# Modeling of Elastic Articulated Objects and Its Parameters Determination from Image Contours

Hailang Pan, Yuncai Liu, and Lei Shi

Institute of Image Processing and Pattern Recognition,
Shanghai Jiao Tong University, P.R.China, 200030
{panhailang, whomliu, sl0030322014}@sjtu.edu.cn

**Abstract.** This paper presents a new method of elastic articulated objects (human bodies) modeling based on a new conic curve. The model includes 3D object deformable curves which can represent the deformation of human occluding contours. The deformation of human occluding contour can be represented by adjusting only four deformation parameters for each limb. Then, the 3D deformation parameters are determined by corresponding 2D contours from a sequences of stereo images. The algorithm presented in this paper includes deformable conic curve parameters determination and the plane, 3D conic curve lying on, parameter determination.

## 1 Introduction

Objects in motion analysis are mainly rigid objects, articulated objects, nonrigid objects and so on. In computer vision research, motion analysis has been largely restricted to rigid objects. However, in the real world, motion of nonrigid objects is the rule[1]. In the past decade, there has been a growing interest in the study of nonrigid motion. In nonrigid motion analysis, dynamic shape modeling provides a mechanism for fitting and tracking visual data. Using deformable models, the seemingly unstructured elastic motion can be compactly represented by a small number of parameters. The task of elastic motion recovery is reduced to the problem of the parameters determination.

In the past researches, a number of method for human deformable body modeling have been proposed. Nahas [2] use B-Spline surfaces to model lissom movements of body and face. Petland [3] introduced a model of elastic nonrigid motion. This model is based on the finite element method. Min [4] proposed a layered structure of the human body to provide a realistic deformation scheme for the human body model. Plankers[5]developed a framework for 3D shape and motion recovery of articulated deformable objects. Smooth implicit surfaces, known as metaballs, are attached to an articulated skeleton of the human body and are arranged in an anatomically-based approximation. Sminchisescu [6] built a human body model which consists of kinematic 'skeletons', covered by 'flesh' built from superquadric ellipsoids. A typical model has around 9 deformable shape parameters for each body part. Apuzzo[7] presents simplified Model of a Limb. Only three ellipsoidal metaballs are attached to each limb skeleton and arranged in an anatomically-based approximation. Each limb has twelve deformation parameters.

Elastic articulated objects are the combination of articulated objects and nonrigid objects. Human body is a typical elastic articulated objects. Our research focuses on two major points: 3D human body modeling and its parameters determination from stereo image contours. The proposed human limb model is composed of two layers: a skeleton layer and a body occluding contour layer. The skeleton layer represents the skeletal structure of the human body, which is composed of the sticks and joints which linking these sticks. The body occluding contour layer is expressed by 3D conic curves, attached to an articulated skeleton and set in an contour-based approximation. The body occluding contour layer deforms with the motion of the skeleton. We can concisely deform the body occluding contour layer during animation by adjusting only four deformation parameters for each human limb. The human arm model is equally applicable to other vertebrates, such as horses, cows and so on. There are two aspects of image-based 3D human arm parameters determination. The first is the skeleton motion parameters estimation. The second is the deformable occluding contours deformation parameters determination. We establish the equations of human arm's movement and deformation, analyze the condition for a solution and solve the nonlinear equations.

One motivation of our research is to build a body model that properly describes human body deformation from a small number of parameters and human 3D shape and motion analysis from 2D image sequence.

## 2  Human Arm Modeling

### 2.1  Establishment of Conic Curves

This algorithm uses a inducting form to express a 2D conic curve restricted by three straight lines [8](see Fig.1).
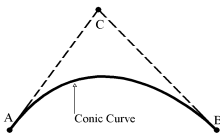


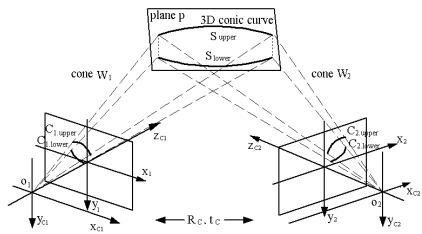**Fig. 1.** 2D conic curve          **Fig. 2.** Coordinate systems and two images of a 3D conic curve

Inducting form of the 2D conic curve is a family of conic curves, which pass through vertices A , B and tangent to the straight lines AC, BC. They have the equation of the form

$$(a_0 + a_1 x + a_2 y)(b_0 + b_1 x + b_2 y) = \rho(u_0 + u_1 x + u_2 y)^2 \tag{1}$$

We can deform the conic curve by adjusting the parameter $\rho$ .The equation (1) can also be represented by the quadratic equation (3).

## 2.2  Conics-Based Stereo

The geometry of the sensors is shown in Fig.2. In stereo vision, we have two coordinate systems $c_1$ and $c_2$ associated with two cameras. The relative position and orientation of $c_1$ and $c_2$ are described by a rotation $R_c$ followed by a translation $t_c$ :

$$X_{c2} = R_c X_{c1} + t_c \tag{2}$$

where $X_{c1}$ and $X_{c2}$ are the coordinates of any point in space respectively by $c_1$ and $c_2$ frames: $X_{c1}= [x_{c1}, y_{c1}, z_{c1}]^T$ and $X_{c2}=[x_{c2}, y_{c2}, z_{c2}]^T$.

The optical center $o_1$ (or $o_2$) is the origin of the camera frame $c_1$ (or $c_2$). The z axis, $o_1z_{c1}$(or $o_2z_{c2}$), is the optical axis and the image plane, $x_1y_1$ plane (or $x_2y_2$ plane), is parallel to the $x_{c1} y_{c1}$ plane (or $x_{c2} y_{c2}$ plane) at a distance $f$ from the origin. We suppose the cameras' intrinsic parameters are known and the pixel coordinates have been transformed to $x_1 y_1$ (or $x_2 y_2$) coordinates.

Suppose there is a 3D conic curve S in space. Its two projections on two images are represented by two quadratic form $C_1$ and $C_2$ [9] :

$$C_i(x_i, y_i) = X_i^T Q_i X_i = 0 \quad i = 1,2 \tag{3}$$

where $X_i=[x_i,y_i,1]^T$.$(x_i,y_i)$ are the 2D coordinates in two images.Two 2D conic curves, projected from the 3D conic curve S and observed by two cameras,are extracted and represented by the matrixes $Q_1$ and $Q_2$ respectively

$$Q = \begin{bmatrix} (a_{i1}b_{i1} - \rho_i \cdot u_{i1}^2) & (a_{i1}b_{i2}+a_{i2}b_{i1}-2\rho_i \cdot u_{i1}u_{i2})/2 & (a_{i0}b_{i1}+a_{i1}b_{i0}-2\rho_i \cdot u_{i0}u_{i1})/2 \\ (a_{i1}b_{i2}+a_{i2}b_{i1}-2\rho_i \cdot u_{i1}u_{i2})/2 & (a_{i2}b_{i2} - \rho_i \cdot u_{i2}^2) & (a_{i0}b_{i2}+a_{i2}b_{i0}-2\rho_i \cdot u_{i0}u_{i2})/2 \\ (a_{i0}b_{i1}+a_{i1}b_{i0}-2\rho_i \cdot u_{i0}u_{i1})/2 & (a_{i0}b_{i2}+a_{i2}b_{i0}-2\rho_i \cdot u_{i0}u_{i2})/2 & (a_{i0}b_{i0} - \rho_i \cdot u_{i0}^2) \end{bmatrix} \quad i = 1,2 \tag{4}$$

Let $X_w=[x_w,y_w,z_w,1]^T$. $(x_w,y_w,z_w)$ is a point in the world coordinate system.Then $X_i$ and $X_w$ are related to each other through the following equation:

$$z_{ci}X_i = z_{ci}\begin{bmatrix} x_i \\ y_i \\ 1 \end{bmatrix} = \begin{bmatrix} f & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \cdot \begin{bmatrix} x_{ci} \\ y_{ci} \\ z_{ci} \\ 1 \end{bmatrix} = \begin{bmatrix} f & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \cdot \begin{bmatrix} R_i & t_i \\ 0^T & 1 \end{bmatrix} \cdot \begin{bmatrix} x_w \\ y_w \\ z_w \\ 1 \end{bmatrix} = M_i X_w \tag{5}$$

Where the matrixes $(R_i , t_i)$, called the extrinsic parameters, is the rotation and translation which relates the world coordinate system to each camera coordinate system. $f$ is the focal length of the cameras.

Substituting the equation (5) into the equation (3), we obtain the representations of the two cones $W_1$ and $W_2$ passing through the 2D conic curve $C_1$(or $C_2$) and the camera center $o_1$(or $o_2$):

$$\begin{cases} W_1(x_w, y_w, z_w) = X_w^T M_1^T Q_1 M_1 X_w = 0 \\ W_2(x_w, y_w, z_w) = X_w^T M_2^T Q_2 M_2 X_w = 0 \end{cases} \tag{6}$$

The 3D conic curve S in space is the intersection curve of two cones $W_1$ and $W_2$.

## 2.3  Establishment of 3D Conic Curve

A 3D conic curve can be obtained by intersecting two cones. The intersection of two cones $W_1$ and $W_2$ of degree 2 is a planar curve lying on the plane p: $rx_1+ sy_1+ z_1+ t=0$, then there exist a scalar factor k and a polynomial $f_c$ of degree1 such that,

$$W_1 - kW_2 = ( rx_1 + sy_1 + z_1 + t) \, f_c \qquad (7)$$

According to this proposition, the polynomial $W_1 - kW_2$ has a linear factor $rx_1 + sy_1 + z_1 + t$. Now the problem of reconstructing the plane p becomes a multivariable polynomial factorization problem. In addition, k in the polynomial $W_1 - kW_2$ is unknown. The representations of two cones $W_1$ and $W_2$ are

$$W_1 = \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix}^T \begin{bmatrix} a_{11} & a_{12} & a_{13} & 0 \\ a_{12} & a_{22} & a_{23} & 0 \\ a_{13} & a_{23} & a_{33} & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} = 0 \quad W_2 = \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix}^T \begin{bmatrix} b_{11} & b_{12} & b_{13} & b_{14} \\ b_{12} & b_{22} & b_{23} & b_{24} \\ b_{13} & b_{23} & b_{33} & b_{34} \\ b_{14} & b_{24} & b_{34} & b_{44} \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} = 0$$

The six equations provided by their resultant are

$$(a_{33} - kb_{33})r^2 - 2(a_{13} - kb_{13})r + a_{11} - kb_{11} = 0 \qquad (8)$$

$$(a_{33} - kb_{33})s^2 - 2(a_{23} - kb_{23})s + a_{22} - kb_{22} = 0 \qquad (9)$$

$$(a_{33} - kb_{33})t^2 - (a_{33} - kb_{33} - 2kb_{34})t - kb_{44} = 0 \qquad (10)$$

$$2(a_{33} - kb_{33})tr - 2(a_{13} - kb_{13})t - (a_{33} - kb_{33} - 2kb_{34})r - 2kb_{14} = 0 \qquad (11)$$

$$(a_{33} - kb_{33})rs - 2(a_{23} - kb_{23})r - (a_{13} - kb_{13})s + a_{12} - kb_{12} = 0 \qquad (12)$$

$$2(a_{33} - kb_{33})st - 2(a_{23} - kb_{23})t - (a_{33} - kb_{33} - 2kb_{34})s - 2kb_{24} = 0 \qquad (13)$$

The variables r, s, and t are well separated in the first three equations. If k can be firstly solved for, then r, s, and t can be independently solved for by these one-variable quadratic equations. Since $W_1 - kW_2$ is a quadratic polynomial and is reducible, it is well known that one of its invariants, D, defined by

$$D = |A - kB| = \begin{vmatrix} \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{12} & a_{22} & a_{23} \\ a_{13} & a_{23} & a_{33} \end{bmatrix} - k \begin{bmatrix} b_{11} & b_{12} & b_{13} \\ b_{12} & b_{22} & b_{23} \\ b_{13} & b_{23} & b_{33} \end{bmatrix} \end{vmatrix} = 0 \qquad (14)$$

should be equal to zero. Therefore k is the eigenvalue of the matrix $B^{-1}A$. After solving for k, we can solve for r, s, and t independently by equations (8) to (10). Since each equation has two solutions, we have eight solutions of (r, s, t) but some of them are not consistent with equations (11) to (13) and can be discarded. It is evident that the explicit parameters of the conic curve in space are completely defined by the cone $W_1$ and the plane p. After solving for p, we can obtain these parameters by simple algebraic operations.

Consequently, a 3D conic curve can be specified by the cone $W_1$ and the plane p, which can represent deformable human body occluding contours.
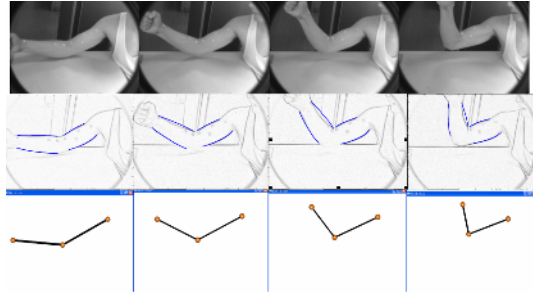
## 2.4   Establishment of Human Skeleton

The skeleton is a stick-model that represents the pose of the person in the image and makes it possible to segment the person into different parts. The motion of joints provides the key to motion estimation and recognition of the whole skeleton. The human skeleton system is treated as a series of jointed links(segments), which can be

modeled as a articulated body. For the specific problem of recovering the motion of a human figure, we describe the body as a stick model consisting of a set of fifteen joints (plus the head) connected by fourteen body segments [10], as shown in Figure 3(a). A closer look at the right arm is shown in Figure 3(b).
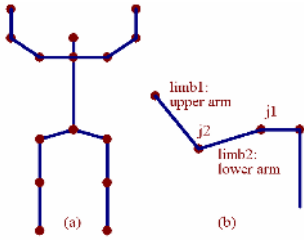


**Fig. 3.** a) The human skeleton     **Fig. 4.** The movement of human arm
b) A closer look at the right arm

## 3   Human Arm Parameters Determination

### 3.1   Deformation Parameters Determination of Human Body Occluding Contours

The image of a solid bounded by a smooth surface is itself bounded by an image curve, called the contour, silhouette or outline of this solid. This curve is the intersection of the retina with a viewing cone whose apex coincides with the pinhole and whose surface grazes the object along a second curve, called the occluding contour, or rim (see Fig.2).

We employ the gradient weighted least-squares estimation to solve the conic curve fitting[11].This method has the very attractive advantage that the eigenvector solution assures the global minimisation in case of convergence. We express deformation of human body occluding contours as the 3D conic curves of Section2.3. We fit one 2D conic curve to a set of points $(x_i, y_i)$, lying on the contour of each limb in left or right images. A 2D conic curve can be expressed in the implicit form:

$$C(x,y) = (a_0 + a_1 x + a_2 y)(b_0 + b_1 x + b_2 y) - \rho(u_0 + u_1 x + u_2 y)^2 = 0 \qquad (15)$$

Supposing $\{(x_i, y_i), i \in 1, \cdots, n\}$ to be the contour points observed, the squared distance of a point $(x_i, y_i)$ to the conic curve can be approximated by

$$dist\{(x_i, y_i), C(x, y)\}^2 \approx \frac{f^2(x_i, y_i)}{\|\nabla f(x_i, y_i)\|^2} \qquad (16)$$

The gradient weighted least-squares criterion for fitting the function to the data set $\{(x_i, y_i)\}$ minimizes the mean square distance, denoted by the cost function

$$\Theta = \frac{1}{n} \sum_{i=1}^{n} dist\{(x_i, y_i), C(x, y)\}^2 = \sum_{i=1}^{n} w_i C^2(x_i, y_i) \qquad (17)$$

where $w_i = 1/\|\nabla C(x_i, y_i)\|^2$ is the "gradient weight".

We minimize the function $\Theta$ to determine the best values of the deformation parameter $\rho$ that describes the model's deformation. That is, we wish to determine the values of the deformation parameter that will make the conic curve pass through the given points data set with minimum error.

We use a stereoscopic vision system. There are two images of each stereo image pair, left and right images. Hence we can get four deformation parameters $\rho_{i \cdot left.upper}$, $\rho_{i \cdot left.lower}$ and $\rho_{i \cdot right.upper}$, $\rho_{i \cdot right.lower}$ of each limb at a time, that can represent the deformation of human arm properly.

### 3.2  Motion Parameters Estimation of Human Skeleton

We pick up the centers of the circles that pass through the vertices of the conic curve and are tangent to the boundary of each limb. We regard the points, lying on the sticks that linking these centers, as the skeleton points. It is well known that rigid motion can be expressed as a rotation R around the joint of the 3D coordinates followed by a translation T [12]. See Fig.3, the motion of the limb $l_i$ is expressed as rotating around the joint $j_i$, then translating. $j_i'$ is the joint after translation. $R_i$ is the rotation matrix of the limb $l_i$ in the 3D world coordinate system. $p_k \leftrightarrow p_k'$ are the correspondent 3D skeleton points on the limb $l_i$. Now, we can estimate the joints $j_i$, $j_i'$ and the rotation matrix $R_i$ over two image views by solving the following system of equations

$$p_k' = R_i(p_k - j_i) + j_i' \qquad j_i' = j_i + T_i \tag{18}$$

where $k = 1, \cdots, m$, $m$ is the number of skeleton points on the limb $l_i$. The motion of the limb $l_i$ can be estimated with only three skeleton points correspondences. We can solve the equation (18) with three skeleton points correspondences.

## 4   Experiment Results

The controlled scenes are acquired by using a stereoscopic vision system. The images used in the experiment are of an indoor scene containing a posing human arm(See Fig.4). The image sequence contains 10 stereo image pairs. However, only four pairs image frames from two cameras are used in this experiment since they show the deformation of body occluding contours typically. The human arm is nicely distinguishable against the simpler background using a combination of edge detection, greyscale, and intensity adjustment. The 2D coordinates of a set of points, lying on the contour of each limb and acquired from the image planes, are used as the input data for determining the deformation parameters of the limbs occluding contours. The gradient weighted least-squares estimation is employed to solve the conic curve fitting. the parameters of the plane, 3D conic curve lying on, are determinated by the method of Section 2.3.

The estimated results are listed in Table 1, Table 2 and Table 3 that describes the model's deformation and motion. The determined deformation parameters from left camera are given in Table 1.The determined plane parameters are given in Table 2. The estimated motion parameters are given in Table 3.

**Table 1.** Estimated deformation parameters

| Deformation parameter | Frame1 | Frame2 | Frame3 | Frame4 |
|---|---|---|---|---|
| $\rho_{1.left.upper}$ | 16143 | 14389 | 846.6720 | 21.9888 |
| $\rho_{1.left.lower}$ | 1.3200 | 1.2825 | 0.7877 | -0.0616 |
| $\rho_{2.left.upper}$ | -8305.8 | 378.4237 | 11.9258 | -0.4749 |
| $\rho_{2.left.lower}$ | 109.2571 | 56.3748 | 2772.9 | 413.6866 |

$\rho_{1.left.upper}$ is the deformation parameter of the upper arm upper contour, $\rho_{1.left.lower}$ is the deformation parameter of the upper arm lower contour, $\rho_{2.left.upper}$ is the deformation parameter of the lower arm upper contour, $\rho_{2.left.lower}$ is the deformation parameter of the lower arm lower contour.

**Table 2.** Estimated parameters of plane p

|  | upper arm ( r, s, t ) | lower arm ( r, s, t ) |
|---|---|---|
| Frame1 | (4.9723,8.5553,-412.8744) | (-1.5019,28.8184,166.2452) |
| Frame2 | (4.2509,7.9248,-386.5332) | (0.8708,-3.8621,-1011.6) |
| Frame3 | (3.6959,7.7719,-302.3469) | (0.2075,-1.7083,-936.8942) |
| Frame4 | (2.6618,6.5545,-268.4768) | (-0.8521, -1.3332,-953.3770) |

**Table 3.** Estimated motion parameters

(a)  Upper arm (UA)

| Time interval | Rotation axis $\vec{n}_1$ | | | Rotation Angle $\theta_1$ | Translation $\overline{T}_1$ | | |
|---|---|---|---|---|---|---|---|
|  | $n_1$ | $n_2$ | $n_3$ |  | X | Y | Z |
| 1—2 | -0.6277 | -0.7728 | 0.0938 | -0.1852 | 7.5911 | 1.5811 | -0.0837 |
| 2—3 | -0.6462 | -0.7576 | 0.0919 | -0.3427 | 2.6562 | -0.7987 | 2.5937 |
| 3—4 | -0.9358 | -0.3509 | 0.0353 | -1.1869 | -1.1896 | -4.2581 | 9.1290 |

(b)  Lower arm (LA)

| Time interval | Rotation axis $\vec{n}_2$ | | | Rotation Angle $\theta_2$ | Translation $\overline{T}_2$ | | |
|---|---|---|---|---|---|---|---|
|  | $n_1$ | $n_2$ | $n_3$ |  | X | Y | Z |
| 1—2 | -0.6419 | 0.4326 | -.9947 | 0.1403 | 0.9377 | 13.5415 | 29.9642 |
| 2—3 | 0.8088 | -0.5181 | -.2782 | -0.5021 | -4.6424 | 18.0008 | 64.0877 |
| 3—4 | 0.6671 | 0.5096 | -.5434 | -0.7036 | -7.8553 | 24.9998 | 236.7208 |

$\vec{n}_1, \theta_1$ and $T_1$ are the rotation axis, rotation angle and translation of the upper arm, $\vec{n}_2, \theta_2$ and $T_2$ are the rotation axis, rotation angle and translation of the lower arm.

Fig.4 depicts the movement of human arm. The first row are the original images of bare arm from left camera. The second row are the corresponding target contours and conic curve fitting results. The third row are the estimated 3D skeleton model.

## 5  Conclusion

We have presented a new method of elastic articulated objects (human bodies) modeling based on a new 3D conic curve. Our research focuses on two major points: 3D human body modeling and its parameters determination.Our experiments have demonstrated that our model can express the deformation of human body occluding contour properly. As future work we want to model other parts of human body, such as upper leg, lower leg and torso, from image sequence and determine their deformation parameters in 3D.

## References

1. J. K. Aggarwal, Q. Cai and W. Liao.Articulated and elastic non-rigid motion: a review.Proc. IEEE Workshop On Motion of Non-rigid and Articulated Objects, 1994, pp. 2-14.
2. Monique Nahas,Herve Huitric,and Michel Saintourens.Animation of a B-Spline Figure. Visual Computer,1988.
3. Alex Petland and Bradley Horowitz.Recovery of Nonrigid motion and structure.IEEE Transaction on PAMI 1991,13(7):730~742.
4. Kyung-Ha Min, Seung-Min Baek, Gun A. Lee, Haeock Choi, and Chan-Mo Park. Anatomically-based modeling and animation of human upper limbs. in Proceedings of International Conference on Human Modeling and Animation, 2000.
5. Ralf Plankers and Pascal Fua. Articulated Soft Objects for Multiview Shape and Motion Capture. IEEE Transaction on PAMI, 2003, 25 (9): 1182 ~ 1187.
6. ESTIMATION ALGORITHMS FOR AMBIGUOUS VISUAL MODELS (C.Sminchisescu), Doctoral Thesis, INRIA, July 2002.
7. D'Apuzzo N, Plänkers R, Gruen A, Fua F and Thalmann D, Modelling Human Bodies from Video Sequences, Proc. Electronic Imaging1999, San Jose, California, January.
8. R.C.Jain, R.Kasturi and B.Schunk, Machine Vision, McGraw-Hill Inc, New York, 1995.
9. Songde MA.Conics-Based Stereo, Motion Estimation, and Pose Determination. Intern. J. Computer Vision, Vol. 10, No.1, 1993.
10. Xiaoyun Zhang, Yuncai Liu and TS Huang. Articulated Joint Estimation from Motion Using Two Monocular Images. Pattern Recognition Letters 25(10): 1097-1106,2004.
11. Z. Zhang. Parameter Estimation Techniques: A Tutorial with Application to Conic Fitting. Image and Vision Computing Journal, 1996.
12. Xiaoyun Zhang, Yuncai Liu and TS Huang. Motion Estimation of Articulated Objects from Perspective Views. Lecture Notes in Computer Science, Vol. 2492, pp.165-176, 2002.