STRUCTURE AND MOTION ESTIMATION FROM DYNAMIC SILHOUETTES UNDER PERSPECTIVE PROJECTION *

Tanuja Joshi Narendra Ahuja Jean Ponce
Beckman Institute, University of Illinois, Urbana, Illinois 61801

Abstract: *We address the problem of estimating the structure and motion of a smooth curved object from its silhouettes observed over time by a trinocular stereo rig under perspective projection. We first construct a model for the local structure along the silhouette for each frame in the temporal sequence. Successive local models are then integrated into a global surface description by estimating the motion between successive time instants. The algorithm tracks certain surface features (parabolic points) and image features (silhouette inflections and frontier points) which are used to bootstrap the motion estimation process. The entire silhouette along with the reconstructed local structure are then used to refine the initial motion estimate. We have implemented the proposed approach and report results on real images.*

1 Introduction

An object's silhouette and its deformations as the object moves with respect to the camera reflect its structure and motion characteristics. For objects with little or no surface detail such as texture, silhouettes serve as the most important cue for the estimation of object structure and motion.

Several methods have been proposed for structure estimation from silhouettes under *known* camera motion [1, 2, 5, 11, 12, 13, 14]. These approaches have demonstrated that given a set of three or more nearby views of a smooth object, the structure of the object up to second order can be obtained along its silhouette. The recovery of structure *and* motion from a monocular sequence of silhouettes has been investigated by Giblin et al. [4]. For the case of a curved object rotating about a fixed axis with constant angular velocity, they have shown that: (1) given a complete set of orthographic silhouettes, the rotation axis and velocity can be recovered, along with the visible surface; and (2) given the silhouettes over a short time interval, the rotation axis can be recovered if the angular velocity is known.

We addressed the problem of estimating structure and motion of a smooth object undergoing arbitrary unknown motion under orthographic projection in [6]. The structure and motion was estimated by tracking the silhouettes observed by a trinocular stereo rig over time. In this paper, we extend the approach to the case of perspective projection. The resulting structure and motion estimation algorithm will be useful in a situation where the viewer has no knowledge or control of the object's motion.

1.1 Problem Statement and Approach

Consider a smooth curved object in motion viewed by an observer. The viewing cone grazes the object surface along the *occluding contour* and intersects the image plane along the *silhouette*. At each point on the occluding contour the surface normal is orthogonal to the viewing ray. This makes the 3D occluding contour and the 2D silhouette viewpoint-dependent. With the object moving relative to the viewer, the silhouettes in successive images are projections of different 3D contours. This is in contrast with an image sequence containing only viewpoint-independent features.

Here, we address the problem of estimating structure and motion of a smooth curved object from its silhouettes observed over time by a trinocular stereo rig under perspective projection. To relate the silhouettes observed at successive time instants, we construct a model for the local structure along the silhouette at each time instant. We use trinocular imagery for our analysis since three images taken from known relative positions are sufficient to recover the local structure (up to second order) along the silhouette [1, 2, 12, 13].

Since successive silhouettes are projections of different 3D contours, there is no true point-to-point correspondence between them. For the triplet of images observed at a given time by the three cameras, the epipolar geometry is known and we establish correspondences between distinct 3D points lying on a common epipolar curve and then estimate local structure parameters.

However for the images taken by a camera at successive time instants, the epipolar geometry is unknown. Unless we know the correspondences we cannot estimate the motion. To bootstrap the process of motion estimation, we use some detectable and trackable silhouette features *(inflections* and *frontier points* [4]) to obtain an initial estimate of the unknown motion, which is then refined using the rest of the silhouette.

*This work was supported in part by the ARPA Grant N00014-93-1-1167 and by the NSF Grant IRI-9224815. J. Ponce was supported in part by the Center for Advanced Study of the University of Illinois. We thank D. Kriegman and B. Vijayakumar for providing the images.

0-8186-7042-8/95 $4.00 © 1995 IEEE

The algorithm for structure estimation using trinocular imagery is described in Sect. 2. The algorithm for motion estimation from dynamic silhouettes is discussed in Sect. 3. We present experimental results on real images in both sections, and conclude with comments in Sect. 4.

2 Structure Estimation Using Trinocular Imagery

2.1 Modeling the Local Structure

The local structure (up to second order) at a surface point P is defined by the 3D location of P in the world coordinate frame (in our case, the coordinate frame of the central camera), the surface normal at P, the two principal directions and the principal curvatures. At each point P, we define a local coordinate frame (X_l, Y_l, Z_l) whose origin is at P, the X_l-axis is aligned with the outward normal, and the Y_l and Z_l-axes are aligned with the principal directions. The local surface up to second order is a paraboloid [3], given by the equation

$$X_l = \frac{1}{2}\kappa_1 Y_l^2 + \frac{1}{2}\kappa_2 Z_l^2 \quad or \quad \boldsymbol{Q_l}^T \boldsymbol{M_l} \boldsymbol{Q_l} = 0, \qquad (1)$$

where κ_1 and κ_2 are the principal curvatures at P, $\boldsymbol{M_l}$ is a symmetric 4×4 matrix and $\boldsymbol{Q_l}$ is the vector of homogeneous coordinates of a point Q on the paraboloid at P.[1] The signs of κ_1 and κ_2 define the point type: if both are same (resp. opposite), P is an elliptic (resp. hyperbolic) point; if either κ_1 or κ_2 is zero, P is a parabolic point; if both κ_1 and κ_2 are zero, P is a planar point.

Let (X, Y, Z) be the camera-centered coordinate frame, where the Z-axis coincides with the optical axis and the XY-plane is the image plane (Fig. 1). Let P, the origin of the local frame, be at (X_0, Y_0, Z_0) in the camera-centered frame. We denote the angle made by the X_l-axis with the normal to the silhouette by ζ, the angle between the X-axis and the normal to the silhouette by θ and the angle between the viewing direction and one of the principal directions (say the Z_l-axis) by γ.

To completely describe the object surface locally, we need to specify the rigid transformation parameters $(\theta, \gamma, \zeta, X_0, Y_0, Z_0)$ defining the local coordinate frame at P with respect to the world coordinate frame, together with κ_1 and κ_2, for each point P on the silhouette. The equation of the paraboloid in the camera-centered frame is:

$$\Sigma = \boldsymbol{Q}^T \boldsymbol{M} \boldsymbol{Q} = 0, \qquad (2)$$

where $\boldsymbol{M} = \boldsymbol{T_0^{-1}}^T \boldsymbol{R_0}\, \boldsymbol{M_l}\, \boldsymbol{R_0^{-1}}\, \boldsymbol{T_0^{-1}}$. Here $\boldsymbol{T_0}$ is the 4×4 matrix for a translation by (X_0, Y_0, Z_0), $\boldsymbol{R_0}$ is the 4×4 matrix for the rotation between the two coordinate frames, and $\boldsymbol{Q_l}$ and $\boldsymbol{Q}$ are the homogeneous coordinate vectors of Q in the local and camera-centered frames respectively.

[1] Notation: All boldface letters denote coordinate vectors or arrays. We use capital letters to denote 3D points in the scene and small letters to denote their projections in the image.

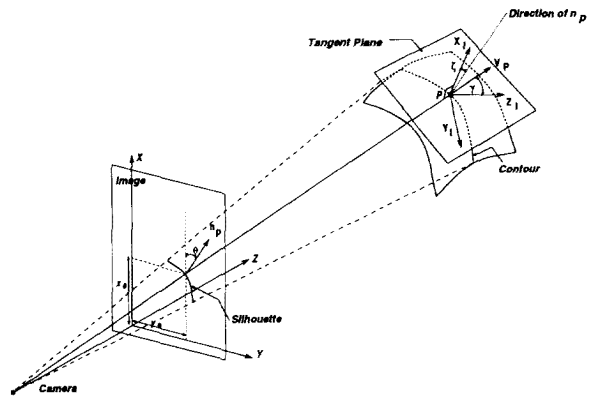


Figure 1: Projection geometry.

By definition, the surface normal at the contour point P is orthogonal to the viewing direction:

$$\boldsymbol{N}_P \cdot \boldsymbol{V}_P = 0, \qquad (3)$$

where $\boldsymbol{N}_P = [\frac{\partial \Sigma}{\partial X}, \frac{\partial \Sigma}{\partial Y}, \frac{\partial \Sigma}{\partial Z}]^T$, and $\boldsymbol{V}_P = [X_0, Y_0, Z_0]^T$. Eliminating Z between (2) and (3) gives the equation of the silhouette in the image coordinates, which is a conic section.

The paraboloid model described in this section is used to model the structure at each point along the silhouette. From a single projection of the paraboloid, we can obtain five constraints on the eight structure parameters $(\theta, \gamma, \zeta, X_0, Y_0, Z_0, \kappa_1, \kappa_2)$: the surface normal $\boldsymbol{N}_P$ at P can be computed as the cross product of $\boldsymbol{V}_P$ and the tangent $\boldsymbol{t}_p$ to the silhouette at p. This gives the angles θ and ζ directly. The coordinates of p give two constraints on X_0, Y_0 and Z_0. The curvature of the silhouette at p gives a constraint on κ_1, κ_2 and γ. To complete the local structure model, we need to estimate the depth Z_0 and obtain two more constraints on κ_1, κ_2, and γ. We obtain these constraints using the matched points from the other two images of the trinocular imagery. The next section presents the method to find correspondences.

2.2 Finding Correspondences

When the relative motion between the object and the camera is known for a pair of images, say I_1 and I_2 taken by two cameras with optical centers C_1 and C_2 respectively, we can estimate the epipolar plane and the epipolar line for each point in each image. Similar to the conventional stereo case, we can match points lying on corresponding epipolar lines, e.g. points p_1 and p_2 in Fig. 2. The difference here is that the two matched image points are not projections of the same 3D point. Similarly if we have a third camera C_3, we can find the epipolar match point p_3 in image I_3. Thus we have a triple of points matched in the three images using the epipolar geometry.

For a continuous relative motion between a camera and an object, the points matched using the instantaneous

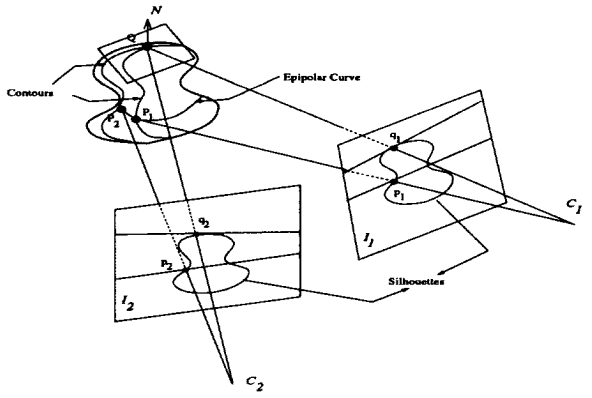


Figure 2: Epipolar geometry.

epipolar geometry trace a 3D curve (called the epipolar curve) on the object such that at each point the curve is tangent to the view line [1]. At points where the surface normal is not parallel to the epipolar plane normal, the epipolar curve is a regular curve e.g. at point P_1. But at points where the surface normal is parallel to the epipolar plane normal, the epipolar curve degenerates into a point, e.g. point Q in Fig. 2. In such a case, the matched points q_1 and q_2 are projections of the same 3D point Q. Such points are the *frontier points* for the corresponding camera motion [4]. We make use of this fact in estimating the translation parameters as explained in Sect. 3.

2.3 Structure Estimation

Previous approaches to structure estimation under known camera motion [1, 2, 12, 13] have used the epipolar parameterization of the surface. Our technique is similar to Szeliski's and Weiss' technique [12]; we choose the epipolar plane for the computation of one of the surface curvatures.

Since the epipolar curve is tangent to the view line at every point, we can estimate its osculating circle by finding the circle that is tangent to the three view lines[2]. The point where this circle touches the central view line provides an estimate of the depth of the 3D surface point. The curvature of the circle is an estimate of the curvature of the epipolar curve. This enables us to compute the normal curvature of the surface along the view line, which in turn gives a constraint on the surface curvatures κ_1, κ_2 and angle γ. Once the depth of the points along the silhouette is computed, we can estimate the direction of the tangent to the 3D contour. This tangent gives one more constraint on κ_1, κ_2 and γ since it is along a direction conjugate to the view line [10]. This constraint, along with the constraints given by the normal curvature and the curvature of the silhouette in the central image give us three equations which are solved to obtain the values of the structural parameters, κ_1, κ_2 and γ. See [7] for further details.

[2]If the three optical centers of the trinocular imagery are collinear, the two epipolar planes will be coplanar. But in general the optical centers will not be collinear, making the three view lines non-coplanar. In such a case the three view lines are projected onto a common plane, and a circle touching the projected lines is estimated. Recently, Boyer and Berger [2] have presented a technique where the assumption of linear camera motion is not necessary.

2.4 Experimental Results

We have applied the structure estimation algorithm to two real image sequences, each taken by a calibrated camera observing an object placed on a turntable. We simulated the trinocular imagery by taking triple of images 10 degrees apart in the sequence. Figure 3(a) shows one such image of one of the objects. Fig. 3(b) and (c) show the reconstructed 3D contour and the recovered Gaussian curvature along the contour respectively. The successive triples of images were taken with 5-degree rotations of the turntable.

3 Motion Estimation

With the results of the structure estimation algorithm presented in the previous section, we have the 3D contours on the object surface in the successive frames $I_1(t)$ and $I_1(t+1)$ of the central camera. Let us assume that the object has undergone a rotation of an unknown angle α about an unknown axis $\mathbf{\Omega} = [\omega_x, \omega_y, \omega_z]^T$ passing through the origin, followed by an unknown translation $[t_x, t_y, t_z]^T$ from t to $t+1$. Let $\boldsymbol{R}$ and $\boldsymbol{T}$ be the matrices corresponding to the unknown rotation and translation respectively.

Motion estimation under orthographic projection was addressed in [6]. Silhouette inflections were tracked over time to get an initial estimate of the rotation parameters. Under orthographic projection the rotational parameters alone determine the epipolar line direction. Thus the estimated rotation was sufficient to identify the frontier points and in turn estimate the translational parameters. This initial motion estimate was then iteratively refined.

In this section, we extend the approach to the more general case of perspective projection. In this case, we cannot separate the estimation of rotational and translational parameters, since the direction of the epipolar line is determined by both rotational and translational parameters. Therefore we have to modify the orthographic motion estimation algorithm.

The motion estimation is done in two steps. We first obtain an initial estimate of the rotation and translation parameters using the inflections and the frontier points. In the second step, the initial estimate is then refined using the rest of the silhouette. The next two sections describe these two steps in detail.

3.1 Obtaining the Initial Estimate

Rotational Parameters: Silhouette inflections are projections of parabolic points on the object surface [9, 13].

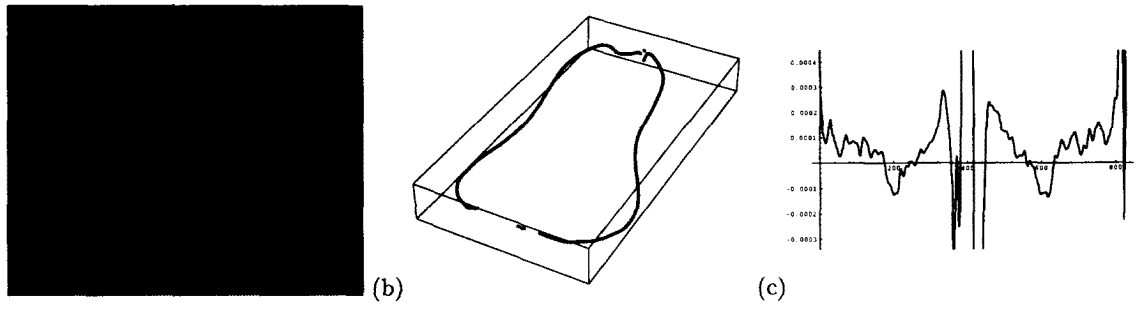


Figure 3: (a) sample image of the central camera; (b) recovered 3D contour; (c) recovered Gaussian curvature along the contour.

On generic surfaces these points lie on continuous curves called parabolic curves. With relative motion between the object and the viewer, the inflections in successive images will be projections of neighbouring points on the parabolic curves. A parabolic point has a single asymptotic direction and the surface is locally cylindrical. Consider the following lemma [7, 8]:

Lemma 1 *At a parabolic point, if we move along the surface in any direction (hence in particular, along the direction tangent to the parabolic curve), the change in the surface normal is perpendicular to the asymptotic direction.*

Consider an inflection point p which is the projection of a parabolic point P in the central image $I_1(t)$. If we track p in successive images, we will be moving along the parabolic curve at P. From Lemma 1, we can see that the change $\boldsymbol{dN}$ in the surface normal is orthogonal to the asymptotic direction, say $\boldsymbol{A}$, at P. We can compute $\boldsymbol{A}$ from the local structure estimated at time t. Let p' be the tracked inflection in $I_1(t+1)$; it is the projection of a neighbouring parabolic point P' with surface normal $\boldsymbol{N'}$. Using Lemma 1 and the fact that $\boldsymbol{A}$ lies in the tangent plane giving $\boldsymbol{A} \cdot \boldsymbol{N} = 0$, we have:

$$\boldsymbol{A} \cdot \boldsymbol{dN} = \boldsymbol{A} \cdot (\boldsymbol{R^{-1}N'} - \boldsymbol{N}) = \boldsymbol{A} \cdot \boldsymbol{R^{-1}N'} = 0. \quad (4)$$

We parameterize the rotation using the angle ϕ made by $\boldsymbol{\Omega}$ with the Z-axis, the angle ψ between its projection in the XY-plane and the X-axis, and the rotation angle α. With $n \geq 3$ inflections present in images $I_1(t)$ and $I_1(t+1)$, we use least-squares minimization with the objective function given by $\sum_{i=1}^{n} \left[\boldsymbol{A_i} \cdot (\boldsymbol{R^{-1}N_i'})\right]^2$.

The minimization is done over the angles ϕ, ψ and α. Note that we have to first find a match between the sets of inflections on the silhouettes in the two images. We match the inflections such that (1) the ordering of the matched inflections along the silhouette is maintained, and (2) the angle between the normals at the matched inflections is small.

Translational Parameters Under orthographic projection, with an estimate of the rotation parameters the epipolar plane normal can be estimated and in turn frontier points can be detected and utilized for the estimation of the translational parameters [6]. However this is not the case under perspective projection, where the rotational parameters alone do not determine the epipolar plane normal. The frontier points cannot be detected unless the translation parameters are estimated first. However our experience shows that the epipolar plane normal direction with the orthographic projection approximation is close to the corresponding epipolar plane normal direction for the perspective case. Given an estimate of the rotation parameters, we estimate the epipolar plane normal with the orthographic projection approximation and use it to detect the frontier points. Needless to say the detected frontier points are an approximation of the true frontier points of the perspective projection, but this approximation gives a good initial estimate of translational parameters. This initial estimate is refined in the second step.

Recall that the matched frontier points are projections of the same 3D point. Therefore if a frontier point F at time t matches the frontier point F' at time $t+1$, then $\boldsymbol{F'} = \boldsymbol{RF} + \boldsymbol{T}$. Thus for a given rotation, estimating the translation becomes a linear problem once the frontier points are detected and matched. This serves as the initial estimate of the translation parameters for a given rotation parameters.

3.2 Refining the Motion Estimate

In the second step, we use the structure along the entire silhouette to refine the estimate of the motion parameters obtained in the first step.

With an estimate of $\boldsymbol{R}$ and $\boldsymbol{T}$, we can determine the epipolar plane for each point in image $I_1(t)$. With an estimate of the structure parameters of the local paraboloids and the epipolar planes, we can estimate the curvature of the epipolar curve at each point. Consider a point q_i on the silhouette in image $I_1(t)$ (Fig. 4). We can *predict* the match point ${}^{p}q_i'$ in frame $I_1(t+1)$ as the projection of a point ${}^{p}Q_i'$ such that ${}^{p}Q_i'$ lies on the estimated osculating

circle of the epipolar curve and the surface normal at ${}^{p}Q_i'$ is orthogonal to the viewing direction. But we can also *detect* the epipolar match point ${}^{d}q_i'$ in image $I_1(t+1)$ as the intersection point of the estimated epipolar line corresponding to q_i with the silhouette at $t+1$.

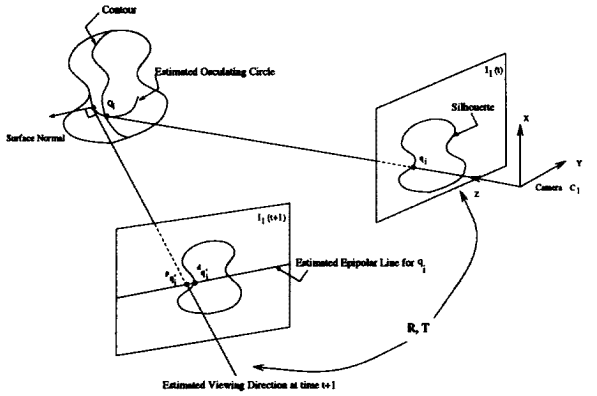


Figure 4: Predicted and detected epipolar matches.

In the refinement step, we iteratively minimize the sum of the squared distance between ${}^{p}q_i'$ and ${}^{d}q_i'$ for all silhouette points q_i. The minimization is over the six-dimensional space of $\boldsymbol{R}$ and $\boldsymbol{T}$ parameters. We iterate over the rotation parameters and at each iteration step we perform non-linear least-squares minimization and determine the best translation parameters that give the minimum sum of distances. In summary, the motion estimation algorithm can be given as follows.

1. Obtain an initial estimate of the rotation parameters $(\alpha_0, \phi_0, \psi_0)$ using tracked inflections. Set $\alpha = \alpha_0, \phi = \phi_0$ and $\psi = \psi_0$.
2. For the given rotation parameters α, ϕ, ψ
   (a) Detect and match the frontier points in the two central frames with the orthographic projection approximation and compute the initial estimate of translation parameters.
   (b) Knowing the local structure at each point on the silhouette, refine the estimate of the translation parameters to minimize the sum $S = \sum_{i=1}^{n} dist^2\ ({}^{p}q_i', {}^{d}q_i')$ of the distances between the predicted and the detected epipolar match points for all the silhouette points q_i.
3. Minimize S by updating the values of the rotation parameters α, ϕ and ψ, and repeating Step 2.

3.3 Implementation and Results

We can potentially consider the entire silhouette for the computation of the sum S. But observing that the structure estimation using epipolar matches gets less reliable as we approach the frontier points, we exclude the points close to the frontier points from the computation of S.

Note that if $\boldsymbol{R}$ and $\boldsymbol{T}$ represent the relative motion from time t to $t+1$, then $\boldsymbol{R}^T$ and $-\boldsymbol{R}^T\boldsymbol{T}$ represent the relative motion from $t+1$ to t. We can use the structural parameters estimated at time $t+1$ to predict the silhouette at time t, making use of all the structural information available. In practice, this has improved the performance of the motion estimation algorithm.

We have applied the method to the sequences of a squash and a bottle mentioned in Sect. 2.4. The rotation axis of the turntable was not parallel to the image plane of the camera. Thus the three effective optical centers were not collinear and the effective motion of the object was a general one.

Table 1 lists the recovered motion parameters after each step on a sample set of frames for the squash. For each step, we also list the angle between the true and estimated rotation axes and the error in the rotation angle. Both the steps of the motion estimation involve non-linear minimization requiring a starting point for the search. The first step of minimization using the tracked inflections was stable with respect to the starting point. The result of this step is used as the starting point for the refinement step.

Figure 5(a-b) shows two views of the global structure of the squash after 30 frames. The detected 3D occluding contours are placed in a common coordinate frame after derotating them using the estimated motion.

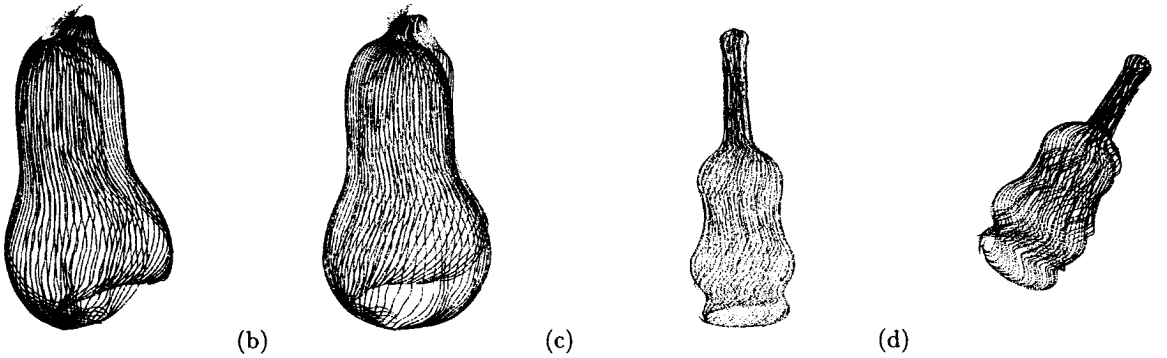
For the squash sequence, the instantaneous motions at successive time frames were estimated independently. To take advantage of the motion continuity, we modified the step of obtaining the initial motion estimate as follows. First, an initial estimate of the motion using the inflection and frontier points at a given time t is computed. We then compare the value of the objective function S using this motion estimage to the value of S using the final motion result of the previous time $t-1$. We then select the motion that yields a smaller value of S among these two, and initiate the refinement step. This modification also avoids any errors in situations when the estimate of the rotation parameters based on inflections may be far off from the solution due to noise or because the number of inflections present is small. Table 1 presents sample results of the modified motion estimation algorithm when applied to the bottle sequence. Figure 5(c-d) shows two views of the global structure of the bottle after 25 frames.

4 Discussion

Although estimating motion from silhouettes is more difficult than using viewpoint-independent features, we have the advantage that we have more information about the surface even from a single silhouette - the surface normal, the sign of the Gaussian curvature and a constraint on the principal curvatures at the surface point. We have used the relationship between certain silhouette features (inflections) and a model of the local surface structure to estimate both the motion and the global surface structure from perspective views of the silhouettes of a moving

		$\Omega(\omega_x, \omega_y, \omega_z)$	Error in Ω	α	Error in α	(t_x, t_y, t_z) in mm	Sum
				Squash			
	True	(0.008,0.94,0.33)	-	5.0	-	(125.6, 0.64, -5.03)	-
1	1st Step	(0.0677,0.953,0.294)	4.	5.59	0.586	(142.,-8.4,-6.15)	25.7
	Final	(0.0312,0.947,0.32)	1.42	5.25	0.248	(132.,-2.57,-5.41)	10.
2	1st Step	(-0.0424,0.942,0.334)	2.92	5.003	0.003	(125.,7.7,-5.08)	19.3
	Final	(-0.11,0.939,0.324)	6.82	5.002	0.002	(125.,17.,-5.49)	14.2
3	1st Step	(-0.0557,0.948,0.314)	3.79	5.04	0.0409	(127.,9.29,-3.23)	132.
	Final	(0.00372,0.944,0.33)	0.268	4.87	0.132	(122.,1.19,-4.88)	8.74
				Bottle			
	True	(-0.042,0.97,0.25)	-	5.0	-	(130.1, 6.99, -5.17)	-
1	1st Step	(-0.07,0.97,0.22)	2.07	8.86	3.86	(232.63, 19.53, -13.14)	259.0
	Final	(0.0312,0.947,0.32)	1.42	5.25	0.25	(132.9, -1.36, -4.6)	12.6
2	1st Step	(-0.085,0.93,0.35)	6.59	6.24	1.24	(142.,-8.4,-6.15)	25.7
	Final	(-0.12,0.96,0.24)	4.47	6.91	1.91	(179.5, 24.7, -8.9)	18.74
3	1st Step	(-0.29,0.93,0.22)	14.34	7.08	2.08	(178.88, 59.10, -11.80)	31.58
	Final	(-0.04,0.97,0.24)	0.52	6.54	1.54	(172.9, 19.03, -9.23)	13.73

Table 1: Result of the motion estimation for the squash and the bottle (all angles are in degrees).



(a) (b) (c) (d)

Figure 5: (a-b)two views of the global structure of the squash after 30 frames; (c-d)two views of the global structure of the bottle after 25 frames.

object. In estimating the motion we have also used another set of points on the silhouette: the frontier points. The results obtained on real images are encouraging and demonstrate the validity of the method.

References

[1] A. Blake and R. Cippola. Surface shape from the deformation of apparent contours. *Int. J. of Comp. Vision*, 9(2), 1992.

[2] E. Boyer and M. O. Berger. 3d surface reconstruction using occluding contours. Technical Report 95-R-013, Crin/Inria Lorraine, 1995.

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

[4] P. Giblin, J. Rycroft, and F. Pollick. Moving surfaces. In *Mathematics of Surfaces V*. Cambridge University Press, 1993.

[5] P. Giblin and R. Weiss. Reconstruction of surfaces from profiles. In *Proc. Int. Conf. Comp. Vision*, 1987.

[6] T. Joshi, N. Ahuja, and J. Ponce. Silhouette-based structure and motion estimation of a smooth object. In *Proc. ARPA Image Understanding Workshop*, 1994.

[7] T. Joshi, N. Ahuja, and J. Ponce. Structure and motion estimation from dynamic silhouettes under perspective projection. Technical Report UIUC-BI-AI-RCV-95-02, Beckman Institute, University of Illinois, 1995.

[8] T. Joshi, J. Ponce, B. Vijayakumar, and D.J. Kriegman. Hot curves for modelling and recognition of smooth curved 3d shapes. In *Proc. IEEE Conf. Comp. Vision Patt. Recog.*, 1994.

[9] J. J. Koenderink. What does the occluding contour tell us about solid shape. *Perception*, 13, 1984.

[10] J. J. Koenderink. *Solid Shape*. MIT Press, MA, 1990.

[11] K. Kutulakos and C. Dyer. Global surface reconstruction by purposive control of observer motion. In *Proc. IEEE Conf. Comp. Vision Patt. Recog.*, 1994.

[12] R. Szeliski and R. Weiss. Robust shape recovery from occluding contours using a linear smoother. In *Proc. IEEE Conf. Comp. Vision Patt. Recog.*, 1993.

[13] R. Vaillant and O. Faugeras. Using extremal boundaries for 3-d object modeling. *IEEE Trans. Patt. Anal. Mach. Intell.*, 14(2), 1992.

[14] J. Y. Zheng. Acquiring 3-d models from sequences of contours. *IEEE Trans. Patt. Anal. Mach. Intell.*, 16(2), 1994.