In Proc. IEEE Conference on Computer Vision and Pattern Recognition (CVPR), San Juan, Puerto Rico, pages 1094-1099, 1997.

Stereo Coupled Active Contours

Tat-Jen Cham Roberto Cipolla

Department of Engineering
University of Cambridge
England
E-mail: {tjc|cipolla}@eng.cam.ac.uk

Abstract

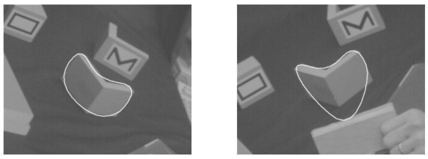
We consider how tracking in stereo may be enhanced by coupling pairs of active contours in different views via affine epipolar geometry and various subsets of planar affine transformations, as well as by implementing temporal constraints imposed by curve rigidity. 3D curve tracking is achieved using a submanifold model, where it is shown how the coupling mechanisms can be decomposed to cater for fixed and variable epipolar geometries. In the case of tracking planar curves, the canonical frame model is developed such that the various geometrical constraints needed in different situations may be efficiently selected. The results show that coupled active contours add consistency and robustness to tracking in stereo.

1. Introduction

The geometry underlying multiple camera views has been well studied, especially since it is the primary tool used in stereo vision and structure from motion techniques. Multi-view geometry has also been used in the *monocular* tracking of objects. The affine active contour first proposed by Blake, Curwen and Zisserman [1] and also used in [5], is one which is constrained to deform only in terms of affine transformations acting globally on all contour points – this is an example where the compatibility between different weak-perspective views of a 2D rigid curve is enforced. However the main thrust of research in this area has been in the incorporation of probability for robust tracking [1] and the development of complex 3D active contour models [8].

The task of tracking the same object in multiple camera views is much less researched. One of the disadvantages of tracking the same object in different views *independently* is that shape inconsistencies cannot be prevented, as exemplified by fig. 1.

Braud, Lapresté and Dhome [2] have considered tracking polyhedral objects with multi-ocular vision, but do not incorporate constraints into the tracking mechanism. Imposing constraints on active contours have also been studied by Fua and Brechbühler [4], but the constraints are mainly used to interpolate fixed points or satisfy tangencies and do not concern tracking. Reynard *et al.* [6] used coupling between different active contours in a monocular image sequence, but this is not based on the geometry between views and requires training of the active contours.



(a1) (a2)

Figure 1. The above figure shows a pair of image frames (a1,a2) taken from *stereo* image sequences. When tracking is carried out separately by using independent active contours, the large flexibility can result in inconsistent shape deformation between pairs of active contours due to the presence of nearby clutter.

Tracking with multiple cameras can benefit from the use of multi-view geometry since the shape consistency of active contours can be enforced over two domains: not only should shape deformation be geometrically compatible across the temporal domain, but the shapes would also have to be compatible across different cameras.

2. Stereo Coupling of Active Contours

The setup considered here is a stereo pair of cameras simultaneously tracking a curve in 3D space using B-spline active contours. In this case the points on the two active contours which have the same spline parameter value are corresponding points. We further make the simplifying assumption that the cameras are *affine* over the regions of the active contours.The following stereo tracking situations will be studied here:

1. tracking non-rigid 3D curves with separate analysis for fixed and variable stereo epipolar geometry;
2. tracking rigid and non-rigid planar curves, again with fixed and variable stereo epipolar geometry; and
3. tracking rigid and non-rigid curves in a fixed plane.

The stereo tracking of rigid 3D curves is also an important case to consider but this has been left for future work.

2.1. Models for Coupled Active Contours

Two different models are used for coupled active contours in the different tracking situations considered above:

1. **Submanifold Model**. The positions of all the control points in the two B-spline active contours are treated as a single vector representing some state in a high-dimensional space. Forcing all iterations of the active contours to satisfy some form of geometry (eg. epipolar or 2D affine) therefore involves projecting the state-change vectors onto the associated lower dimension manifold. This general model encompasses the affine subspace projection method by Blake *et al.* [1].
2. **Canonical Frame Model**. In this model, the two 'slave' active contours in the image frames are affine transformed versions of a 'master' active contour which lies in a canonical frame. This formulation intrinsically decomposes the deformation of the active contours into two steps: a deformation in the master active contour which drives the deformation in the slave active contours, and a change in the affine transformations relating the canonical frame to the image frames. See fig. 2. It is *not* satisfactory to treat one of the active contours in an image frame as a master, since this leads to a bias in tracking errors (the slave contour will have larger errors). This model is only suitable for tracking planar curves.

3. Affine Epipolar Coupling Mechanisms

Epipolar geometry may be used to couple stereo active contours tracking 3D curves via the submanifold model. The geometry subspace may be determined according to the *affine epipolar constraint* [7]:

$$\begin{bmatrix} x & y & x' & y' & 1 \end{bmatrix} \begin{bmatrix} f_1 \\ f_2 \\ f_3 \\ f_4 \\ 1 \end{bmatrix} = 0 \qquad (1)$$

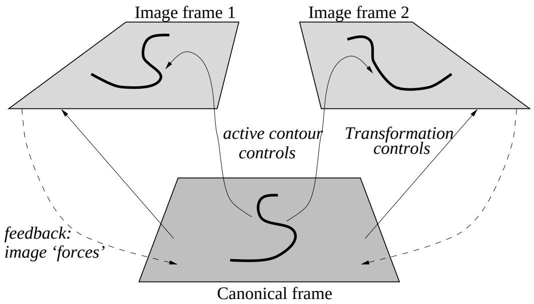


Figure 2. The canonical frame model for stereo-coupled active contours comprises a 'master' active contour in a canonical reference frame, which controls the 'slave' active contours in the image frames. See text for details.

where $f_1 \ \ldots \ f_4$ are the elements of some *affine fundamental matrix*, while $\boldsymbol{x}$ and $\boldsymbol{x'}$ are corresponding image points which may be considered to be corresponding control points of the B-spline active contours in the current context.

Instead of computing the best state for a pair of *unconstrained* active contours (represented by an augmented vector involving control point positions) and projecting this state onto the submanifold, we can ensure that the active contours are started such that the initial state lies on the submanifold and thereafter force all state updates to similarly lie on the submanifold.

Suppose the changes in state $\boldsymbol{d}_j$'s computed for the pair of *unconstrained* active contours during a tracking process is given by

$$\boldsymbol{d}_j = \begin{bmatrix} \Delta p_{jx} & \Delta p_{jy} & \Delta p'_{jx} & \Delta p'_{jy} \end{bmatrix}^T \ , \ j = 1 \ldots N \qquad (2)$$

where $\boldsymbol{p}_j = [\ p_{jx} \ p_{jy}\ ]^T$ and $\boldsymbol{p'}_j$ are the jth corresponding control points (in vector form) on the splines out of a total of N.

From (1), the changes in state $\hat{\boldsymbol{d}}_j$'s for *epipolar-coupled* active contours must necessarily satisfy

$$\begin{bmatrix} \boldsymbol{p}_j^T & \boldsymbol{p'}_j^T \end{bmatrix} \boldsymbol{\Delta f} + \hat{\boldsymbol{d}}_j^T (\boldsymbol{f} + \boldsymbol{\Delta f}) = 0 \qquad (3)$$

where $\boldsymbol{\Delta f}$ is the change in the vectored affine fundamental matrix $\boldsymbol{f}$ given by

$$\boldsymbol{f} = \begin{bmatrix} f_1 & f_2 & f_3 & f_4 \end{bmatrix}^T \qquad (4)$$

which may be fixed throughout the tracking (eg. fixed cameras) or updated at each time-step.

The preferred manner in which to calculate the changes in state $\hat{\boldsymbol{d}}_j$'s *and* the change in epipolar geometry $\boldsymbol{\Delta f}$ is to carry this out in two steps:

1. **Fixed Epipolar Geometry**. Compute $\hat{\boldsymbol{d}}_{ju}$, the components of $\hat{\boldsymbol{d}}_j$ perpendicular to the current $\boldsymbol{f}$. These are

exactly the components of the unconstrained changes of state $\boldsymbol{d}_j$'s perpendicular to $\boldsymbol{f}$, *ie. the components which satisfy the current epipolar geometry*. Since the epipolar geometry does not change, neither does $\boldsymbol{f}$.

2. **Updating Epipolar Geometry**. Compute the optimal change in epipolar geometry $\boldsymbol{\Delta f}$ via a least-squares operation on the components of $\boldsymbol{d}_j$'s parallel to the current $\boldsymbol{f}$. $\hat{\boldsymbol{d}}_{jv}$'s, the components of $\hat{\boldsymbol{d}}_j$'s parallel to the *updated* $\boldsymbol{f}$ may be found via (3).

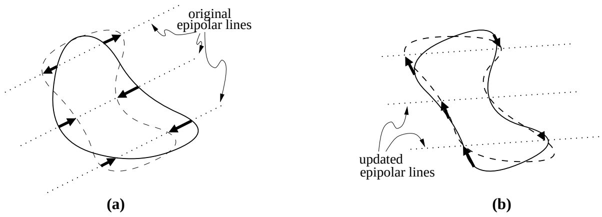


Figure 3. If the active contour in one image frame is fixed, the deformation of its opposite stereo half may be decomposed into: (a) a deformation along the current epipolar lines, and (b) a deformation perpendicular to the current epipolar lines with corresponding updates to the epipolar lines.

The final epipolar-coupled changes in state are given by

$$\hat{\boldsymbol{d}}_j = \hat{\boldsymbol{d}}_{ju} + \hat{\boldsymbol{d}}_{jv} \tag{5}$$

In order to provide an intuitive idea for this decomposition, fig. 3 shows the deformation of an active contour S in one image frame for the case in which its opposite active contour S' remains fixed. Step 1 adjusts deforms the active contour along the epipolar lines thereby retaining the epipolar geometry, while step 2 further adjusts the epipolar geometry based on components of 'image forces' perpendicular to the epipolar lines. The advantage of such a decomposition is that if the epipolar geometry is known to be fixed, only step 1 needs to be carried out. Details are given in [3].

3.1. Results

In fig. 4, results for tracking a 3D curve under fixed epipolar geometry are presented. The epipolar consistency enforced in the coupled active contours help to overcome tracking distraction caused by clutter in directions perpendicular to the epipolar lines (the active contours behave like unconstrained active contours in directions parallel to the epipolar line as would be expected). Without the inbuilt epipolar constraints, the active contours are much more likely to be attracted to neighbouring strong edges, as was shown in fig. 1.

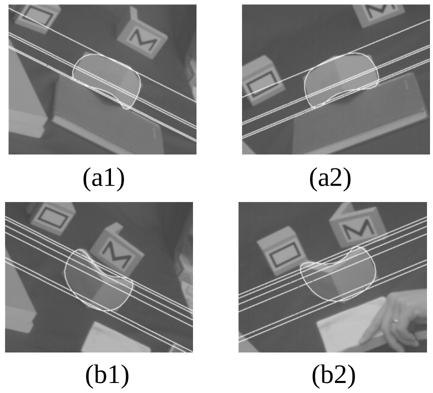


Figure 4. The pairs of active contours shown in figures (a1),(a2) and (b1),(b2) are constrained to share the same affine epipolar geometry. Note the white edge in (b2) does not distract the tracker as was the case in fig. 1.

In fig. 5, the results for tracking 3D curves with a variable epipolar geometry, as would be necessary when the cameras are moving, are shown. The constraints in this case are comparatively weaker than those in the previous case with fixed epipolar geometry, in that there are an additional four degrees of freedom. More details are provided in [3].

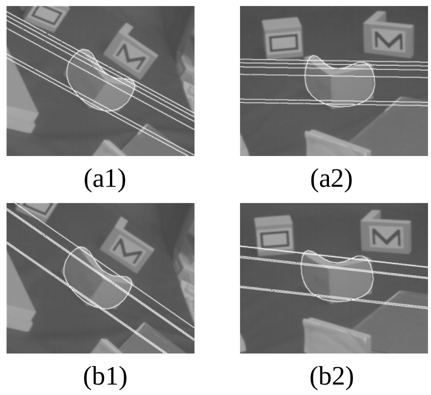


Figure 5. When the epipolar geometry is not fixed but iteratively updated, the additional degrees of freedom cater for changes in camera configuration, but at a loss of tracking robustness.

4. Coupling with 2D Affine Transformations

When simultaneously tracking planar curves in cameras with small fields of view, the active contours can be constrained to share the same affine structure. In the case, the control points $\boldsymbol{p}_j$'s and $\boldsymbol{p'}_j$'s of corresponding splines will be related by an affine transformation given by

$$\boldsymbol{p'}_j^T = \boldsymbol{p}_j^T \boldsymbol{A} + \boldsymbol{t}^T \quad (6)$$

where

$$\boldsymbol{A}^T = \begin{bmatrix} a_1 & a_2 \\ a_3 & a_4 \end{bmatrix}, \ \boldsymbol{t} = \begin{bmatrix} t_1 \\ t_2 \end{bmatrix} \quad (7)$$

and $a_1 \ldots a_4$, $t_1 \ldots t_2$ are parameters of the affine transformation.

Similarly, both $\boldsymbol{p}_j$'s and $\boldsymbol{p'}_j$'s may be related to control points $\boldsymbol{q}_j$'s in some *canonical frame* such that

$$\boldsymbol{p}_j^T = \boldsymbol{q}_j^T \boldsymbol{A_1} + \boldsymbol{t_1}^T \quad (8)$$

$$\boldsymbol{p'}_j^T = \boldsymbol{q}_j^T \boldsymbol{A_2} + \boldsymbol{t_2}^T \quad (9)$$

where $\boldsymbol{A_1}$, $\boldsymbol{A_2}$ are matrices and $\boldsymbol{t_1}$, $\boldsymbol{t_2}$ are vectors such that

$$\boldsymbol{A} = \boldsymbol{A_1}^{-1} \boldsymbol{A_2} \quad (10)$$

$$\boldsymbol{t}^T = \boldsymbol{t_2}^T - \boldsymbol{t_1}^T \boldsymbol{A_1}^{-1} \boldsymbol{A_2} \quad (11)$$

The control points of the *master* active contour in the canonical frame are given by $\boldsymbol{q}_j$'s. The deformation of the slave active contours in the image frames are therefore effected by manipulating the master active contour via $\boldsymbol{q}_j$'s *and* the affine transformation parameters $\boldsymbol{A_1}$, $\boldsymbol{A_2}$, $\boldsymbol{t_1}$ and $\boldsymbol{t_2}$.

The advantage of the Canonical Frame Model lies in the ease in choosing different modes of operation. Table 1 show how different tracking modes can be effected. For example if a rigid curve is to be tracked under variable epipolar geometry (eg. if cameras are not stationary), the master active contour in the canonical frame is fixed by keeping $\boldsymbol{q}_j$'s constant, while the affine transformation parameters relating the canonical frame to the image frames $\boldsymbol{A_1}$, $\boldsymbol{A_2}$, $\boldsymbol{t_1}$ and $\boldsymbol{t_2}$ can be optimally updated without constraints in the tracking process.

4.1. 2D Affine Transformations with Variable Epipolar Geometry

The equations for updating the parameters $\boldsymbol{q}_j$'s, $\boldsymbol{A_1}$, $\boldsymbol{A_2}$, $\boldsymbol{t_1}$, and $\boldsymbol{t_2}$ are fairly straightforward and are given in [3].

The results for tracking 2D affine curves using the canonical frame model are shown in fig. 6,7. For fig. 6(a1,a2,b1,b2), the active contours are tracking in the *rigid curve mode* in that only affine deformations of the contours are allowed. This is very much similar to the stereo affine-deforming active contours used in [5], except that in this case the two contours share the same affine structure.

This is more evident in fig. 7(a1,a2,b1,b2) in which the tracking is carried out using the *non-rigid curve mode*. The active contours are not constrained to deform affinely, but are still required to share the same affine structure. *This is particularly useful for initialising the active contours in situations when the shapes of the contours being tracked are not known in advance.*

Tracking Mode?		**Action Required**	
Class of Geometry	*Rigidity of curve*	*Update action for master active contour*	*Update action for affine transformation*
Variable epipolar	Non-rigid	Unconstrained	Unconstrained
Variable epipolar	Rigid	Fixed	Unconstrained
Fixed epipolar	Non-rigid	Unconstrained	Epipolar-consistent
Fixed epipolar	Rigid	Fixed	Epipolar-consistent
Fixed planar	Non-rigid	Unconstrained	Fixed
Fixed planar	Rigid	Affine Deformation	Fixed

Table 1. Modes of operation for the Canonical Frame Model. The two left columns show the choices of tracking operation available. The two right columns show the update actions which must be performed for the desired tracking operation.

4.2. 2D Affine Transformations with Fixed Epipolar Geometry

It is possible to restrict the changes in the affine transformation parameters such that the epipolar geometry remains fixed, as shown by the following theorem:

Theorem 1 (2D Affine Geometry – Epipolar Geometry Relation)
If a set of points related by a 2D affine transformation with parameters a_1, a_2, a_3, a_4, t_1 and t_2 as defined in (6) are also satisfying the affine epipolar geometry with parameters f_1, f_2, f_3 and f_4 as in (1), the parameters are necessarily related in the following way:

$$\begin{bmatrix} f_3 & f_4 & 0 & 0 & 0 & 0 & f_1 \\ 0 & 0 & f_3 & f_4 & 0 & 0 & f_2 \\ 0 & 0 & 0 & 0 & f_3 & f_4 & 1 \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ a_3 \\ a_4 \\ t_1 \\ t_2 \\ 1 \end{bmatrix} = \mathbf{0} \quad (12)$$

In particular we see that the six independent parameters in the affine transformation in (6) is confined to *three* degrees of freedom in this case. Proof of theorem 1 and further details may be found in [3].

Figure 8 shows the tracking of rigid planar curves in which the epipolar geometry is fixed. The additional constraints imposed provide greater resilience against clutter and minor occlusions.

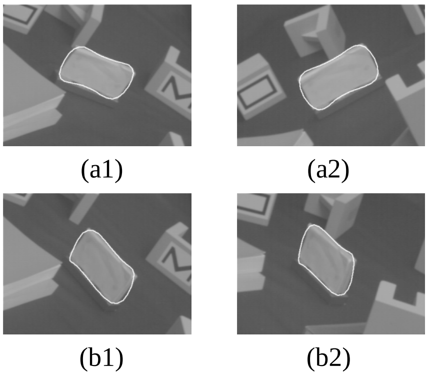


Figure 6. (a1,a2,b1,b2) In rigid curve mode, pairs of stereo active contours are not only required to share the same 2D affine structure, but are also required to deform affinely in time.

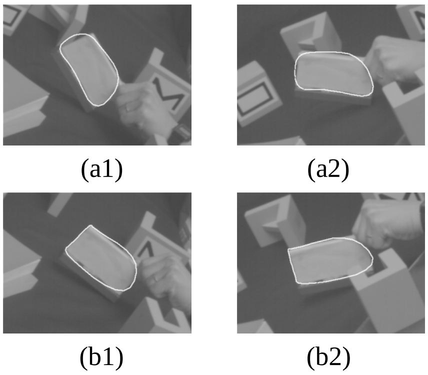


Figure 7. In (a1,a2,b1,b2), the pairs of active contours share the same affine structure but are not required to deform affinely in time. This is useful for tracking non-rigid 2D curves.

4.3. Results for Tracking Curves in a Fixed Affine Plane

In some situations it is useful to track planar curves moving in a fixed plane. For example it may be useful to track the roof outlines of moving cars in a traffic scene, or the boundaries of biological cells.

In fig. 9, the tracking of rigid planar curves which lie in some fixed plane is shown. In this case, the planar curve being tracked lies in the ground plane. While fig. 9(a1,a2,b1,b2) show that movement of the curve is tracked, fig. 9(c1,c2) demonstrate that the tracking is robust to occlusion.

4.4. Affine Symmetry

The main purpose for developing affine symmetrically coupled active contours is that the tracking of objects that

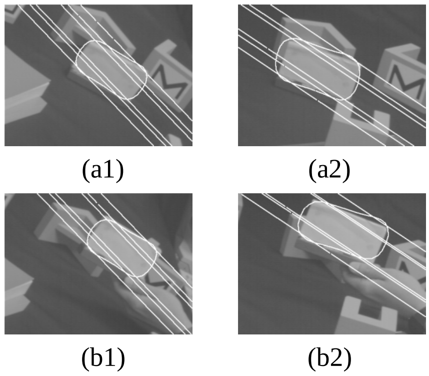


Figure 8. If the pair of stereo active contours is not only affine transformation-related but also operate under a fixed epipolar geometry, robustness of tracking to background clutter improves.

are surfaces of revolution is best achieved by following their occluding contours, which are affine symmetric pairs[1]. These active contours differ from the ones developed in the previous sections in that pairs of coupled contours will be located in the same image, and therefore may be used for a monocular tracking system. Moreover, it is also possible to enforce the ends of both active contours to be joined, such that they may be treated as a single affine symmetric active contour. The canonical frame model is, as with other 2D coupled contours, best employed here, since an affine symmetrical transformation may be represented via a canonical frame model formulation, such that

$$\boldsymbol{q}_j^T = \boldsymbol{p}_j^T \boldsymbol{Z_1} + \left[\begin{array}{cc} C & 0 \end{array} \right] = \boldsymbol{p'}_j^T \boldsymbol{Z_2} - \left[\begin{array}{cc} C & 0 \end{array} \right] \quad (13)$$

where

$$\boldsymbol{Z}_1 = \left[\begin{array}{cc} -\sin\phi & \sin\alpha \\ \cos\phi & -\cos\alpha \end{array} \right], \; \boldsymbol{Z}_2 = \left[\begin{array}{cc} \sin\phi & \sin\alpha \\ -\cos\phi & -\cos\alpha \end{array} \right] \quad (14)$$

and α, ϕ, and C are independent affine symmetry parameters. The canonical frame corresponds to a symmetry-rectified coordinate frame such that the y-axis is the axis of symmetry, and all lines of symmetry are parallel to the x-axis, while the points $\boldsymbol{q}_j$'s correspond to one half of the pairs of the affine symmetrical image points $\boldsymbol{p}_j$'s and $\boldsymbol{p'}_j$'s. Once again, details of the updating mechanisms are given in [3].

In fig. 10, we show the monocular tracking of affine symmetric contours. The associated symmetry axis and angle of skew are also represented. In this case, the affine symmetric tracker treats the two symmetric contours as a single closed curve, and is used in the tracking of a surface of revolution (a lampshade).

[1] With fixating cameras, the occluding contours of surfaces of revolution are bilateral symmetric.

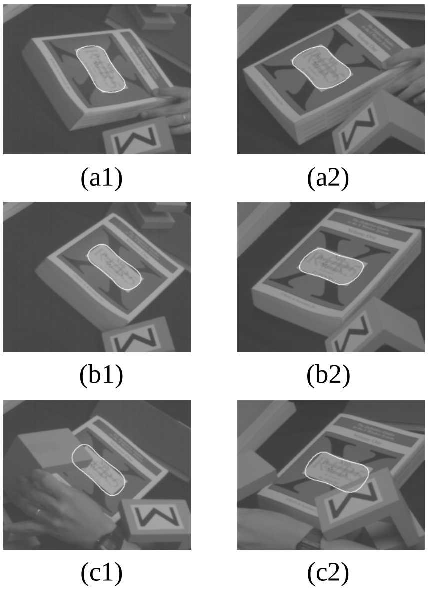


Figure 9. When tracking planar curves confined to a fixed plane, the affine transformation relating the pair of stereo active contours is also fixed. By further enforcing temporal affine deformation, robustness to occlusion can be achieved.

5. Conclusions

Geometrically coupled active contours are means by which geometric constraints can be consistently applied across camera-temporal domains when tracking image contours. In particular the submanifold and canonical frame models were proposed as useful representations for coupled active contours. The submanifold model may be used for coupling the contours via affine epipolar geometry, with the deformation mechanisms decomposed in a way such that choosing between fixed and variable epipolar geometry is simplified. The results show that epipolar-coupled contours are useful for tracking the image contours of space curves. The canonical frame model is particularly suited for tracking planar curves for which the image contours are related by planar affine transformations, since there is considerable ease in switching between rigid or flexible curve modes as well as selecting the various geometries of coupling which comprise of planar affine transformations, planar affine transformations under fixed affine epipolar geometry, fixed planar affine transformations and affine symmetry. Results obtained demonstrate that by maximising the use of geometrical constraints, additional robustness to occlusion and clutter can be achieved in tracking.

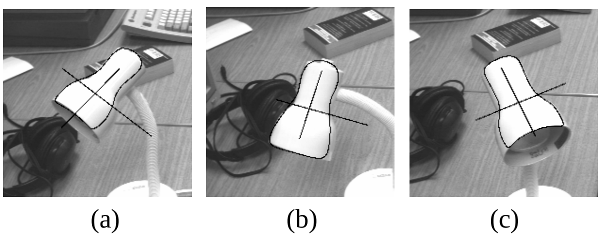


Figure 10. The pair of affine symmetric curves are tracked over an image sequence as a single closed curve by combining the control points. Tracking is fairly robust against background clutter.

Acknowledgements

The authors would like to thank Olivier Faugeras, Nick Kingsbury and Andrew Zisserman for their comments and suggestions.

References

[1] A. Blake, R. Curwen, and A. Zisserman. A framework for spatiotemporal control in the tracking of visual contours. *Int. Journal of Computer Vision*, 11(2):127–145, 1993.

[2] P. Braud, J.-T. Lapresté, and M. Dhome. Recognition, pose and tracking of modelled polyhedral objects by multi-ocular vision. In B. Buxton and R. Cipolla, editors, *Proc. 4th Euro. Conf. on Computer Vision, Cambridge (England)*, volume 1065 of *Lecture Notes in Computer Science*, pages 455–464. Springer-Verlag, 1996.

[3] T.J. Cham. *Geometric Representation and Grouping of Image Curves*. PhD thesis, Department of Engineering, University of Cambridge, Aug 1996.

[4] P. Fua and C. Brechbühler. Imposing hard constraints on soft snakes. In B. Buxton and R. Cipolla, editors, *Proc. 4th Euro. Conf. on Computer Vision, Cambridge (England)*, volume 1065 of *Lecture Notes in Computer Science*, pages 495–506. Springer-Verlag, 1996.

[5] N. Hollinghurst and R. Cipolla. Uncalibrated stereo hand-eye coordination. *Image and Vision Computing*, 12(3):187–192, 1994.

[6] D. Reynard, A. Wildenberg, A. Blake, and J. Marchant. Learning dynamics of complex motions from image sequences. In B. Buxton and R. Cipolla, editors, *Proc. 4th Euro. Conf. on Computer Vision, Cambridge (England)*, volume 1064 of *Lecture Notes in Computer Science*, pages 357–368. Springer-Verlag, 1996.

[7] L.S. Shapiro, A. Zisserman, and M. Brady. 3D motion recovery via affine epipolar geometry. *Int. Journal of Computer Vision*, 16:147–182, 1995.

[8] D. Terzopoulos and D. Metaxas. Tracking nonrigid 3D objects. In A. Blake and A. Yuille, editors, *Active Vision*, chapter 5, pages 75–89. MIT Press, 1992.