Reconstruction of 3D-Curves from 2D-Images Using Affine Shape Methods for Curves*

Rikard Berthilsson, Kalle Åström
Dept of Mathematics, Lund University
Box 118, S-221 00 Lund, Sweden

Key words: Projective geometry, subspace methods, structure from motion, curves, affine shape, invariancy

Abstract

In this paper, we propose an algorithm for doing reconstruction of general 3D-curves from a number of 2D-images taken by uncalibrated cameras. No point correspondences between the images are assumed except for the end points. The curve and the view points are uniquely reconstructed, modulo common projective transformations. Furthermore, the algorithm is independent of the choice of coordinates, as it is based on orthogonal projections and aligning subspaces. The ideas behind the algorithm are based on an extension of affine shape of finite point configurations to more general objects, like for example curves and surfaces.

1 Introduction

Affine shape of finite point configurations, has been introduced and studied in a series of papers by Sparr, see for example [Spa96], where a reconstruction algorithm, for arbitrary numbers of points and images, was proposed. This was based on aligning subspaces by using orthogonal projections and maximising some of the largest eigenvalues of the sum of these projections. That algorithm is here extended and modified to handle curves as well. Furthermore, no point correspondences between the different images are needed. For simplicity, however, correspondences are assumed at the ends of each curve. The algorithm is independent of the choice of coordinates and is generic in the numbers of images taken. Furthermore, the method presented here does not require any derivatives of the curves and it is enough to require integrability, which makes it rather robust. For convenience though, we assume continuity.

Another way of reconstructing curves can be found in [FP93, PF96], where a theory for computing the structure and motion of three-dimensional curves is developed. It is shown there that the image of three-dimensional curves obey several constraints at each point. These constraints involve high order spatio-temporal derivatives of the image curve motion and camera motion, which therefore might lead to numerical problems.

Yet another contribution can be found in [PP91, Car94, ÅCG96], where image-motion constraints that hold for certain points on the curve are developed. These constraints seem more robust than the ones mentioned above and can be applied to the silhouette of curved surfaces, but they do not exploit the full structure of the curve reconstruction problem.

2 Extension of Affine Shape to Curves

Our aim, in this section, is to define affine shape of curves and to obtain conditions for projective reconstruction of general open curves when the point correspondences are known between the images. By an open curve, we loosely mean that it has different starting and ending points. Later on, these conditions will enable us to construct an algorithm for reconstruction when *no* point correspondences between the images are assumed except at the end points of each curve. We would also like the conditions for projective reconstruction to be independent of the choice of coordinates, generic in choosing number of images, robust and easy to extend to recurrent `recursive` algorithms, where more and more images are taken into account. First we give a few definitions.

Let $\mathcal{X}$ be an open curve in $\mathbb{R}^n$. It is then possible to define a continuous and bijective map $\phi : I \to \mathcal{X}$, with $I = [0, 1]$ and where $\phi(t) = (\phi_1(t), \dots, \phi_n(t))$ and each

*This work has been done within then ESPRIT Reactive LTR project 21914, CUMULI and the Swedish Research Council for Engineering Sciences (TFR), project 95-64-222

$\phi_i(t)$ is a coordinate for the point $\phi(t)$ on the curve.

Introduce the scalar product of two functions $f, g \in L^2(I)$ by

$$\langle f|g\rangle_I = \int_I fg dx,$$

when g is a real valued function and

$$\langle f|\phi\rangle_I = (\langle f|\phi_1\rangle_I, \ldots, \langle f|\phi_n\rangle_I),$$

when $\phi : I \to \mathbb{R}^n$ is a vector valued function. For a bijective map $L^2(I) \ni \phi : I \to \mathcal{X}$, we now define the affine shape of ϕ as the linear space

$$s(\phi) = \left\{ f \mid \langle f|\phi\rangle_I = 0,\ \langle f|1\rangle_I = 0,\ f \in L^2(I) \right\}.$$

It can be shown that $s(\phi)$ is a complete affine invariant and of course is infinite in dimension. Furthermore, the shape $s(\phi)$ is not entirely an intrinsic property of $\mathcal{X}$, as it changes with different choices of ϕ. This is why we denote it by $s(\phi)$ instead of $s(\mathcal{X})$. This non intrinsic property causes problems when no point correspondences between the images are assumed.

We will now use $s(\phi)$ in the pinhole camera model. For that, let $\mathcal{X}$ be a curve in $\mathbb{R}^3$ and $\mathcal{Y}$ its image in $\mathbb{R}^2$. Fix the set $I = [0, 1]$ once and for all and define bijective and continuous maps $\phi : I \to \mathcal{X}$ and $\psi : I \to \mathcal{Y}$, such that for each $t \in I$, $\psi(t)$ is the image of the point $\phi(t)$. (We always reserve ϕ for the object curve and ψ for the image curve.) Then there exists a function $\alpha : I \to \mathbb{R}$, called the depth of ψ and a camera center $c \in \mathbb{R}^3$, such that

$$\overrightarrow{c\phi(t)} = \alpha(t)\overrightarrow{c\psi(t)}, \quad t \in I.$$

It can be shown that $\mathcal{Y}$ is the image of $\mathcal{X}$ with corresponding maps ψ and ϕ respectively and depths α if and only if

$$\alpha s(\phi) \subseteq s(\psi).$$

Furthermore, it can be shown that if $\{\mathcal{Y}_i\}_0^{m-1}$ is a set of images of $\mathcal{X}$, that have been taken by uncalibrated pinhole cameras, with corresponding maps $\psi_i : I \to \mathcal{Y}_i$, $\phi : I \to \mathcal{X}$ and depths $\alpha_i : I \to \mathbb{R}$, then

$$s(\phi) \subseteq s(\psi_0) \bigcap_1^{m-1} q_i s(\psi_i). \tag{1}$$

Here we have introduced $q_i = \alpha_0/\alpha_i$, which we call kinetic depths. However, in (1), it is necessary to have point correspondences, i.e. $\psi_i(t)$ is the image of $\phi(t)$ for each $t \in I$ and all i. It should actually be $\alpha s(\phi)$ on the left hand side of (1), but since ϕ is unknown, we can replace $\alpha_0 s(\phi)$ by $s(\phi)$. This freedom in replacing $\alpha_0 s(\phi)$ by $s(\phi)$ is just the non uniqueness of projective reconstructions. As reconstructions, we take the orthogonal complement of $s(\phi)$ and all other equally valid reconstructions can be obtained by multiplying both sides of (1) with functions α, such that the product is still a shape space, that is, all its elements have integral zero.

3 Applications to curve reconstruction

Assume that $\mathcal{X} \subset \mathbb{R}^3$ is an open curve of finite length, which is closed as a subset of $\mathbb{R}^3$ and is such that $\mathcal{X}$ does not belong to any affine plane. We assume that the curve is open for simplicity, as it makes the algorithm simpler when we have point correspondences between the end points in all images $\mathcal{Y}_i$, but this is also all we need. The points in between on the curves are now not assumed to have known correspondences. We also assume that the images $\mathcal{Y}_i$ are taken from general locations and not only from some affine plane.

If we would have had point correspondences and because of the generality of $\mathcal{X}$ and the viewing locations, a dimensionality argument implies that (1) becomes the equality

$$s(\phi) = s(\psi_0) \bigcap_1^{m-1} q_i s(\psi_i), \tag{2}$$

or by considering the orthogonal complement

$$s^\perp(\phi) = s^\perp(\psi_0) + \sum_1^{m-1} q_i^{-1} s^\perp(\psi_i), \tag{3}$$

where $q_i^{-1} = 1/q_i$. In fact, since $\mathcal{X}$ is a general curve, $s(\phi)$ has codimension equal to four and in the same way $q_i s(\psi_i)$ has codimension three for all i. Since these $q_i s(\psi_i)$ do not coincide the left hand side of (1) has codimension four as well.

Since we do not have point correspondences, the problem is to find parametrisations of $\mathcal{Y}_i$ (i.e. ψ_i) for obtaining point correspondences as well as the kinetic depths q_i for obtaining reconstruction.

Introduce the following notation. Let P_X be the orthogonal projection onto $s(\phi)$ and Q_i the orthogonal projection onto $q_i^{-1} s^\perp(\psi_i)$. These projection operators can explicitly be written using orthonormal bases. For example, let $\mathcal{X}$ be written in extended coordinates, that is we add the number 1 as a fourth coordinate for each point and let $\phi = (\phi_1, \phi_2, \phi_3, \phi_4)$ be a parametrisation with $\phi_4(t) = 1$. Let $\{\tilde{\phi}_1, \tilde{\phi}_2, \tilde{\phi}_3, \tilde{\phi}_4\}$ be an orthonormal basis for $\mathrm{lh}(\phi_1, \phi_2, \phi_3, \phi_4)$, where lh denotes the *linear hull* or *linear span* of a set. Then the projection operator P_X can be written

$$P_X(f) = f - \sum_{k=1}^{4} \left\langle \tilde{\phi}_k | f \right\rangle \tilde{\phi}_k \quad .$$

Note that from (3), the operator $P_X Q_i$ should be zero for every i. For that sake, introduce the *proximity measure*

$$\mu = \sum_{i=0}^{m-1} \|P_X Q_i\|_F^2 \quad ,$$

where the the operator norm F is the Frobenius norm. It can be shown that by choosing orthonormal bases $\{\tilde{\psi}_{i,1}, \tilde{\psi}_{i,2}, \tilde{\psi}_{i,3}\}$ of each $q_i^{-1} s^{\perp}(\psi_i)$, μ can be written

$$\mu = \sum_{i=0}^{m-1} \sum_{k=1}^{3} \left\| P_X \tilde{\psi}_{i,k} \right\|^2 \quad .$$

Furthermore, μ is independent of the choice of orthonormal bases, $\{\tilde{\psi}_{i,1}, \tilde{\psi}_{i,2}, \tilde{\psi}_{i,3}\}$.

4 Algorithm construction

We propose the following algorithm which is based on repeatedly finding $s(\phi)$, adjusting the kinetic depths q_i and the parametrisations of $\mathcal{Y}_i$. In the following, we let all the maps ϕ and ψ_i be in extended coordinates and introduce $Y_i = \mathrm{lh}\{\psi_{i1} q_i^{-1}, \psi_{i2} q_i^{-1}, \psi_{i3} q_i^{-1}\}$ and $X = \mathrm{lh}\{\phi_1, \phi_2, \phi_3, \phi_4\}$. These spaces are usually referred to as depth spaces, see [Spa96].

I Initialise: Choose one parametrisation ψ_i in each image curve by for example using the image based arc-length. Set $q_i(x) = 1$ and $Y_i = \mathrm{lh}(q_i^{-1}\psi_i)$, for all i.

II Update X: Keeping Y_i fixed for all i, find P_X that minimises μ.

III Update q_i: Keeping X and Y_i fixed, find q_i such that $q_i^{-1} Y_i$ minimises μ. Set $Y_i := q_i^{-1} Y_i$.

IV Update parametrisation: Keeping X and Y_i fixed, find a continuous bijection $\gamma_i : I \to I$, such that $Y_i \circ \gamma_i$ minimises μ. Set $Y_i := Y_i \circ \gamma_i$ and go to II.

It is difficult to minimise μ with respect to all parameters. It is, however, reasonably fast to solve each of the three steps II to IV approximately as will be demonstrated. Since we iterate the procedure over and over, we do not have to be very precise in each step.

Step I. Initialisation

In the experiments, each image curve $\psi_i(t) = (\psi_{i1}(t), \psi_{i2}(t), 1)$ has been parametrised using image arclength $t \in I$ so that $(\psi'_{i1})^2 + (\psi'_{i2})^2 = 1$. The depth is initially chosen as $q_i(t) = 1$ for all points in all curves and we set $Y_i = \mathrm{lh}(\psi_{i,1}, \psi_{i,2}, 1)$.

Step II. Computation of X given Y_i

Let $\{\psi_{i,1}, \psi_{i,2}, \psi_{i,3}\}$ be an orthonormal basis for the 3-dimensional linear space Y_i. From (3), the 4-dimensional linear space $X = s^{\perp}(\phi)$ is spanned by $\sum_{i=0}^{m-1} Y_i$, thus we would like to solve

$$\min_{\dim X = 4} \mu = \min_{\dim X = 4} \sum_{i=0}^{m-1} \sum_{k=1}^{3} \|P_X \psi_{i,k}\|^2 \quad . \qquad (4)$$

Remember that P_X is the orthogonal projection onto $s(\phi)$. Set

$$v = (\psi_{0,1}, \ldots, \psi_{0,3}, \ldots, \psi_{m-1,0}, \ldots, \psi_{m-1,3})$$

and introduce the positive matrix $M_1 = \{\langle v_i | v_j \rangle\}$. Let u_i be the 4 eigenvectors corresponding to the 4 largest eigenvalues of M_1 and let $\phi_i = \sum_j u_{ij} v_j$ form a basis for X. It can be shown that this X solves (4).

Step III. Computation of kinetic depths

Let $\{\psi_{i1}, \psi_{i2}, \psi_{i3}\}$ be an orthonormal basis for Y_i and let $\{\phi_1, \phi_2, \phi_3, \phi_4\}$ be an orthonormal basis for the 4-dimensional linear space $X = s^{\perp}(\phi)$. We would like to minimise

$$\sum_{k=1}^{3} \left\| P_X \psi_{ik} q_i^{-1} \right\|^2 ,$$

with $\|q_i^{-1}\| = 1$ for $i = 1, \ldots m-1$. For convenience we drop the index i for the index of the image, since each image can be treated separately. Parametrise q^{-1} using a finite basis $\{f_j\}$ according to

$$\mathbb{R}^n \ni x \to q^{-1}(x) = \sum_{j=1}^{n} x_j f_j \quad .$$

Set $\vartheta_{kj} = P_X(\psi_k f_j)$ and form the sum of positive matrices $M_2 = \sum_{k=1}^{3} \{\langle \vartheta_{kj} | \vartheta_{kl} \rangle\}$. The eigenvector of unit length, corresponding to the smallest eigenvalue of M_2, can be shown to give the desired x. We set

$$Y_i = \mathrm{lh}\{q_i^{-1}(x)\psi_{i1}, q_i^{-1}(x)\psi_{i2}, q_i^{-1}(x)\psi_{i3}\}.$$

Step IV. Reparametrise each image curve

Select orthonormal bases for Y_i and X as in step II. We want to find continuous bijections $\gamma_i : I \to I$, such that

$$\sum_{k=1}^{3} \|P_X(\psi_{ik} \circ \gamma_i)\|^2 , \quad i = 1, \dots, m-1 \qquad (5)$$

is minimised over some set of reparametrisations. Again, we drop the index i for convenience, since each image i can be treated independently. Parametrise some permissible γ by using a finite basis g_j according to

$$I \times \mathbb{R}^n \ni (t, x) \to \gamma(t, x) = t + \sum_j x_j g_j(t).$$

The function $\gamma(t, x)$ should be monotonous in the variable t for all x. By taking the derivative $\partial\gamma/\partial t$, we see that this is the case if

$$|x|^2 < \frac{1}{\max_t \sum_j |g_j(t)|^2} \quad . \qquad (6)$$

Set $\Theta_k(x) = P_X(\psi_k \circ \gamma(x))$ and linearise $\Theta_k(x)$ around $x = 0$. The x that minimises (5) can be found by using the Gauss-Newton iteration, see [Fle87]. If $|x|$ is too large, we replace x by βx so that (6) holds. Finally, we set

$$Y_i = \mathrm{lh}\left\{\psi_{i1} \circ \gamma_i(t, x), \psi_{i2} \circ \gamma_i(t, x), \psi_{i3} \circ \gamma_i(t, x)\right\} \quad .$$

5 Experimental Validation

The following experimental validation was performed. A simulation was made with 6 images of a curve as shown in Figure 1.

In each image i the curve was parametrised according to arclength $(\psi_{i1}(s), \psi_{i2}(s), 1)$. Note that the start and endpoints of these curves were assumed to be known, but that points with the same curve parameter s in different images are not necessarily in correspondence. For each curve $\mathcal{Y}_i$ a space $Y_i = \mathrm{lh}\{\psi_{i,1}, \psi_{i,2}, 1\}$ was chosen. Twenty iterations of the algorithm were then performed and after each step, the proximity measure μ was stored, and the reconstructed curve was compared with ground truth. Figure 2 shows the reconstructed curve after the first and after the 20'th iteration. Notice the relatively good alignment already after the first iteration.

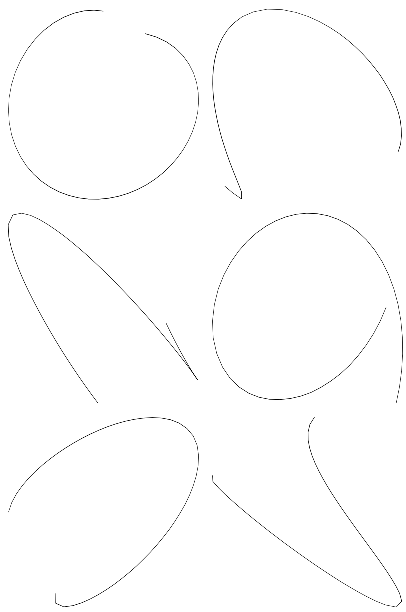

Figure 1: Six images of a curve.

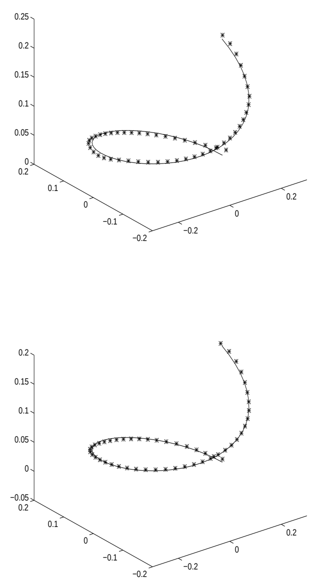

Figure 2: The reconstructed curve (*) and ground truth (-) after the first and the 20'th iteration.

A comparison was also made to the algorithm with step IV, (reparametrisation) omitted. Figure 3 shows the proximity measure after each step in the algorithm both with and without reparametrisation. Figure 4 shows the RMS residual between the reconstructed curve and the true curve as a function of iteration.

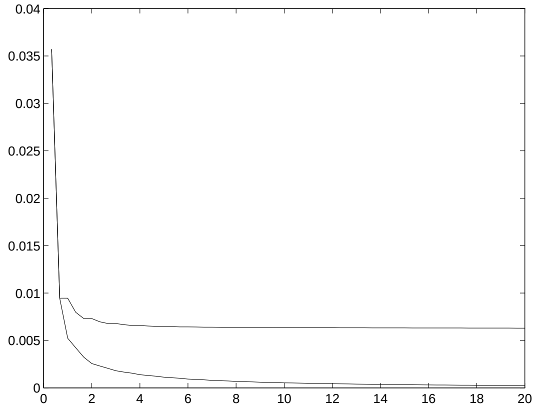

Figure 3: The proximity measure as a function of the number of iterations.

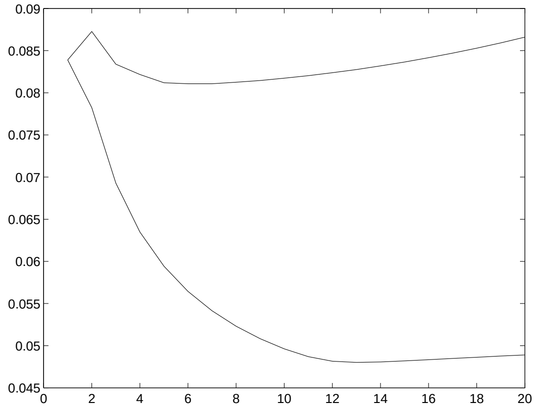

Figure 4: The RMS residual between reconstructed curve and ground truth as a function of the number of iterations.

6 Conclusions

In this paper we have presented an extension of affine shape to measurable sets and shown its applicability to the reconstruction of three-dimensional curves from its two-dimensional projections. The resulting algorithms seem to give reasonable good estimates already after a few iterations. The convergence properties of the algorithms have, however, not been examined thoroughly. This is an interesting aspect for continued research.

The idea of using proximity measures that are invariant under choice of coordinate systems is very appealing. However, if additional information are given, e.g. about the noise of the measured image curves, then this could be used to estimate structure and motion on a statistical basis. Such algorithms typically involve optimisation and can be sensitive on the choice of initial estimates. The algorithm presented in this paper seems to be efficient at giving good initial solutions. An interesting extension would be to incorporate such statistical refinements.

Another interesting and important aspect is to incorporate further experiments on real image sequences to obtain more information about the strengths and weak points of the algorithm. Although the algorithms used so far assume that the two end-points of the curve can be detected in each image, it is possible to extend the theory to deal with closed curves and partially occluded curves. This will be investigated in the future.

7 Acknowledgements

The authors thank Gunnar Sparr, Anders Heyden and Bo Bernardsson for valuable discussions.

References

[ÅCG96] K. Åström, R. Cipolla, and P. J. Giblin. Generalised epipolar constraints. In *Proc. 4th European Conf. on Computer Vision, Cambridge, UK*, 1996.

[Car94] S. Carlsson. Sufficient image structure for 3D motion and shape estimation. In J-O. Eklundh, editor, *Proc. 3rd European Conf. on Computer Vision, Stockholm, Sweden*, volume I, pages 83–91. Springer–Verlag, 1994.

[Fle87] R. Fletcher. *Practical Methods of Optimization.* John Wiley & Sons, 1987.

[FP93] O. D. Faugeras and T. Papadopoulo. A theory of the motion fields of curves. *Int. Journal of Computer Vision*, 10(2):125–156, 1993.

[PF96] T. Papadopoulo and O. Faugeras. Computing structure and motion of general 3d curves from monocular sequences of perspective images. In B Buxton and R. Cipolla, editors, *Proc. 4th European Conf. on Computer Vision, Cambridge, UK*, pages 696–708. Springer-Verlag, 1996.

[PP91] J. Porrill and S. B. Pollard. Curve matching and stereo calibration. *Image and Vision Computing*, 9(1):45–50, 1991.

[Spa96] G. Sparr. Simultaneous reconstruction of scene structure and camera locations from uncalibrated image sequences. In *13th International Conference on Pattern Recognition*, volume 1, pages 328–333, 1996.