COMPUTER VISION AND IMAGE UNDERSTANDING
Vol. 70, No. 2, May, pp. 142–156, 1998
ARTICLE NO. IV970620

Nonrigid Motion Analysis: Articulated and Elastic Motion*

J. K. Aggarwal, Q. Cai, W. Liao, and B. Sabata†

Computer and Vision Research Center, Department of Electrical and Computer Engineering, The University of Texas at Austin, Austin, Texas 78712

Received July 17, 1995; accepted February 19, 1997

Motion of physical objects in the world is, in general, nonrigid. In robotics and computer vision, the motion of nonrigid objects is of growing interest to researchers from a wide spectrum of disciplines. The nonrigid objects being studied may be generally categorized into three groups according to the degree of deformation of body parts: articulated, elastic, and fluid. In articulated motion, individual rigid parts of an object move independently of one another and the motion of the whole object is nonrigid in nature. Elastic motion is nonrigid motion that conforms to a certain degree of continuity or smoothness. Fluid motion violates even the continuity assumption and may involve topological variations and turbulent deformations. This paper presents an overview of existing work on articulated and elastic motion, motivated by problems relating to the motion of the human body and of an animal heart, respectively. We study various approaches for recovering the 3D structure and motion of objects through a sequence of images in a bottom-up fashion, a strategy widely employed by various investigators. These approaches are classified as (1) motion recovery without shape models, and (2) model-based analysis. In the discussion of each algorithm, we also include a description of the complexity of feature and motion constraints, which are highly related to each other.

1. INTRODUCTION

In computer vision research, motion analysis has been largely restricted to the study of the motion of rigid objects. However, in the real world, nonrigid motion of objects is far more common. Driven by a wide range of applications such as medical imaging, sports training, image compression, graphics animation, video conferencing, and content-based query of multimedia databases, there is a growing interest in the study of nonrigid motion.

Nonrigid motion may be generally classified into three groups [41] according to the degree of deformation of body parts: the motion of rigid parts, the motion of coherent objects, and the motion of fluids. Motion of rigid parts occurs in situations where the individual rigid parts of an object move independently of one another. In this case, the motion of each constituent part is rigid, but the motion of the whole object is nonrigid. This type of motion is known as *articulated* motion. On the other hand, nonrigid motion of coherent objects includes motions such as the beating of a heart, the waving of a cloth, or the bending of a metal sheet, where the shape of the object deforms within certain continuity constraints. This type of nonrigid motion is classified as *elastic* motion. In this paper, we consider these two specific types of motion. Our study is motivated by problems encountered in computer vision studies of the human gait and the motion of the human body as well as the motion of an animal heart, where the constraints are related to continuity and smoothness. In both cases, it is of considerable importance to characterize both the structure and motion of the studied object. The motion of fluids is not considered in this paper.

The analysis of articulated motion, especially human motion, has been motivated by a large number of applications. For instance, automatic tracking of human motion using cameras can be applied to surveillance and traffic monitoring. If we process the motion analysis further to gait level and recover the 3D structures of articulated objects, it may be useful to applications such as (1) kinematic analysis of the movements of athletes in scientific training, (2) clinical evaluation of patients in biomechanics, (3) computer graphics and animation in entertainment, and (4) man–machine interfaces via gesture signaling. The applications of elastic motion also span a wide spectrum. For example, applications related to facial recognition and recovery, model-based image compression, animation of nonrigid objects, and clinical examination of the left ventricle, all involve the analysis and 3D structure recovery of moving objects with elastic properties.

Since nonrigid motion encompasses a wide range of possible motion transformations, a general paradigm for estimating motion parameters would be extremely difficult, if not impossible, to develop. Investigators have proposed various approaches to deal with different motion transformations, along with certain restrictions imposed on the object behavior. This paper examines the trends in the research on articulated and elastic motion. A taxonomy of motion types is included to give an overview of the general types of nonrigid motion and their formal definitions.

This paper is an extension of an earlier study [1] and is organized as follows: Section 2 delineates the classification schemes proposed by Kambhamettu *et al.* [50]. Sections 3 and 4 provide

* The research reported in this paper was supported in part by the Texas Higher Education Coordinating Board Advanced Technology Program (Project 442) and the Army Research Office (Contract DAAL-04-94-G-0017).

† Current address: SRI International, 333 Ravenswood Ave., Menlo Park, CA 94025.

1077-3142/98 $25.00

reviews of *articulated* and *elastic* motion, respectively. Finally, our conclusions on the current state of nonrigid motion analysis are presented in Section 5.

2. TAXONOMY OF MOTION TYPES

Nonrigid motion was first classified into three categories—articulated motion, elastic motion, and fluid motion—by Huang [41]. Goldgof *et al.* [34] refined the classification according to the mean and Gaussian curvature changes of the object surface and defined three restricted classes: isometric, homothetic, and conformal motion.

Kambhamettu *et al.* [50] proposed an extended classification scheme based on the degree of nonrigidity of the objects. Figure 1 depicts the classification tree for the motion of objects according to Kambhamettu's definition. The motion of objects is generally divided into two classes: rigid and nonrigid motion. Nonrigid motion is further subdivided into the restricted and general classes according to the degree of deformation. Articulated motion falls into the restricted class, whereas elastic motion falls into the general classes of nonrigid motion.

The definitions of the different motion classes are briefly described as follows:

- *Rigid motion* preserves all distances and angles and has no associated nonrigidity.
- *Articulated motion* is piecewise rigid motion. The rigid parts conform to the rigid motion constraints, but the overall motion is not rigid.
- *Quasi-rigid motion* restricts the deformation to be small. A general motion is quasi-rigid when viewed in a sufficiently short interval of time.
- *Isometric motion* is defined as motion that preserves the distances along the surface and the angles between the curves on the surface.
- *Homothetic motion* is motion with a uniform expansion or contraction of the surface.
- *Conformal motion* is nonrigid motion which preserves the angles between the curves on the surface, but not the distances.

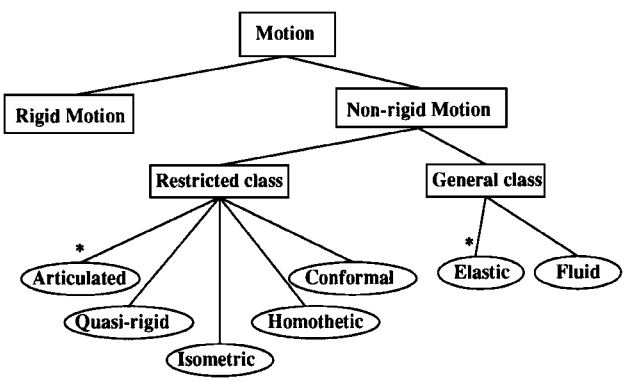

FIG. 1. The classification tree for motion of objects defined by Kambhamettu *et al.* The asterisk indicates the two areas we focus on in this paper.

- *Elastic motion* is nonrigid motion whose only constraint is some degree of continuity or smoothness.
- *Fluid motion* violates even the continuity assumption. It may involve topological variations and turbulent deformations.

The following review is restricted to two categories of nonrigid motion, namely, articulated and elastic motion, due to the interests of the authors. In discussing each type of motion, we follow the bottom-up strategy which most investigators apply to accomplish various high level tasks.

3. ARTICULATED MOTION

The focus of this section is on *articulated motion*, in particular, motion analysis which utilizes the characteristics of object's shape and jointed body parts. A number of researchers [3, 13, 17, 33, 78] have focused on high-level processing of motion analysis using temporal data of joints in 2D or 3D. Articulated motion is considered as an ordered sequence of various states of phase vectors [13, 17]. Recognition of different gestures can be reduced to finding the specific order of these atomic motion states corresponding to prototype gestures. We do not present a detailed discussion of this type of approach because the nonrigidity of the human shape is not considered.

Studies of articulated motion may be classified as those that use *a priori* shape models and those that do not. This classification relies on whether knowledge of the object shape is employed in the motion analysis. Both classes follow certain standard procedures to accomplish the defined tasks by using a bottom-up approach: (1) feature extraction, (2) feature correspondence, and (3) motion analysis. In cases where the *a priori* shape of the object of interest is unknown, the analysis usually starts by extracting object features, such as joint or corner points, edges, surfaces, and 3D blocks. Certain assumptions about the object motion are then imposed to establish feature correspondence. Based on the identified feature correspondence, high-level tasks such as moving body segmentation, joint location, and 3D recovery of the structure can be performed. In approaches using *a priori* shape models, the feature correspondence phase is replaced by matching the 2D image sequence to the 3D geometrical model data. Therefore, the model configuration determines the complexity of the matching. Typically, the shape of the object of interest (usually human form) is modeled by either stick figures or volumetric graphics. Human motion is normally measured by the movements of the lower limbs [12, 52, 69, 101], such as the velocities of the hip, knee, and ankle, or the angular velocity of various body parts. After the relationship between the image and model data is established, the 3D structure of the object can be fully recovered.

Although the two methodologies described above are based on different scenarios, they also can be combined to accomplish complex and high-level tasks. For example, a complete system can start with extracting coarse features at low level and perform the basic identification. Then a model with finer details can be used to make the feature correspondence and motion analysis

more accurate. Using the *a priori* shape model also verifies the former identification. Niyogi and Adelson's work [69] fits this scenario closely. Other work discussed subsequently more or less focuses on a certain subset of this scenario.

3.1. Motion Analysis without *a Priori* Shape Models

Among the approaches that do not use *a priori* shape models, there is a large diversity in the feature characteristics and motion assumptions that determine the complexity of feature correspondence, motion analysis, and other high level processes. These problems are discussed below.

3.1.1. Feature Characteristics

The simplest feature is probably the well known *moving light display* (MLD) [31, 46, 82]. Johansson [46] first showed that human eyes can interpret a moving human-like structure with MLDs only. Further studies by Rashid [82] found that the strong relationship between the movements of related points can be used directly for 3D reconstruction. Similar to Johansson's experiments, Rashid used projected MLDs of a person walking along different paths as input. Along this vein, Webb and Aggarwal [100] extended the study of the Johansson-type figures for a 2D image sequence. The object studied by Qian and Huang [80] is similarly treated as a collection of rigid lines linked by points. The difference is that the stick figures in [80] are obtained from synthetic 2D data. In [38], Hel-Or *et al.* measured 3D positions of the feature points from stereo image pairs instead of using the 2D projection of 3D points as the input. Motion analysis is conducted based on the location of each point in the local coordinates of the component where the point belongs and the orientation of this component relative to the viewer coordinates.

Recent work by Kurakake *et al.* [57] attempts to locate the joints of articulated objects through motion using extracted ribbons. Their use of 2D data from real images as input requires more preprocessing of the image, such as edge detection, line fitting, and ribbon extraction, to achieve satisfying results. Similarly, Kakadiaris *et al.* [48] used 2D contours as shape representations for segmentation and motion estimation. Joints are determined as the overlapping area of two moving subparts based on motion and shape information.

The 3D block is another feature being studied. The motion analysis by Asada *et al.* [7, 8] deals with line drawings generated by a computer, which serve as the orthogonal projection of a time-varying block's world. Recent work of Eggert *et al.* [30] utilizes range images of 3D blocks and converts the point information into complete 3D boundary descriptions. In [30], the object is made up of three portions: part, connection, and linkage. All the vertices used for feature matching are characterized by their 3D coordinates.

3.1.2. Motion Assumptions

It appears that no versatile algorithms exist for the nonrigid motion analysis of complex and variable nonrigid objects. Instead, algorithms use certain assumptions to solve problems in a particular area, mostly to simplify the process of feature correspondence. One widely used assumption is known as the *small motion assumption*, i.e., the assumption that the motion between consecutive images is small. This is addressed or indicated in a fairly large number of publications, e.g., [7, 8, 30, 48, 57, 100]. Another approach is to add constraints on the change of movements spatially or temporally. For example, smooth motion constraints [82], such as constant velocity or constant acceleration [7, 8], can reduce the ambiguity of feature matching between adjacent frames. In addition to smooth motion constraints, Rashid [82] also adopted the heuristic assumption that points belonging to the same object should have a higher correlation in the projected position and velocity. In a similar scenario, Webb and Aggarwal [100] imposed the *fixed axis assumption*, i.e., that "every rigid object movement consists of a translation plus a rotation about an axis that is fixed in direction for short periods of time." In their work, the motion of the object is assumed to be in a plane parallel to the image plane, which justifies the use of orthogonal projection. The rigidity of the connected parts makes it possible to recover the 3D structure of jointed objects. Qian and Huang [80] assume perspective projection instead. Other constraints, such as coplanar motion, fixed axis, and at least one known joint, are also imposed to aid the construction of their system. In [38], constraints were imposed based on the rigidity of articulated parts, such as constant distance, parallelism, and coplanarity between certain points.

3.1.3. Feature Correspondence

Feature correspondence is usually the most important and difficult task in motion analysis. Since the input is usually unlabeled and not explicitly identified in each frame, tracking features from one frame to the next becomes especially challenging. As mentioned earlier, well-defined constraints are usually introduced to eliminate invalid matches and distinguish unique connections. The characteristics of the shape representation also influence the complexity of the matching algorithm. There is always a trade-off between the extraction of features and the establishment of feature correspondence. For higher level features, the tracking and matching process become relatively easier since the number of features is much smaller compared to low level ones. Extraction of stable high-level features from an image, however, remains a tough problem.

Rashid [82] made good use of the smooth motion assumption for matching between consecutive frames. Feature correspondence is established by minimizing the difference between the expected position of each point and the actual position of the corresponding point. MLDs of simple objects can be tracked accurately, but velocity information must be incorporated for complex objects. He tried to get around the difficulty of determining the initial velocity by simply assuming that tracking always starts at a clearly interpreted frame. Hel-Or *et al.* [38] applied the Extended Kalman Filter to enforce the imposed constraints into the pose solution. The correspondence is considered valid when the

error between the predicted position and measured position is minimized subject to the rigidity constraints of the subparts.

Compared to points, using 2D contours as a high-level feature reduces the possibility of false matching but increases the burden of preprocessing tasks. After successfully extracting ribbons and their symmetric axes, Kurakake *et al.* [57] classify ribbons into layers in terms of their widths. Small interframe motion ensures correspondence between ribbons in the same layer with a similar axis angle. Matching one ribbon in one frame with two ribbons in the next frame is allowed in case the articulation appears or disappears. Detected articulations are verified using matching results. In Kakadiaris's work [48], feature correspondence is solved by accumulating knowledge of the 2D deformable model, which is derived solely from previous images. The model is initially defined as a single part object. As different postures of the object occur, the object evolves into various subparts based on the *bending deformation*, consisting of three zones with different motion types: the *fixed zone*, the *bending zone*, and the *relocation zone*. Therefore, the model deforms into different part models accordingly. At the end of the process, joints are located as the connections of the subparts based on motion.

Tracking the motion of 3D blocks is more complicated due to the possible occlusion of surfaces, edges, or vertices of objects. Based on the work of Huffman [42], Clowes [22] and Waltz [99], Asada *et al.* [8] applied a transition table of junction labels and contextual information between consecutive frames to analyze the structural change. The vertices are used as the primitive for feature matching. All the labeled lines are segmented into objects, and the correspondence of each junction throughout the image sequence is established using an object-to-object matching method. The 3D vertex matching in Eggert [30] is conducted by using the maximal point matching scheme of Chen and Huang [20]. The algorithm produces groups of three or more points satisfying local distance and angular constraints in each group. Constraints such as shape rigidity, planarity and small interframe motion are incorporated in order to eliminate mismatched sets and identify the proper correspondence. The optimal match is found by selecting the minimum among the sums of the square error of all possible attributes. In the final stage, the global consistency of point sets is guaranteed throughout the entire image sequence, eliminating the possibility of "accidental" alignment.

3.1.4. Motion Analysis

Once the feature correspondence is established, motion vectors are recovered to reveal the underlying structure of the object. High-level tasks such as segmentation, image compression, gait recognition, or structure recovery can be subsequently pursued.

Rashid [82] took advantage of the heuristic observation that velocities and positions of points belonging to the same object in consecutive frames are highly correlated. He constructs a graph with four dimensional nodes, which is formed by the projected positions of points along with their projected velocities. The edges of the graph are weighted by the Euclidean distance between the nodes. Then a *minimal spanning tree* is formed and the graph is separated into clusters based on the cost of the cut function [72]. Similar techniques are applied to determine the intraobject relationship, except that rotation becomes an important factor in the analysis.

In Kurakake's work [57], ribbon motion is used to locate the joints of the moving object. After ribbons are extracted, those representing the background or still objects are removed from consideration according to the displacement between adjacent frames. The articulations are located among connected or close ribbons. Finally, the description of the structure of the object of interest is integrated. Recent work of Kakadiaris *et al.* [48] uses a similar scenario. They combine the three courses, which are segmentation, shape, and motion estimation, together in the evolution of the data-driven deformable model. Joints of moving parts are identified as the connections of the existing segments.

The work by Asada *et al.* [7, 8] assumes that each object consists of a main body and its subparts (see Fig. 2). Therefore, they focus on multiaxis motion analysis and object segmentation based on the rigidity of the part shape. 3D geometrical parameters such as orientation and edge lengths are considered as matching primitives for the motion. Motion is segmented into subsequences of translation and rotation only. They use a technique called *Gaussian sphere mapping* [44] to find the number of axes along with the rotation vectors from an image sequence. Additional constraints such as constant velocity or acceleration of the vertices are employed to resolve ambiguity. Along this vein, Eggert *et al.* [30] sought to obtain the potential connections between the parts after the feature correspondence is established. Linkages are detected by computing causation ratios, which account for the relationship between the movement of two parts.

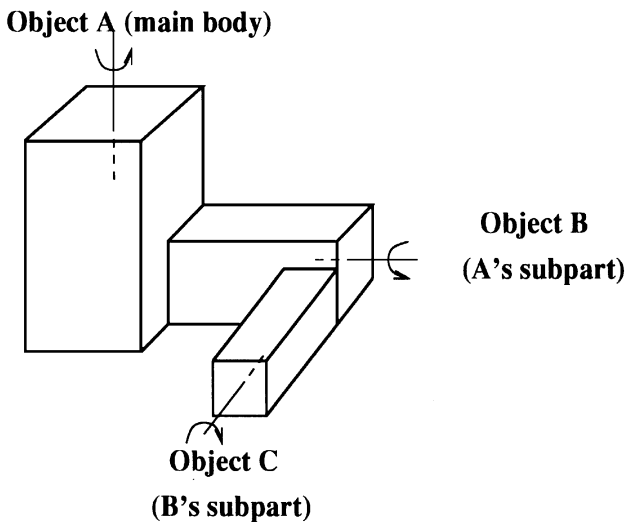

FIG. 2. The main body and its subparts (derived from Asada [7, 8]).

Among the high-level processes in motion analysis, one of the conventional tasks is to reconstruct 3D structure, using either spatial or temporal information. Kakadiaris *et al.* [47] attempted to reconstruct the 3D structure of a human body part from three mutually orthogonal views. Since the use of multiple cameras increases the computational cost, monocular methods are normally used in the study of articulated motion when depth information is needed. In Webb and Aggarwal's work [100], the *fixed axis assumption* indicates that a point connected to the fixed axis lies in a circle during its movement. When the circle is projected into the 2D image plane, it becomes an ellipse. The ellipse is determined from a real image sequence. Thus, the 3D structure of points belonging to the rigid parts of an articulated object can be computed through the fitting of the ellipse. Since the validity of the system depends on the shape of the projected ellipse, the method is void if there is no rotation, or if the rotational axis is perpendicular or parallel to the image plane. In these cases, either there is no circle, or the projected circle is a circle without depth information, or a line. Qian and Huang's work [80] focuses on the high-level theoretical analysis of articulated motion. They attempt to estimate the depth of the joints in a stick figure via decomposition of the object. The system equations are established on the assumptions of the rigidity of the parts and coplanar motion. Uniqueness of the solution is analyzed for three cases: (1) two points in four frames, (2) two points in three frames, and (3) three points in two frames. Then, they adopt the continuation method to estimate the depth information of the whole system. Experiments on synthetic data have verified their algorithms. In their method, the system equations are hard-constrained and thus sensitive to noise corruption.

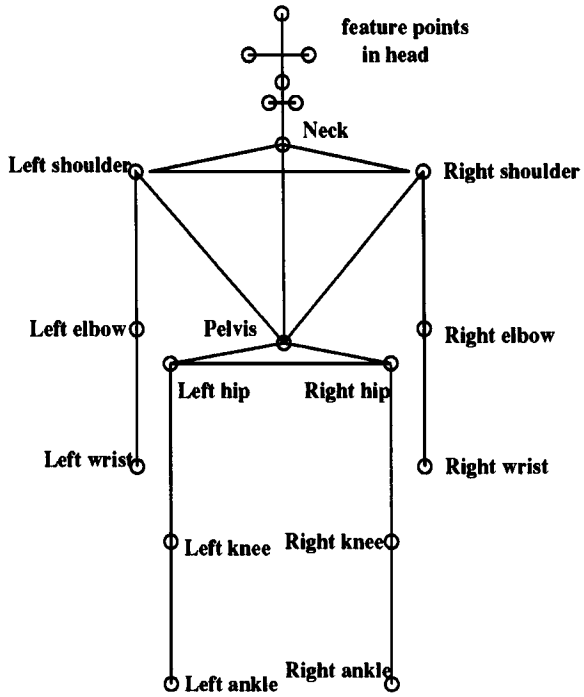

FIG. 3. A stick-figure human body model (based on Chen and Lee's work [7]).

3.2. Approaches Using *a Priori* Shape Models

Although motion recovery without the use of an *a priori* shape model is highly desirable, it introduces additional complexity to the tracking procedure. Sometimes, difficulties in feature correspondence cannot be overcome without using *a priori* knowledge of the object shape. Everyday experience indicates that human beings usually "see" an object by comparing it to prior knowledge of similar objects in the memory. Therefore, model-based approaches come up naturally. Since most of the previous work in this area has dealt with human body movements, we restrict our review to the human form as the object of study.

In the rest of this subsection, we will first discuss various models of the human body and the analysis of human motion. Then, we will address the high-level procedures such as body matching, part recognition, and 3D structure determination.

3.2.1. Human Model Definitions

Conventionally, the human body is represented either by a stick figure (Fig. 3) or a volumetric model (Fig. 4). The stick figure representation [2, 12, 21] is based on the observation that human motion is essentially the movement of the human skeleton brought about by the attached muscles. Volumetric models (such as 2D ribbons [58], generalized cones [2], elliptical cylinders [39, 84], and spheres [71]) better represent the shape of the human body, but require more parameters for computation. Each model can be scaled according to the height of the subject. Akita [2] and Perales *et al.* [77] incorporate both stick figures and volumetric models to perform different levels of matching. Table 1 lists the researchers and the human models used in their research.

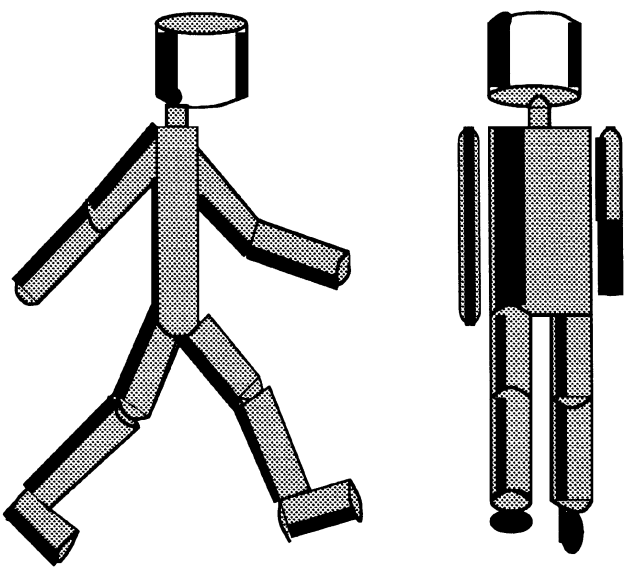

FIG. 4. A volumetric (cylinder) human body model (viewed from two directions), derived from Hogg [14].

TABLE 1
Researchers and Their Human Models

Researcher's name	Human model: Stick figure	Human model: Volumetric model
O'Rourke and Badler		√ sphere
Hogg		√ elliptical cylinder
Akita	√	√ cone
Chen and Lee	√	
Okawa and Hanatani		√ 2D mask
Yamato *et al.*		√ 2D mesh
Kinzel		√ 2D region
Leung and Yang		√ 2D ribbon
Rohr		√ elliptical cylinder
Bharatkumar *et al.*	√	
Perales and Torres	√	√ various 3D primitives

Chen and Lee's 3D stick figure model [21] contains 17 segments and 14 joints for the head, torso, hip, arms, and legs. Both torso and hip parts are assumed to be rigid. The whole stick figure model is described with the 3D space of the joints and the length of each rigid segment. The identification of a pose of human motion in this model involves the joint location in 3D coordinates and the 3D joint angle measurements. Bharatkumar *et al.* [12] use a 2D stick figure to model the lower limb of the human body, where joints such as hip, knee, and ankle are considered.

A large body of literature concerns the shape recognition of 2D contours of human parts, especially for lower limbs. The 2D features vary from the lower level masks [70] and meshes [101], to higher level regions [52] and ribbon representations [58]. Among them, the 2D ribbon model proposed by Leung and Yang [58] is the most complicated. It contains two components: the basic body model and the extended body model. The basic body model outlines the structural and shape relationships between the body parts. Several rules are imposed along with the structure, such as structural constraints, shape constraints, and balancing constraints, to ease the body labeling process. The extended model consists of three patterns: the support posture model, the side view kneeling model, and side horse motion model. It attempts to resolve ambiguities in the interpretation process by identifying a certain pattern from the outlined picture.

A collection of elliptical cylinders is also commonly used in modeling nonrigid forms. The models in both Hogg [39] and Rohr [84] fall into this category; more specifically, they are the cylinder models originated by Marr and Nishihara [61]. In this model, the human body is described by 14 elliptical cylinders. Each cylinder is controlled by three parameters: the length of the axis, and the major and minor axes of the ellipse cross section. The origin of the coordinate system is located at the center of the torso. Along the same vein, Rehg *et al.* [83] rendered two occluded fingers with several cylinders, and the center axes of the cylinders are projected into the center line segments of the 2D finger images.

O'Rourke and Badler [71] use an elaborate sphere model consisting of 24 rigid segments and 25 joints. The surface of each segment is defined by a collection of overlapping sphere primitives. A coordinate system is embedded in the segments. The model also includes the constraints of human motion, such as restrictions on joint angles, and an algorithm to detect collisions between nonadjacent segments.

Recent work by Goncalves *et al.* [35] addressed the problem of motion estimation of a human arm in 3D using a calibrated camera. Both the upper and lower arm are modeled as truncated circular cones, and the shoulder and elbow joints are assumed to be spherical joints. The hand tip is considered to be an extension of the forearm axis.

As mentioned previously, both stick figures and volumetric models can be integrated to form a more comprehensive system for coarse-to-fine processing. In Akita's work [2], key frames of 2D stick figures are used to guide the approximate order of the motion and spatial relationships between the body parts. These key frames represent the crucial moment of a changing pose. Thus, each frame differs from its predecessor and successor in the place where the body segments cross or uncross each other. The stick figure has six segments: head (close to a single point), torso, two arms, and two legs. No joint is explicitly defined. A cone model is included to provide knowledge of the rough shape of the body parts. Each part corresponds to the counterpart of the stick figure model. Recent work by Perales and Torres [77] introduces a predefined library with two levels of biomechanics graphical models. Level One is a stick figure tree with nodes for body segments and arcs for joints. Level Two is composed of descriptions of surface and body segments constructed with various 3D primitives used in computer graphics. Both levels of the model are applied in different matching stages.

In contrast, Kuch *et al.* [55] build a generic model of the surface of a human hand using cubic B-splines to model the surface of the hand. Their model consists of 300 control points for rendering into a smoothed surface. Any given hands can be approximated by adjusting the variations of these control points based on the calibration results from three camera views. The focus of their work is on modeling and then rendering it to real human hands. Tracking of human hands is achieved by minimizing the error between the real hand image and the model projection based on the assumption that the initial orientation of the hand is available.

3.2.2. Articulated Motion Models

A substantial amount of research on articulated motion has been focused on human motion. Human motion can be described in terms of kinetics or kinematics. Kinetics involves the study of the forces/torques involved in generating the movements. Kinematics, on the other hand, concerns the geometry of the object, such as its position, orientation, and deformation. Most of the model-based approaches in computer vision are concerned with studies of the kinematic patterns.

In kinematics, human motion is usually characterized by joint angles and the horizontal/vertical displacements of joints, as initiated in medical studies by Murray [66]. Along this vein, Rohr [84] describes human walking with the joint angles of the hip, knee, shoulder, and elbow within a walking cycle. In his experiments, raw kinematic data are interpolated by periodic cubic splines to create smoothed patterns. Among the motion of different body parts, the movement of the lower limbs has been extensively studies in the past [12, 52, 69, 101]. Compared to the upper trunk and arms, the lower limbs usually move more regularly and maintain their shapes in a more consistent fashion. This is indeed the reason why the prototyped hidden Markov model (HMM) succeeds in recognition of human activities in an image sequence based on simple mesh symbols [101]. In [12], Bharatkumar *et al*. [12] used kinesiology data as the basis for their walking human model. They measure the projected 2D angles of the hip and the knee for one walking cycle and fit them with cubic splines as the model data. Horizontal displacement of the hip and ankle is also used as a part of the model. A high correlation has been found between the model and the stick figure of a subject using medial axis transformation. Similarly, Kinzel [52] employs the hip and knee angles for modeling limb movements. In contrast, Niyogi and Adelson's [69] exploit the repetition information of the lower limb trajectory for coarse recognition of a human walking. They assume that the human is walking at a roughly constant speed and parallel to the image plane, thus no depth information is considered. An XT-slice (x axis vs time) of the cube near the ankle is used as a braided signature for walking patterns. The XT-slice of the head shows a nearly straight line in an image sequence. These XT-slices are utilized to outline the contour of a walking human based on the observation that the human body is spatially contiguous. A reliable stick figure of the human body is extracted from the outline images. The corresponding gait was found and compared to the patterns in the database. Finally, the walking human is recognized from his gait signature.

Kinematic models of articulated motion also help in tracking self-occluding articulated objects. Rehg *et al.* [83] assumed an invariant visibility order of 2D templates of two occluded fingers to the viewing camera. The hand is assumed to rotate along the middle finger axis, thus there are three possible occlusion relations: the second finger occluded by the first, disjoint, and the first occluded by the second. The definition of this kinematic motion model simplifies the motion prediction and matching process between the 3D shape model and its projection in the image plane.

3.2.3. Part Location

Recognition of body parts is essential for high-level processes such as segmentation, tracking, and object recognition. It usually involves region tracking and body labeling. Region tracking comprises extracting the shape primitive of the subject and determining its location from frame to frame. The labeling process matches the body parts to their counterparts in the model, so that a meaningful description of the human motion and structure can be obtained.

Leung and Yang [58] applied the *different picture* method of Jain and Nagel [45] along with their own coincidence edge detection algorithm to generate a complete outline of the moving object images. They proposed a spatiotemporal relaxation process to determine which side of the moving edge belongs to the moving object. Two sets of 2D ribbons on each side of the moving edge, either a part of the body or that of the background, are identified according to their shape changes over time. The body parts are labeled according to the human body model. Then, a description of the body parts and the appropriate body joints is obtained. 2D human stick figures are produced as the final output.

In Akita's work [2], the first step is to obtain the outline image by an edge detector and set the viewer's space coordinate system in each frame. According to the stability of the different body parts, the labeling process is executed in the order of legs, head, arms, and trunk. Two methods are combined to establish the correspondence between frames regarding the degree of body movement. If the change of the body segments is small, the *window code* distance is used to locate the body parts, incorporated with their position in previous frames. Otherwise, the key stick frames are applied for dramatic changes of the posture due to occlusion.

Recently, Goncalves [35] used perspective projection of a 3D arm model to fit the blurred image of a real arm. Matching is conducted by recursively minimizing the error between the model projection and the real image by dynamically adapting the size and orientation of the model. Rehg [83] apply a similar approach except that the *window function* was adopted to exclude the effects of occlusion and locate the templates of finger images to be matched.

3.2.4. 3D Structure Determination

Since 3D structure reconstruction that depends solely on the analysis of real image data is extremely difficult and unreliable, most approaches [21, 39, 77, 84] establish a match between the 2D real image and the 3D graphical model. Once the relationship between the graphical and the real image data is discovered, one can directly employ the noise-free synthetic model data to simplify higher level tasks [71].

Chen and Lee [21] attempted to recover the 3D configuration of a moving subject according to its projected 2D image. A set of possible interpretations in 3D stick figures is obtained from the basic analysis of the gait. To eliminate a great number of false solutions, they propose a computational model using graph search theory. The search space is constrained by the smoothness and continuity of human motion. Physical constraints, such as angle constraints, distance constraints, and collision-free constraints, are imposed to exclude invalid matches. The transformation matrix from the body coordinate system to the viewer's coordinate system is constructed using the specific six feature points in the head. Then, the locations of the joints in the model are

transformed into the viewer's coordinates. Finally, the 3D structure of the whole body is recovered.

Hogg [39] presented a computer program called the WALKER, which attempts to "see" a walking person. The system is made up of three phases. First, the raw images sequence is input to a differencing algorithm to produce an isolated region of the moving object. The region including the moving object serves as an indication of the object's size, location, and rough posture. Next, the outline of the gray-scaled object is extracted using the Sobel operator and a fixed threshold. Then, an exhaustive search is performed to find the corresponding posture of the model that is the best match of the current image. The matching process is implemented by maximizing the plausibility value using a pixel-by-pixel Least Mean Squares (LMS) fitting. Finally, the description of the 3D structure for the walking person is generated. Rohr's work [84] is quite similar to Hogg's, except for the method of mapping the gray-scaled image to the corresponding model. After the outline of the moving object is extracted, Rohr applied the eigenvector line fitting [29] to approximate the edges of the image. Then, a similarity measure is calculated to locate the proper model. A recent study by Perales [77] follows a similar strategy, except that they define the model in a more elaborate way according to the input image and the user's assignment.

In O'Rourke and Badler's experiments [71], the images were taken from a computer-controlled human model. The segments of the head, hands, and feet were lightly colored to present a sort of *moving light display*. Predicted 3D regions corresponding to various body parts are fed into the *image analysis* phase to produce a more accurate location. After the range of the predicted 3D location of the body parts becomes smaller, the *parser* fits these location–time relationships into certain linear functions. Then, the *prediction* phase estimates the position of the parts in the future frame based on the determined linear functions. Finally, a *simulator* embedded with extensive knowledge of the human body translates the prediction data into corresponding 3D regions, which will be verified by the *image analysis* phase in the next loop.

4. ELASTIC MOTION

Elastic motion refers to the type of nonrigid motion whose only constraint is some degree of continuity or smoothness. In this domain of general deformable motion, there are no constraints other than topological invariance. Most approaches that deal with *elastic* motion assume an object model and then try to model the deformations as variations to the model parameters. A notable example of elastic motion is the motion of the heart (specifically, the left ventricle). Much of the research on elastic motion has concentrated on the analysis of left ventricular motion, due to its importance in aiding the understanding of the physiology of the normal heart and detection of cardiovascular diseases.

In general, model-based approaches have the advantage of constraining the degrees of freedom exhibited by the deformable objects. The correspondence problem is simplified as the incorporation of shape models significantly reduces the search space in the feature matching process. In some cases, the correspondence problem is transformed to a parameter estimation problem. Using physically based modeling primitives has the added advantage of resolving the 3D shape reconstruction problem usually encountered in nonrigid motion analysis. These methods are most effective when *a priori* knowledge about the motion structure is available. If the motion characteristics are not known in advance, applying such methods could cause errors in motion estimation as well as in shape recovery. In such cases, approaches that assume no prior knowledge about the motion or object structure can be utilized to determine the general properties of the nonrigid motion; then proper modeling primitives can be selected to refine or enhance the analysis. Thus, in some sense, these two general approaches can be integrated to yield a more reliable scheme for analyzing general nonrigid motion.

4.1. Motion Recovery without *a Priori* Shape Models

In this subsection, we discuss motion recovery strategies based on the most general assumptions of elastic motion, i.e., coherence of bodies and smoothness of motion. Since no explicit shape models are used, one must rely on image features to track the nonrigid motion. The analysis usually consists of three major stages: (1) feature extraction, (2) feature correspondence, and (3) motion and structure recovery. The overall scheme appears very similar to that used by stereo vision or rigid motion estimation. Nonetheless, we will see remarkable differences in each stage of processing as we proceed.

4.1.1. Feature Extraction

Establishing the correspondence between image features is usually the primary step in recovering structure from motion. Unlike situations in which the rigidity assumption is valid, when deformation is allowed, there usually are no image features that bear reliable structural and geometrical information about the objects. This significantly limits the type of matching primitives that can be used to analyze nonrigid motion. In general, two levels of features have been employed, i.e., low-level features such as points and high-level features such as contours.

Point features are widely used, not only because they are usually the only type of features found in many biomedical images, but also because they are invariant with respect to deformable transformations. Since point features are a fairly low-level representation, the accuracy of the matching results depends heavily on the validity of the motion smoothness assumption. In other words, large deformations cannot be properly addressed.

The contour of a 2D object or the bounding surface of a 3D object are classified as high-level features. These representations provide richer information than edges or collections of points. Therefore, in establishing the correspondence between contours (2D data) or surfaces (3D data), assumptions such as local rigidity or small interframe motion can be somewhat relaxed. In addition, since the contour contains structural information about the object, one expects the matching results to be more robust than

those obtained from low-level feature matching. A major problem, however, is that the extraction of reliable contours/surfaces from a sequence of nonrigid shapes is itself a difficult problem, especially in biomedical images that exhibit poor contrast. In such cases, incorporation of parametrically deformable models may assist the task of boundary detection [90].

4.1.2. Feature Correspondence

There are two general approaches for feature correspondence in nonrigid motion analysis: *explicit* and *implicit* matching methods. In explicit matching, the finite set of image features extracted from image sequences are examined and compared to arrive at one-to-one mappings based on certain likelihood measures. In implicit matching, distinct features are not isolated. Instead, an energy functional incorporating internal forces (e.g., smoothness constraint) and external forces (e.g., feature energy) is formulated and the feature tracking problem is converted to an optimization problem. In the following, we discuss strategies for matching low-level and high-level features using explicit matching methods. We also touch upon the topic of implicit, low-level feature matching. Implicit, high-level matching is closely related to physically based shape models and will be treated in latter subsections.

Explicit, low-level feature matching. Strickland and Mao [93] developed a relaxation algorithm to compute correspondence in image sequences containing nonrigid shapes. A special representation scheme consisting of a network of nodes connected by branches is used to describe the edge map. Critical points, such as branching points and corners, are selected as candidates of matching primitives. These points are matched using an iterative relaxation procedure, where the matching probability is updated according to the similarity between the geometry of neighboring nodes, the averaged motion vector, and the relative location of neighbors with respect to the current node. Even though all the above constraints favor local rigidity, the particular image structure permits global deformation between image sequences. This technique has been applied to track fluid flow features in combustion image sequences [92].

Shapiro *et al.* [87] presented a parallel strategy for tracking *corner* features on independently moving (and possibly nonrigid) objects. Their system consists of two components: a *matcher* and a *tracker*. The matcher computes the correspondence based on local patch correlation, while the tracker supervises the matcher over time, maintaining the motion trajectory as well as feeding the predictions to the matcher. The key assumption in this work, however, is that the displacement between two image frames is small, which is a severe restriction in many applications. Moreover, *corner* features might not be present in many practical situations, and other reliable means to extract physically meaningful points have to be utilized.

Liao *et al.* [59] have addressed the analysis of three-dimensional shape and shape change in nonrigid biological objects imaged via a stereo light microscope. They proposed a cooperative spatial and temporal feature matching process for stereo and motion analysis. The correspondence problem is cast into a four-dimensional maximum weighted matching problem, which belongs to the class of NP-complete problems. To get around NP-completeness, they use a relaxation labeling technique to effectively reduce the number of possible matches, and then apply a local search algorithm that returns a near-optimal solution to the modified problem. The approach is applicable to general feature types. Experimental results on real images of a frog's ventricle using *point* features are demonstrated.

Explicit, high-level feature matching. In [9], Bajcsy and Kovacic proposed a multiresolution elastic matching algorithm for medical applications. Contours of the brain atlas and the CT (Computerized Tomography) brain from three different scales are matched in a coarse-to-fine manner. After proper registration, matching starts at the coarsest level and the result is propagated through the next finer level. The finest level solution is then used to incrementally deform the model using an elastic constraint equation. This procedure continues until a satisfactory match is found. It is shown that by integrating information from different scales, robustness is achieved and the restriction of small deformation is alleviated.

Amini *et al.* [4] computed the displacement vectors between two successive contours using a two-stage process. The first stage involves the initial estimation of the displacement vectors by modeling the contour as an elastic rod and minimizing the bending energy. In the second stage, a smooth vector flow field is obtained by refining the initial displacement vectors based on another energy function. In [5, 6], the method is extended to allow surface tracking of nonrigid objects undergoing conformal motion.

Cohen *et al.* [23] presented a method to track points on deformable contours using curvature information. High curvature points on the contour are used as landmarks to guide the matching process. Assuming that the boundaries have been successfully extracted, they formulate an energy functional which preserves the matching between high curvature points while ensuring smooth flow field everywhere. This approach appears to be very similar to [6], except that an explicit description of the mapping between the contours to be matched is provided.

Serra and Berthod [86] introduced an algorithm for optimal subpixel matching of contour chains and segments. The algorithm relies only on the geometrical properties of the contours, eliminating the need for parametrization of the deformation. Each contour is treated as a collection of small linear segments. Contour matching is achieved by minimizing a *deformation measure* between the linear segments using a dynamic programming framework.

Implicit, low-level feature matching. In implicit matching, extraction of feature sets is not required. Instead, matching is accomplished on a global basis, where a one-to-one relationship between local features might not exist or is unimportant. A classic example of this type of matching strategy is the computation of optical flow by Horn [40]. Due to its sensitivity to occlusions,

discontinuities, and noise, however, this method has shown only limited success in practical situations.

Bartels *et al*. [10] introduced the concept of material coordinates in modeling shape changes in microscopic images. Even though the basic assumption that brightness does not change too much between image frames is similar to that in optical flow analysis, their algorithm seeks a deformation transformation between the material coordinates rather than computing the flow vector explicitly. Concepts from differential geometry are applied to assist the analysis. The variational technique employed to optimize the energy functional, however, yields complicated nonlinear, second-order, coupled partial differential equations. Convergence of the numerical solution is thus an issue of concern.

Kambhamettu and Goldgof [49] studied the recovery of point correspondence in *homothetic* and *conformal* motion—two restricted classes of nonrigid motion [50]. In these classes of confined nonrigid motion, there is a mapping relating the Gaussian curvatures of the surface at corresponding points. By hypothesizing all possible point correspondences in a small neighborhood and minimizing the error function which captures the deviation from the underlying motion class for each hypothesis, locally consistent matches between points are obtained. Once the point correspondence is recovered, stretching parameters can be estimated.

Chaudhuri and Chatterjee [18] established implicit point correspondence by means of subset matching. Subsets are created by grouping points lying on a single surface. Thus, the correspondence problem is reduced to identifying the same surface at different time intervals. The results of matching are used to compute translation, rotation, and deformation parameters for nonrigid objects.

Huttenlocher *et al*. [43] addressed the tracking of nonrigid objects in complex scenes. They decompose the image of a moving object into two components: 2D shape change and 2D motion. The 2D shape is represented by a collection of points. Correspondence between shapes is achieved by comparing the Hausdroff distance between two group of points.

4.1.3. Motion and Structure Recovery

Motion recovery without the explicit use of shape models fails to address the important problem of structure modeling. Establishing explicit correspondence between image features only gives flow vectors at certain locations. Interpolation techniques must be applied to obtain a dense map [93]. Implicit point matching algorithms produce denser optical flow fields, but the results are usually noisy and necessitate some smoothing process. On the whole, additional processing is required to reconstruct the motion transformation of the entire body. Unfortunately, this crucial issue has been largely ignored by most researchers.

4.2. Model-Based Approaches

In nonrigid motion analysis, dynamic shape modeling provides the mechanism for fitting and tracking visual data. It plays the dual role of assisting motion estimation as well as structure recovery. Using deformable models, the seemingly unstructured elastic motion can be compactly represented by a small number of parameters. The task of motion recovery is often reduced to the problem of parameter estimation. In this subsection, we will discuss two classes of shape modeling primitives: *parametric* models and *physically based* models. Parametric models are suitable for describing the shape of static objects and, with proper modification, can also be used to model deformable surfaces. Physically based models are fundamentally dynamic and are governed by the laws of rigid and nonrigid dynamics expressed through a set of Lagrangian equations of motion. They are more versatile in terms of modeling complex local deformations.

4.2.1. Parametric Models

A number of parametric surface models have been proposed for geometric shape representation. Parametric models concisely capture the global shape of the objects, which sometimes implies that heuristic, *a priori* knowledge about the object must be provided. Consequently, most parametric models are able to represent only a limited class of objects and are not well suited to model dynamically deformable objects without proper modification. In the following, we briefly discuss these representation schemes and comment on their usage and limitations in dealing with nonrigid shapes.

Polynomials. Polynomials have been extensively used because of their simplicity [11]. Spheres, cylinders, and cones are some simple second degree surfaces that are commonly used. In most cases, these simplified shape primitives are very coarse abstractions and can, at best, provide some qualitative measure of the visual data. For a discussion of dynamic shape representation using polynomials, the reader is referred to [53].

Generalized cylinders. Generalized cylinders [79] are constructed by sweeping an arbitrary two-dimensional set along an arbitrary axis $a(s)$, called the *spine*, in 3D space. They can represent either a solid or a surface. Many parameters such as surface area or volume can be easily computed from this representation. However, determining the axis $a(s)$ and fitting the visual data to the model are nontrivial tasks, especially when deformation of the object shape is likely to happen. They are better suited for modeling simple geometric shapes than complex, dynamic objects.

Spherical harmonics. Spherical harmonic functions are a complete, orthogonal set of functions on the sphere under the inner product

$$\langle f_1, f_2 \rangle = \int f_1 f_2 \sin\phi \, d\phi \, d\theta. \tag{1}$$

Therefore, any radial or stellar surface (surface obtained by deforming a sphere by moving points in the radial direction) can be represented by a sum of these basis functions $U_n^m(\phi, \theta)$ and $V_n^m(\phi, \theta)$

$$R(\phi,\theta)=\sum_{n=0}^{N}\sum_{m=0}^{n}A_{nm}U_n^m(\phi,\theta)+B_{nm}V_n^m(\phi,\theta). \quad (2)$$

Because of its ability to capture the details of local surface patch, the spherical harmonic function has been used by Chen *et al.* [19] as a surface modeling primitive for interpolating the local deformation of the left ventricle.

Splines. Splines have been widely used in computer-aided geometrical design and computer graphics since the early 1960's. Recently, they have found their way into modeling and data reduction applications in computer vision [96]. In fact, the *snake* model we will discuss later reduces to a spline function when external forces are removed. (We will present a detailed discussion on the snake model in the next section, due to its strong influence on the study of nonrigid shapes.) In a series of papers [14–16], Bookstein illustrated the potential applications of thin-plate splines, including the modeling of biological shape change, production of biomedical atlases, and image feature extraction. He demonstrated the decomposition of deformations by *principal warps*, which are geometrically independent, affine-free deformations of progressively smaller scales. Such a decomposition scheme is conceptually similar to conventional orthogonal functional analysis.

Superquadrics. Another class of modeling primitives that has received much attention recently is superquadrics, a family of parametric shapes derived from the parametric forms of the quadric surfaces. Solina and Bajcsy [89] discussed the recovery of parametric models from range images using superquadrics with global deformations. Mathematically, the equation for a superquadric surface is given as

$$e(\eta,\omega)=\begin{bmatrix}a_1C^{\epsilon_1}(\eta)C^{\epsilon_2}(\omega)\\ a_2C^{\epsilon_1}(\eta)S^{\epsilon_2}(\omega)\\ a_3S^{\epsilon_1}(\eta)\end{bmatrix},\quad \frac{-\pi}{2}\le\eta\le\frac{\pi}{2},\quad -\pi\le\omega\le\pi, \quad (3)$$

where $C^{\epsilon}(\eta)=sgn(\cos\eta)|\cos(\eta)|^{\epsilon}$ and $S^{\epsilon}(\eta)=sgn(\sin\eta)|\sin(\eta)|^{\epsilon}$. The variables η and ω correspond to latitude and longitude in a spherical coordinate system. ϵ_1,ϵ_2 are known as *squareness* parameters and a_1, a_2, and a_3 are scaling factors. By varying these five parameters, a variety of shapes can be obtained. While superquadrics are capable of modeling global shapes, they usually cannot give accurate descriptions of natural objects. Later, we will discuss how global deformations (bending, tapering, twisting, and concavity) and local deformations can be incorporated to enhance the modeling capabilities of superquadric surfaces. Hyperquadrics [37] are generalizations of superquadrics that allow smooth deformation from shapes with convex polyhedral bounds, although no explicit parameterized form is possible.

Implicit algebraic surfaces. An implicit surface S is defined as the zero set of some function f

$$S=\{\mathbf{x}\mid f(\mathbf{x})=0\},\quad \mathbf{x}=(x,y,z)^T. \quad (4)$$

In [94], Sullivan *et al.* used the parameterized algebraic surface defined by

$$f(x,p)=\sum_{j=1}^{q}p_jx^{k_j}y^{l_j}z^{m_j}=0 \quad (5)$$

for 3-D object modeling and recognition. The parameter vector $\mathbf{p}=(p_1,\ldots,p_q)$ is estimated from the visual data by minimizing the mean-squared algebraic distance between the model and the observed points. According to their definition, a quadric surface will possess a total of 34 coefficients and can be used to model fairly complex shapes. (In comparison, a superquadric surface, which is an implicit surface, but not an algebraic surface, has only 5 parameters.) Motion and deformation of the object is modeled by an affine transformation between two algebraic surfaces in that it preserves the surface degree.

Fourier decomposition. Staib and Ducan [90] proposed a parametrically deformable model for boundary finding based on the elliptic Fourier decomposition of a contour

$$\begin{bmatrix}x(t)\\ y(t)\end{bmatrix}=\begin{bmatrix}a_0\\ c_0\end{bmatrix}+\sum_{k=1}^{\infty}\begin{bmatrix}a_k & b_k\\ c_k & d_k\end{bmatrix}\begin{bmatrix}\cos kt\\ \sin kt\end{bmatrix} \quad (6)$$

The parameter vector to be recovered is $\mathbf{p}=(a_0,c_0,a_1,b_1,c_1,d_1,\ldots)$. The strength of their model stems from the combination of flexible constraints in the form of probabilistic models. Boundary detection is formulated as an estimation problem using a Maximum *A Posteriori* (MAP) objective function. In [91], the 2D model is extended to allow global shape parametrization of smoothly deformable 3D objects. Surface detection is formulated as an optimization problem, although no probabilistic model is involved. Deng and Wilson [28] also used a Fourier descriptor to represent the global shape of an object, but they allow local forces to interact with the global boundary so that local deformation is permissible.

4.2.2. Physically Based Models

As we have mentioned previously, parametric models are inadequate for the analysis and representation of complex, dynamic real-world objects. This deficiency is apparent in the case of nonrigid motion, and has motivated recent research into modeling methods based on computational physics. A number of physically based models have been developed for image analysis, including *snakes*, symmetry-seeking models, deformable superquadrics, deformable templates, and modal models.

The *snake* model is perhaps the most well known and widely used physically based model. A great deal of attention has been devoted to the extension of the named prototype first proposed

by Kass *et al.* [51]. The original snake model is a class of active contours that evolve under the influence of external potentials but are constrained by internal energies. When augmented by Lagrangian mechanics, dynamic snakes with intuitive physical behaviors are developed [97]. In [32], temporal context is explicitly incorporated to form a spatiotemporal solid called the *active tube*. Not surprising, the snake model can also be generalized to deal with 3D images, as shown by Cohen *et al.* in [24, 25]. This newly formed class of deformable surfaces, called *balloons*, is able to conform to image features and external forces in a way similar to the original snake model. An evolution equation similar to that in dynamic *snakes* has also been formulated. Cohen *et al.* have successfully applied this model to the segmentation of 3D MRI images as well as to establishing correspondence between a deformable surface and an anatomical atlas. McInerney *et al.* [62] have addressed similar problems using a slightly different balloon model. Their focus, however, is on the numerical solution of the dynamic deformable surface model.

A more recent extension to *snake* models can be found in [95], where the essential elements of physically based models and probabilistic approaches are incorporated. A Bayesian framework is introduced and the original energy-minimizing problem is transformed to an MAP problem. To further exploit the power of probabilistic modeling, Szeliski *et al.* [95] have developed a sequential estimation algorithm using the Kalman filter. Known as the *Kalman snake*, this dynamic system is able to integrate nonstationary, noisy observations over time. It provides the flexibility to design behaviors that may not be possible with purely physically based models. Moreover, model parameters can be derived from statistical models of sensors, rather than chosen heuristically. It would be useful to generalize their development to form *Kalman balloons* that are capable of tracking dynamic surfaces.

A significant amount of work in left ventricle (LV) boundary detection/tracking has been influenced by the *snake* model. For example, Ranganath [81] developed an automatic contour extraction procedure using a specially configured active contour model. Contour propagation between image slices (spatial continuity) and image phases (temporal continuity) is considered to enhance the reliability of the extracted boundaries.

Kumar and Goldgof [56] also based their feature-tracking algorithm on snake models. Spatiotemporal tracking of the SPAMM grid in cardiac MR images is achieved by using multiple parallel and vertical snakes. The intersections of snakes are used as markers for establishing point correspondences. After the correspondence problem is solved, a thin-plate spline model is utilized to interpolate the local motion vectors.

Along a similar vein, Young *et al.* [102, 103] proposed deformable models to reconstruct 2-D and 3-D heart wall motion from tagged MR images. Their approach appears to be very similar to that reported by Kumar and Goldgof, except that the latter used modified internal and image energies, while Young *et al.* used the original snake formulation and minimization procedure developed by Kass *et al*. Finite element method is employed to model the geometry and deformation of the LV. Verification of the proposed algorithms is achieved using a silicon gel phantom undergoing controlled deformation.

The semi-automatic tag tracking algorithm proposed by Kraitchman *et al.* [54] represents a combination of template matching and active contour modeling. Tag intersections are first detected by computing the normalized cross-correlation between an idealized template and the tagged MR image. An "active spring mesh" structure is then formulated to track the identified intersection points. The spring mesh has a different internal energy from the original snake model since discrete points, rather than lines or contours are being tracked. The image energy is chosen to be the correlation image to attract the intersection points to regions of high correlation.

Lobregt and Viergever [60] presented a *discrete* dynamic contour model with special properties to avoid undesirable effects such as shrinking and vertex clustering, which are common problems in existing active contour models. The internal energy of the model depends on local contour curvature, while the external energy is derived from image features. The resulting model is both conceptually and computationally simple, and has proven to be useful for segmenting MI images. Extension of their methods to 3D images, however, remains an open question.

The free-form deformable surface model proposed by Delingette *et al.* [26, 27] is conceptually similar to the active contour model [51]. They model an object as a closed surface that is deformed subject to attractive fields generated by input data points and features. A fundamental conflict in shape representation is that a modeling primitive should be general enough to handle a wide variety of scenes, yet simple enough to be usable for tasks such as recognition and manipulation. To balance these conflicting requirements, the authors suggested a coarse/fine approach where features affect the global shape while data points control its local shape.

Deformable templates, such as those employed by Yuille *et al.* [104] to extract facial features, mark a blend of parametric representation and physical-based methods. The template is described by a parametrized geometrical model. The goodness of fit between the deformable model and the image is measured by the interaction energy, which contains contributions from various image features. Optimal fit is obtained when the energy is minimized. In [36], the authors improved the original model by designing a more sophisticated cost functional.

Deformable superquadrics [98] are dynamic surface models with global and local deformation properties inherited from superquadric ellipsoids and membrane splines. The combined local/global representation is aimed at solving the conflicting goals of shape reconstruction and recognition we addressed earlier. Additional deformational degrees of freedom are gained from the incorporation of global deformation such as tapering, twisting, and bending [63]. By casting the fitting of time-varying visual data into the Lagrangian mechanical framework, the equations of motion governing the behavior of the deformable superquadrics can be developed. When augmented by Kalman

filter theory [64], the dynamic system becomes a recursive shape and motion estimator which employs the Lagrange equation of dynamic surfaces as a system model. In [65], deformable superquadrics that combine Kalman filter with additional constraints are employed to track articulated objects.

Inspired by modal analysis in linear mechanical systems, Pentland [74] developed a system that is capable of automatically recovering deformable part models based on the finite element method (FEM). Nonrigid object behavior is described by *modal dynamics*, i.e., by the superposition of its natural strain or vibration modes. By limiting the number of modes used in the representation, the analysis of nonrigid motion can always be transformed to an overconstrained problem. Later, the same model combined with an extended Kalman filter is applied to recover nonrigid motion and structure from contour [76] as well as optical flow data [75]. A major limitation of the modal framework is that objects must be described in term of the modes of some prototype shape. Such a procedure implicitly imposes an *a priori* parameterization upon the sensor data. It is thus more suitable for *modeling* than for *tracking* purposes. To address this problem, Sclaroff and Pentland [85] recently developed a new method that computes the object's vibration modes directly from the image data.

Nastar and Ayache [67] followed similar physics principles and developed elastic models for nonrigid motion tracking. The notable property of their model is that the governing dynamic equations are linear and decoupled for each coordinate, regardless of the amplitude of deformation. Algorithmic complexity is therefore significantly reduced. In [68], they have applied this improved model to study the temporal evolution of 3D nonrigid objects.

Perk *et al.* [73] have developed a volumetric model for analyzing 3D motion of the LV from MRI-SPAMM images. This new class of modeling primitive is physically motivated, and is capable of describing complex volumetric motion of the LV, such as the twisting deformation, using only a small number of *parameter functions*. The values of the parameter functions are estimated using the physics-based approach introduced in [65].

Shi *et al.* [88] proposed a model-based approach for the tracking myocardial deformation based on *continuum mechanics*. The model is embeded in a finite element framework, whose strain–stress relationship obeys the Hook's law. Velocity information (obtained within mid-well region of LV) and shape information (obtained on the inner and outer boundaries of LV) are integrated to provide a complete description of the elastic motion. A feedback mechanism is incorporated to improve the accuracy of motion estimation.

5. CONCLUSION

In this paper, we have reported the past developments on nonrigid motion analysis in two categories, *articulated* and *elastic* motion. This work can be generally categorized as (1) methods using no *a priori* shape models, and (2) model-based approaches. Motion tracking assuming no *a priori* knowledge about the motion or object shape is necessary when dealing with an unknown object. The difficulty of establishing feature correspondence is the major obstacle to this type of approach. Researchers either focus on high level processing, assuming that matching is known *a priori*, or impose constraints on the object's behavior to get around the problem. Model-based approaches, which have the advantage of knowing the approximate object shapes in advance, simplify this problem or transform it into other tractable issues. However, these methods are not applicable if knowledge of the object shape is not available.

Certain success has been achieved in the study of the human gait and the analysis of left ventricular motion using various image analysis techniques, as described in this paper. Overall, nonrigid motion analysis is still in its infancy. However, many investigators have realized the importance of dealing with nonrigidity in motion analysis. The applications of nonrigid motion analysis extend from teleconferencing, gesture recognition, face recognition, material deformation studies, biomedical applications, and geological formation studies to weather prediction and image compression. We expect novel schemes will be presented to deal with different types of nonrigid motion in the future.

REFERENCES

1. J. K. Aggarwal, Q. Cai, W. Liao, and B. Sabata, Articulated and elastic non-rigid motion: A review, in *Proc. of IEEE Computer Society Workshop on Motion of Non-Rigid and Articulated Objects, 1994*, pp. 16–22.
2. K. Akita, Image sequence analysis of real world human motion, *Pattern Recognit.* **17**(1), 1984, 73–83.
3. M. Allmen and C. R. Dyer, Cyclic motion detection using spatiotemporal surface and curves, in *Proc. ICPR, 1990*, pp. 365–370.
4. A. A. Amini and J. S. Duncan, Pointwise tracking of left-ventricular motion in 3D, in *Proc. CVPR, 1991*, pp. 294–299.
5. A. A. Amini and J. S. Duncan, Differential geometry for characterizing 3D shape change, in *Proc. SPIE Vol. 1768 Mathematical Methods in Medical Imaging, 1992*, pp. 170–181.
6. A. A. Amini, R. L. Owen, P. Anandan, and J. S. Duncan, Non-rigid motion models for tracking the left-ventricular wall, *Information Processing in Medical Images, 12th International Conference, 1991*, pp. 341–357.
7. M. Asada and S. Tsuji, Representation of three-dimensional motion in dynamic scenes, *Comput. Vision Graphics Image Process.* **21**, 1983, 118–114.
8. M. Asada, M. Yachida, and S. Tsuji, Analysis of three-dimensional motions in blocks world, *Pattern Recognit.* **17**(1), 1984, 57–71.
9. R. Bajcsy and S. Kovacic, Multiresolution elastic matching, *Comput. Vision Graphics Image Process.* **46**, 1989, 1–21.
10. K. Bartels, A. Bovik, S. J. Aggarwal, and K. R. Diller, Shape change analysis of confocal microscope images using variational techniques, in *Proc. SPIE Vol. 1660: Biomedical Image Processing and Three-Dimensional Microscopy, 1992*, pp. 618–629.
11. P. J. Besl, Geometric modeling and computer vision, in *Proc. IEEE* **76**(8), 1988, 936–958.
12. A. G. Bharatkumar, K. E. Daigle, M. G. Pandy, Q. Cai, and J. K. Aggarwal, Lower limb kinematics of human walking with the medial axis transformation, in *Proc. of IEEE Computer Society Workshop on Motion of Non-Rigid and Articulated Objects, 1994*, pp. 70–76.

13. A. F. Bobick and A. D. Wilson, A state-based technique for the summarization and recognition of gesture, in *Proc. of 5th Intl. Conf. on Computer Vision, 1995*, pp. 382–388.
14. F. L. Bookstein, Principal warps: Thin-plate splines and the decomposition of deformations, *IEEE Trans. PAMI* **11**, 1989, 567–585.
15. F. L. Bookstein, Thin-plate splines and the atlas problem for biomedical images, *Information Processing in Medical Images, 12th International Conference, 1991*, pp. 326–342.
16. F. L. Bookstein and W. D. K. Green, Edge information at landmarks in medical images, in *Proc. SPIE Vol. 1808: Visualization in Biomedical Computing, 1992*, pp. 242–258.
17. L. W. Campbell and A. F. Bobick, Recognition of human body motion using phase space constraints, in *Proc. of 5th Intl. Conf. on Computer Vision, 1995*, pp. 624–630.
18. S. Chaudhuri and S. Chatterjee, Estimation of motion parameters for a deformable object from range data, in *Proc. IEEE Conf. on CVPR, 1989*, pp. 291–295.
19. C. W. Chen and T. S. Huang, Surface modeling in heart motion analysis, in *Proc. SPIE Vol 1610: Curves and Surfaces in Computer Vision and Graphics, 1991*, pp. 360–371.
20. H. H. Chen and T. S. Huang, Maximal matching of 3-D points for multiple-object motion estimation, *Pattern Recognit.* **21**, 1988, 75–90.
21. Z. Chen and H. J. Lee, Knowledge-guided visual perception of 3D human gait from a single image sequence, *IEEE Trans. Systems Man Cybernetics* **22**(2), 1992, 336–342.
22. M. B. Clowes, On seeing things, *Artif. Intell.* **2**, 1971, 79–116.
23. I. Cohen, N. Ayache, and P. Sulger, Tracking points on deformable objects using curvature information, in *Proc. ICPR, 1992*, pp. 458–466.
24. I. Cohen, L. Cohen, and N. Ayache, Introducing new deformable to segment 3D images, in *Proc. CVPR, 1991*, pp. 738–739.
25. L. D. Cohen and I. Cohen, Deformable models for 3D medical images using finite elements and balloons, *Proceedings of ICPR, 1992*, pp. 592–598.
26. H. Delingette, M. Hebert, and K. Ikeuchi, Deformable surfaces: A free-form shape representation, in *Proc. SPIE Vol. 1570: Geometric Methods in Computer Vision, 1991*, pp. 21–30.
27. H. Delingette, M. Hebert, and K. Ikeuchi, Shape representation and image segmentation using deformable surfaces, in *Proc. CVPR, 1991*, pp. 467–472.
28. K. Deng and J. N. Wilson, Contour estimation using global shape constraints and local forces, in *Proc. SPIE Vol. 1570: Geometric Methods in Computer Vision, 1991*, pp. 227–233.
29. R. O. Duda and P. E. Hart, *Pattern Classification and Scene Analysis*, Wiley, New York, 1973.
30. D. Eggert, D. Goldgof, and K. W. Bowyer, Reconstructing CAD models of articulated objects, in *Second CAD-Based Vision Workshop, 1994*, pp. 98–105.
31. E. Fauvet, Human movement analysis with image processing in real time, in *Proc. Intl. Congress on High-Speed Photography and Photonics, 1990*, SPIE Vol. 1358, pp. 620–630.
32. R. Furukawa, M. Imai, and T. Uno, Robust algorithm for motion analysis based on active tubes, in *Proceedings of the Workshop on Motion of Non-rigid and Articulated Objects, 1994*, pp. 200–205.
33. N. H. Goddard, Incremental model-based discriminiation of articulated movement from motion features, in *Proc. of IEEE Computer Society Workshop on Motion of Non-Rigid and Articulated Objects, 1994*, pp. 89–95.
34. D. B. Goldgof, H. Lee, and T. S. Huang, Motion analysis of nonrigid surfaces, in *Proc. CVPR, 1988*, pp. 375–380.
35. L. Goncalves, E. D. Bernardo, E. Ursella, and P. Perona, Monocular tracking of the human arm in 3D, in *Proc. of 5th Intl. Conf. on Computer Vision, 1995*, pp. 764–770.
36. P. W. Hallinan, Recognizing human eyes, *SPIE Proc. Vol. 1570: Geometric Methods in Computer Vision, 1991*, pp. 214–226.
37. A. J. Hanson, Hyperquadrics: Smoothly deformable shapes with convex polyhedral bounds, *Comput. Vision Graphics Image Process.* **44**, 1988, 191–210.
38. Y. Hel-Or and M. Werman, Constraint-fusion for interpretation of articulated objects, in *Proc. CVPR, 1994*, pp. 39–45.
39. D. Hogg, Model-based vision: A program to see a walking person, in *Image and Vision Computing, 1983*, pp. 5–20.
40. B. K. P. Horn, *Robot Vision*, MIT Press, Cambridge, MA, 1986.
41. T. S. Huang, Modeling, analysis, and visualization of nonrigid object motion, in *Proceedings of 10th ICPR, 1990*, pp. 361–364.
42. D. Huffman, Impossible objects as nonsense sentences, *Machine Intelligence* (B. Meltzer and D. Michie, Eds.), Vol. 6, pp. 295–323, 1971.
43. D. P. Huttenlocher, J. J. Noh, and W. J. Rucklidge, Tracking non-rigid object in complex scenes, in *Proc. of 12th ICPR, 1993*, pp. 93–101.
44. K. Ikeuchi and B. K. P. Horn, Numerical shape from shading and occluding boundaries, in *Artificial Intelligence*, Vol. 17, pp. 141–184, 1981.
45. R. Jain and H. H. Nagel, On the analysis of accumulative difference pictures from image sequences of real world scenes, *IEEE Trans. PAMI* **1**(2), 1979, 206–214.
46. G. Johansson, Visual motion perception, *Sci. American* **232**(6), 1975, 76–88.
47. I. A. Kakadiaris and D. Metaxas, 3D human body model acquisition from multiple views, in *Proc. of 5th Intl. Conf. on Computer Vision, 1995*, pp. 618–623.
48. I. A. Kakadiaris, D. Metaxas, and R. Bajcsy, Active part-decomposition, shape and motion estimation of articulated objects: A physics-based approach, in *Proc. CVPR, 1994*, pp. 980–984.
49. C. Kambhamettu and D. B. Goldgof, Point correspondence recovery in non-rigid motion, in *Proc. CVPR, 1992*, pp. 222–227.
50. C. Kambhamettu, D. B. Goldgof, D. Terzopoulos, and T. S. Huang, Nonrigid motion analysis, *Handbook of PRIP: Computer Vision*, Vol. 2, 1994.
51. M. Kass, A.Witkin, and D. Terzopoulos, Snakes: Active contour models, *Int. J. Comput. Vision* **1**(4), 1987, 312–331.
52. W. Kinzel, Pedestrian recognition by modelling their shapes and movements, in progress in image analysis and processing III, in *Proc. of the 7th Intl. Conf. on Image Analysis and Processing 1993, Singapore, 1994*, pp. 547–554.
53. J. J. Koenderink and A. J. van Doorn, Dynamic shape, *Biol. Cybern.* **53**, 1986, 383–396.
54. D. L. Kraitchman, A. A. Young, C. N. Chang, and L. Axel, Semi-automatic tracking of myocardial motion in MR tagged images, *IEEE Trans. Med. Imaging* **14**(3), 1995, 422–433.
55. J. J. Kuch and T. S. Huang, Vision based hand modeling and tracking for virtual teleconferencing and telecollaboration, in *Proc. of 5th Intl. Conf. on Computer Vision, 1995*, pp. 666–671.
56. S. Kumar and D. Goldgof, Automatic tracking of SPAMM grid and the estimation of deformation parameters from cardiac mr images, *IEEE Trans. Med. Imaging* **13**(1), 1994, 122–133.
57. S. Kurakake and R. Nevatia, Description and tracking of moving articulated objects, in *11th Intl. Conf. on Pattern Recognition, 1992*, pp. 491–495.
58. M. K. Leung and Y. H. Yang, *An Empirical Approach to Human Body Motion Analysis*, Technical Report 94-1, University of Saskatchewan, Saskatchewan, Canada, 1994.
59. W. Liao, S. J. Aggarwal, and J. K. Aggarwal, Reconstruction of dynamic

3D structures of biological objects using stereo microscopy, in *Proc. of 1st Intl. Conf. on Image Processing, 1994*, Vol. III, pp. 731–735.

60. S. Lobregt and M. A. Viergever, A discrete dynamic contour model, *IEEE Trans. Med. Imaging* **14**(1), 1995, 12–24.
61. D. Marr and H. K. Nishihara, Representation and recognition of the spatial organization of three-dimensional shapes, *Proc. R. Soc. London, B*, 1978, 269–294.
62. T. McInerney and D. Terzopoulos, A finite element method for 3D shape reconstruction and non-rigid motion tracking, in *Proc. of 12th ICPR, 1993*, pp. 518–523.
63. D. Metaxas and D. Terzopoulos, Constrained deformable superquadrics and non-rigid motion tracking, in *Proc. CVPR, 1991*, pp. 337–343.
64. D. Metaxas and D. Terzopoulos, Recursive estimation of shape and nonrigid motion, *IEEE Workshop on Visual Motion, 1991*, pp. 306–311.
65. D. Metaxas and D. Terzopoulos, Constrained deformable superquadrics and non-rigid motion tracking, *IEEE Trans. PAMI* **15**(6), 1993, 580–591.
66. M. P. Murray, Walking patterns of normal men, *J. Bone and Joint Surgery* (46-A), 1964, 335–360.
67. C. Nastar and N. Ayache, A physically based analysis for deformation in 3D images, *Geometric Methods in Computer Vision II, SPIE* Vol. 2031, 1993, pp. 182–192.
68. C. Nastar and N. Ayache, Spatio-temporal analysis of nonrigid motion from 4D data, in *Proceedings of the Workshop on Motion of Non-rigid and Articulated Objects, 1994*, pp. 146–151.
69. S. A. Niyogi and E. H. Adelson, Analyzing and recognizing walking figures in XYT, in *Proc. CVPR, 1994*, pp. 469–474.
70. Y. Okawa and S. Hanatani, Recognition of human body motions by robots, in *Proc. IEEE/RSJ Intl. Conf. Intelligent Robots and Systems, Raleigh, NC, 1992*, pp. 2139–2146.
71. J. O'Rourke and N. I. Badler, Model-based image analysis of human motion using constraint propagation, *IEEE Trans. PAMI* **2**, 1980, 522–536.
72. C. H. Papadimitriou and K. Steiglitz, *Combinatorial Optimization: Algorithms and Complexity*, Prentice-Hall, Englewood Cliffs, NJ, 1982.
73. J. Park, D. Metaxas, and L. Axel, Volumetric deformable models with parameter functions: A new approach to the 3D motion analysis of the LV from MRI-SPAMM, in *Proc. of 5th Intl. Conf. on Computer Vision, 1995*, pp. 700–705.
74. A. Pentland, Automatic extraction of deformable part models, *Int. J. Comput. Vision* **4**, 1990, 107–126.
75. A. Pentland and B. Horwitz, Recovery of non-rigid motion and structures, *IEEE Trans. PAMI* **13**(7), 1991, 730–742.
76. A. Pentland, B. Horwitz, and S. Sclaroff, Non-rigid motion and structure from contour, in *Proc. CVPR, 1991*, pp. 288–293.
77. F. J. Perales and J. Torres, A system for human motion matching between synthetic and real image based on a biomechanic graphical model, in *Proc. of IEEE Computer Society Workshop on Motion of Non-Rigid and Articulated Objects, 1994*, pp. 83–88.
78. R. Polana and R. Nelson, Low level recognition of human motion (or how to get your man without finding his body parts, in *Proc. of IEEE Computer Society Workshop on Motion of Non-Rigid and Articulated Objects, 1994*, pp. 77–82.
79. J. Ponce, Straight homogeneous generalized cylinders: Differential geometry and uniquness results, *Int. J. Comput. Vision* **4**(1), 1990, 79–100.
80. R. J. Qian and T. S. Huang, Estimating articulated motion by decomposition, in *Time-Varying Image Processing and Moving Object Recognition* (V. Cappellini, Ed.), Vol. 3, pp. 275–286, 1994.
81. S. Ranganath, Contour extraction from cardiac MRI studies using snakes, *IEEE Trans. Med. Imaging* **14**(2), 1995, 328–338.
82. R. F. Rashid, Towards a system for the interpretation of moving light display, *IEEE Trans. PAMI* **2**(6), 1980, 574–581.
83. J. M. Rehg and T. Kanade, Model-based tracking of self-occluding articulated objects, in *Proc. of 5th Intl. Conf. on Computer Vision, 1995*, pp. 621–617.
84. K. Rohr, Towards model-based recognition of human movements in image sequences, *CVGIP: Image Understanding* **59**(1), 1994, 94–115.
85. S. Sclaroff and A. Pentland, Modal matching for correspondence and recognition, *IEEE Trans. Pattern Anal. Machine Intell.* **17**(6), 1995, 545–561.
86. B. Serra and M. Berthod, Optimal subpixel matching of contour chains and segments, in *Proc. of 5th Intl. Conf. on Computer Vision, 1995*, pp. 402–407.
87. L. S. Shapiro, H. Wang, and J. M. Brady, A matching and tracking strategy for independently moving objects, in *Proc. British Machine Vision Conference, 1992*, pp. 306–315.
88. P. Shi, G. Robinson, R. T. Constable, A. Sinusas, and J. Duncan, A model-based integrated approach to track myocardial deformation using displacement and velocity constraints, in *Proc. of 5th Intl. Conf. on Computer Vision, 1995*, pp. 687–692.
89. F. Solina and R. Bajcsy, Recovery of parametric models from range images: The case for superquadrics with global deformations, *IEEE Trans. PAMI* **12**(2), 1990, 131–147.
90. L. H. Staib and J. S. Duncan, Boundary finding with parametrically deformable models, *IEEE Trans. PAMI* **14**(11), 1992, 1061–1075.
91. L. H. Staib and J. S. Duncan, Deformable fourier models for surface finding in 3D images, in *Proc. SPIE Vol. 1808: Visualization in Biomedical Computing, 1992*, pp. 90–104.
92. R. N. Strickland and Z. Mao, Feature matching in image sequences, in *Proc. 6th Scandinavian Conf. on Image Analysis, 1989*, pp. 729–732.
93. R. N. Strickland and Z. Mao, Computing correspondences in a sequence of non-rigid shapes, *Pattern Recognit.* **25**(9), 1992, 901–912.
94. S. Sullivan, L. Sandford, and J. Ponce, Using geometric distance fits for 3-D object modeling and recognition, *IEEE Trans. PAMI* **16**(12), 1994, 1183–1196.
95. R. Szeliski and D. Terzopoulos, Physically-based and probabilistic models for computer vision, in *Proc. SPIE Vol. 1570: Geometric Methods in Computer Vision, 1991*, pp. 140–152.
96. D. Terzopoulos, Regularization of inverse visual problems involving discontinuities, *IEEE Trans. PAMI* **8**, 1986, 413–424.
97. D. Terzopoulos, On matching deformable models to images: Direct and iterative solutions, in *Topical Meeting on Machine Vision, Technical Digest Series*, Vol. 12, pp. 160–167, 1987.
98. D. Terzopoulos and D. Metaxas, Dynamic 3D models with local and global deformations: Deformable superquadrics, *IEEE Trans. PAMI* **13**(7), 1991, 703–714.
99. D. Waltz, The psychology of computer vision, in *Understanding Line Drawings of Scenes with Shadows*, pp. 19–95, McGraw-Hill, New York, 1971.
100. J. A. Webb and J. K. Aggarwal, Structure from motion of rigid and jointed objects, in *Artificial Intelligence, Vol. 19, 1982*, pp. 107–130.
101. J. Yamato, J. Ohya, and K. Ishii, Recognizing human action in time-sequential images using hidden markov model, in *Proc. IEEE Conf. CVPR, June 1992*, pp. 379–385.
102. A. A. Young, D. L. Kraitchman, L. Dougherty, and L. Axel, Tracking and finite element analysis of stripe deformation in magnetic resonance tagging, *IEEE Trans. Med. Imaging* **14**(3), 1995, 413–421.
103. A. A. Young, D. L. Kraitchman, and L. Axel, Deformable models for tagged MR images: Reconstruction of two and three-dimensional heart wall motion, in *Proceedings of the IEEE Workshop on Biomedical Image Analysis, 1994*, pp. 317–323.
104. A. L. Yuille, D. S. Cohen, and P. W. Hallinan, Feature extraction from faces using deformable templates, in *Proc. IEEE Conf. CVPR, 1989*, pp. 104–109.