *Methods of Mathematical Physics* I (Interscience, London, 1953).
43. Witkin, A., Scale space filtering, in: *Proceedings IJCAI-83*, Karlsruhe, F.R.G. (1983) 1019–1021.
44. Duda, R.O. and Hart, P.E., *Pattern Classification and Scene Analysis* (Wiley, New York, 1973).
45. Lapidus, L. and Pinder, G.F., *Numerical Solution of Partial Differential Equations in Science and Engineering* (Wiley, New York, 1982).
46. Press, W.H., Flannery, B.P., Teukolsky, S.A. and Vetterling, W.T., *Numerical Recipes: The Art of Scientific Computing* (Cambridge University Press, Cambridge, England, 1986).
47. Terzopoulos, D., Matching deformable models to images: Direct and iterative solutions, in: *Topical Meeting on Machine Vision*, Technical Digest Series **12** (Optical Society of America, Washington, DC, 1987) 160–167.
48. Rosenberg, D., Levine, M.D. and Zucker, S.W., Computing relative depth from occlusion cues, in: *Proceedings Fourth International Joint Conference on Pattern Recognition*, Tokyo, Japan (1978).
49. Blum, H.A., A transformation for extracting new descriptions of shape, in: *Proceedings Symposium on Models for Perception of Speech and Visual Form*, Boston, MA (1964) 362–380.
50. Brady, J.M. and Asada, H., Smoothed local symmetries and their implementation, *Int. J. Rob. Res.* **3** (1984) 33–61.