

Reconstruction of Curves in \mathbb{R}^3 , using Factorization and Bundle Adjustment*

R. Berthilsson, K. Åström, A. Heyden
Centre for Mathematical Sciences
Lund University
P.O. Box 118, S-221 00 Lund, Sweden

12th August 1999

Abstract

In this paper, we extend the notion of affine shape, introduced by Sparr, from finite point sets to curves. The extension makes it possible to reconstruct 3D-curves up to projective transformations, from a number of their 2D-projections. We also extend the bundle adjustment technique from point features to curves.

The first step of the curve reconstruction algorithm is based on affine shape, is independent of choice of coordinates, robust, does not rely on any preselected parameters and works for an arbitrary number of images. In particular this means that a solution is given to the aperture problem of finding point correspondences between curves. The second step takes advantage of any knowledge of measurement errors in the images. This is possible by extending the bundle adjustment technique to curves.

Finally, experiments are performed on both synthetic and real data to show the performance and applicability of the algorithm.

1 Introduction

Projective reconstruction of a 3D-scene from a set of 2D-images is a central problem in computer vision. Usually a finite set of feature points are extracted, that are characteristic for the object at hand. These points are identified in the different images and a reconstruction is computed. By the algorithm proposed here, it is possible to use also 3D-curves as characteristic features in a reliable way. Furthermore, no point correspondences between the different images are needed for the curves beforehand, but are computed by the algorithm. This is a difficult problem often referred to as the correspondence problem or the *aperture problem*.

*This work has been done within the VISIT program of Swedish Foundation for Strategic Research (SSF) and ESPRIT Reactive LTR project 21914, CUMULI.

A first algorithm is presented that is independent of choice of affine coordinates, and works with any number of images, taken by uncalibrated cameras, i.e. the camera parameters are allowed to vary in an arbitrary way between imaging instants. The algorithm makes no assumptions on the image curves that are not met in practice. In fact, neither smoothness nor continuity is required. Furthermore, the images need not be taken close to each other. Experiments have also proved that it is robust to noise and converges in only a few iterations.

A second algorithm is also presented that take advantage of knowledge of the measurement errors in the images. This algorithm requires a good starting point, but have super linear convergence. For finite point configurations, the corresponding algorithm is known as bundle adjustment algorithm, in computer vision. Its first extension to curves can be found in [4]. Experiments show that the errors in the reconstruction are of the order of measurement errors in the images.

Affine shape of finite point configurations, has been introduced and studied in a series of papers by Sparr, cf. [11], [12]. It has proved to be an important tool when analyzing the geometry of cameras and scene. Instead of working with camera matrices the method is based on finding relations between linear subspaces, and the composition of two perspective transformations corresponds to an algebraically simple operation of multiplication of the depths. So far this analysis has been confined to finite point sets.

A reconstruction algorithm, by using affine shape, has been proposed in [12], [8], [5], [9]. It works for arbitrary but finite numbers of points and camera views and is based on aligning subspaces by using orthogonal projections and maximizing some of the largest eigenvalues of the sum of these projections. The algorithm presented here is based on the algorithm in [12] and an extension of affine shape of finite point configurations to curves.

Earlier versions of some of the ideas in this paper can be found in [2], where the reconstruction of synthetically generated open 3D-curves are computed. A restriction is that both endpoints have to be identifiable in all images. In [4] an algorithm for closed 3D-curves is presented, but it is assumed that the images are taken closed to each other in order to get a good initial estimate. None of these restriction is required here. Furthermore, it is also shown how occlusion can be handled, by reconstructing parts of curves. A preliminary study on the possibility to extend the bundle adjustment technique to curves can also be found in [4]. However, no solution to the problem of finding initial values are given there.

Another approach can be found in [7, 10], where the structure and motion of 3D-curves are computed. That method relies on a calibrated camera and high order spatio-temporal derivatives, which are difficult to compute in practice due to measurement errors and numerical problems.

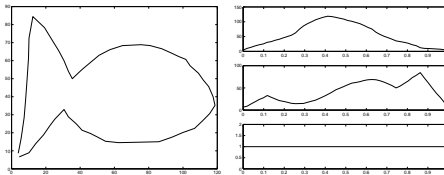


Figure 1: From left: 2D-curve $(x(t), y(t))$, and 3 basis functions $x(t)$, $y(t)$ and $1(t)$, for the affine shape of the curve

2 Affine shape of curves

In this section, we will give the necessary mathematical tools that are needed for the first reconstruction algorithm. The notion of affine shape of curves is introduced and, by using this, the condition for when a 2D-curve is a projective image of a 3D-curve is given. Furthermore, a proximity measure on the affine similarity of two curves, is introduced, which is to be minimized by the algorithm. One appealing property of these tools is that they are independent of the choice of affine coordinates in the images and space.

Let \mathcal{X} be a curve in \mathbb{R}^n , and let

$$\phi_{\mathcal{X}}(t) = (\phi_{\mathcal{X},1}(t), \dots, \phi_{\mathcal{X},n}(t)) \in \mathcal{X} \subset \mathbb{R}^n ,$$

where $t \in I = [0, 1]$, be a parametrisation of \mathcal{X} . The **affine shape** of is defined as the linear subspace of real valued functions

$$s(\phi_{\mathcal{X}}) = \left\{ f \mid \begin{array}{l} \int_I f \phi_{\mathcal{X}} dt = 0 \\ \int_I f dt = 0 \end{array} , f \in L^2(I) \right\} .$$

It can be shown, cf. [1, 3], that $s(\phi_{\mathcal{X}})$ is a complete affine invariant. Furthermore, the shape $s(\phi_{\mathcal{X}})$ is not an intrinsic property of \mathcal{X} , as it changes with different parametrisations $\phi_{\mathcal{X}}$. This is why we denote shape by $s(\phi_{\mathcal{X}})$ instead of $s(\mathcal{X})$. This ambiguity causes problems, especially when we are computing the reconstruction of a 3D-curve, from a set of 2D-images of it, and no point correspondences between the images are assumed. The orthogonal complement of $s(\phi_{\mathcal{X}})$ is called the **depth space**, $d(\phi_{\mathcal{X}})$. It is a finite dimensional space and is often easier to work with than $s(\phi_{\mathcal{X}})$. If the parametrisation $\phi_{\mathcal{X}}$ is written in extended coordinates, i.e. $\phi_{\mathcal{X}}(t) = (\phi_{\mathcal{X},1}(t), \dots, \phi_{\mathcal{X},n}(t), \phi_{\mathcal{X},n+1}(t))$, with $\phi_{\mathcal{X},n+1}(t) \equiv 1$, then

$$d(\phi_{\mathcal{X}}) = \text{linhull}\{\phi_{\mathcal{X},j}\}_{j=1}^{n+1} ,$$

i.e. the set of all linear combinations of elements from $\{\phi_{\mathcal{X},j}\}_{j=1}^{n+1}$. Figure 1 shows how a curve in \mathbb{R}^2 is represented by the depth space. We will now use $s(\phi_{\mathcal{X}})$ in the pinhole camera model. Let $\mathcal{X} \subset \mathbb{R}^3$ be a 3D-curve and let $\mathcal{Y} \subset \mathbb{R}^2$ its projective image. Fix the set $I = [0, 1]$ once and for all and let $\phi_{\mathcal{X}} : I \rightarrow \mathcal{X}$ and $\psi_{\mathcal{Y}} : I \rightarrow \mathcal{Y}$ be parametrisations in extended coordinates, such that $\psi_{\mathcal{Y}}(t)$ is the projective image of the point $\phi_{\mathcal{X}}(t)$ for each $t \in I$. Then there exists

a function $\alpha : I \rightarrow \mathbb{R}$, called the depth of $\phi_{\mathcal{X}}$ and a 3×4 matrix P (camera matrix), such that

$$\alpha(t)\psi_{\mathcal{Y}}(t) = P\phi_{\mathcal{X}}(t), \quad t \in I ,$$

which is another form of the well-known camera equation. Below we will drop the index \mathcal{X} of $\phi_{\mathcal{X}}$. We always reserve ϕ for the object curve and ψ for the image curve. It can be shown, [1], that \mathcal{Y} is the projective image of \mathcal{X} , with corresponding parametrisations ψ and ϕ respectively, and depths α as above, if and only if

$$\alpha s(\phi) \subseteq s(\psi) . \quad (1)$$

However, to use this, the correspondence between the parametrisations of \mathcal{X} and \mathcal{Y} must have been established. This is a difficult problem, referred to as the **correspondence problem** or **aperture problem**. Assuming for the moment that the correspondence problem is solved, let $\{\mathcal{Y}_i\}_0^{m-1}$ be a sequence of projective images of \mathcal{X} , taken by uncalibrated pinhole cameras, with corresponding parametrisations $\psi_i : I \rightarrow \mathcal{Y}_i$, $\phi : I \rightarrow \mathcal{X}$ and depths $\alpha_i : I \rightarrow \mathbb{R}$. Then, by (1)

$$s(\phi) \subseteq \bigcap_1^{m-1} \frac{1}{\alpha_i} s(\psi_i) . \quad (2)$$

If we assume that $\mathcal{X} \subset \mathbb{R}^3$ is a closed curve that does not belong to any affine plane and if the images \mathcal{Y}_i are taken from general locations, then (2) holds with equality, i.e.

$$s(\phi) = \bigcap_{j=0}^{m-1} \frac{1}{\alpha_j} s(\psi_j), \quad \text{or} \quad d(\phi) = \sum_{j=0}^{m-1} \alpha_j d(\psi_j) . \quad (3)$$

Let \mathbf{P}_ϕ , \mathbf{P}_j , \mathbf{Q}_ϕ and \mathbf{Q}_j be the orthogonal projections from $L^2(I)$ onto $s(\phi)$, $s(\psi_j)/\alpha_j$, $d(\phi)$ and $\alpha_j d(\psi_j)$, respectively. These projection operators can explicitly be written using orthonormal bases. Introduce the notation

$$\langle f|g \rangle = \int_0^1 f g dt, \quad \|f\|^2 = \int_0^1 |f|^2 dt .$$

Using extended coordinates, $\phi = (\phi_1, \phi_2, \phi_3, \phi_4)$, where $\phi_4 \equiv 1$, let $\{\tilde{\phi}_1, \tilde{\phi}_2, \tilde{\phi}_3, \tilde{\phi}_4\}$ be an orthonormal basis for $\text{linhull}(\phi_1, \phi_2, \phi_3, \phi_4)$. Then

$$\mathbf{Q}_\phi(f) = \sum_{k=1}^4 \langle \tilde{\phi}_k | f \rangle \tilde{\phi}_k, \quad \mathbf{P}_\phi(f) = (I - \mathbf{Q}_\phi)(f) . \quad (4)$$

It is easily seen that the conditions (3) implies that the operator $\mathbf{P}_\phi \mathbf{Q}_j$ is the zero operator for every j . Therefore, introduce the **proximity measure**

$$\mu = \sum_{j=0}^{m-1} \|\mathbf{P}_\phi \mathbf{Q}_j\|_{HS}^2 .$$

where HS denotes the Hilbert-Schmidt norm, see [6]. The HS -norm is independent of the choice of orthonormal basis. By choosing the first three basis vectors $\{e_1, e_2, e_3\}$ to span $\alpha_j d(\psi_j)$, it follows that

$$\|\mathbf{P}_\phi \mathbf{Q}_j\|_{HS}^2 = \sum_{k=1}^3 \|\mathbf{P}_\phi e_k\|^2 .$$

Thus, if $\{\tilde{\psi}_{j,1}, \tilde{\psi}_{j,2}, \tilde{\psi}_{j,3}\}$ is an orthonormal basis of $\alpha_j d(\psi_j)$, $j = 0, \dots, m-1$, then

$$\mu = \sum_{j=0}^{m-1} \sum_{k=1}^3 \|\mathbf{P}_\phi \tilde{\psi}_{j,k}\|^2 .$$

In the next section, we will construct an algorithm that solves (3), by finding parametrisations ψ_j and depths α_j , in order to obtain a reconstruction ϕ . The measure of how well this is achieved is obtained by the proximity measure μ .

3 Reconstruction of 3D-curves, using affine shape

In the following, let $I = [0, 1]$ and $\{\mathcal{Y}_j\}_0^{m-1}$ be a sequence of projective images of an unknown 3D-curve \mathcal{X} . Furthermore, let all the parametrisations $\phi : I \rightarrow \mathcal{X}$ and $\psi_j : I \rightarrow \mathcal{Y}_j$, $j = 0, \dots, m-1$, be expressed in extended coordinates. We propose the following algorithm, which is based on repeatedly finding $s(\phi)$, adjusting the depths α_j , and the curve parametrisations ψ_j . To obtain a reconstruction, the problem is to find both the parametrisations ψ_j , and the depths α_j in (3).

- I **Initialization:** Choose one parametrisation ψ_j in each image curve, for example by using the image based arc-length. Set $\alpha_j(x) \equiv 1$ and $d(\psi_j) = \text{linhull}(\psi_j)$, $j = 0, \dots, m-1$. One can also use step IV to update the parametrisation of the curves directly.
- II **Update the 3D structure of the curve, $d(\phi)$:** Keeping $d(\psi_j)$ fixed for all j , find \mathbf{P}_ϕ that minimizes μ .
- III **Update the depth parameters α_j :** Keeping $d(\phi)$ and $d(\psi_j)$ fixed, find α_j such that $\alpha_j d(\psi_j)$ minimizes μ . Set $d(\psi_j) := \alpha_j d(\psi_j)$, $j = 0, \dots, m-1$.
- IV **Update curve parametrisation τ globally:** Keeping $d(\phi)$ and $d(\psi_j)$ fixed, find a transformation τ , such that $\tau \circ d(\psi_j)$ minimizes μ . Set $d(\psi_j) := \tau \circ d(\psi_j)$, $j = 0, \dots, m-1$.
- V **Update curve parametrisation locally:** Keeping $d(\phi)$ and $d(\psi_j)$ fixed, find a continuous bijection $\gamma_j : I \rightarrow I$, such that $d(\psi_j) \circ \gamma_j$ minimizes μ . Set $d(\psi_j) := d(\psi_j) \circ \gamma_j$, $j = 0, \dots, m-1$ and go to II.

Before going into the details below, the iterative algorithm can be explained more intuitively as follows. Start by an arc-length based initial parameterization

of the curves. Then iterate by updating the structure (object parameters), i.e. $d(\phi)$, the motion (camera parameters), i.e. α_j and the correspondences (parameterization of the curves), i.e. $d(\Psi_j)$. Updating the correspondences is done in two steps. In IV the parameterization of each curve is fixed apart from a cyclic permutation of the points. Here the goal is to find correspondences between the present points on the curve. In V the parameterization of each curve is updated, giving new points on the curve to continue with. Step IV is not essential when a fairly good estimate of the correspondences has been obtained. Step IV might also be used in the initialization to find a good initial correspondence, by substituting $d(\Phi)$ with $d(\Psi_1)$.

3.1 Step I. Initialization

Let each image curve $\psi_j(t) = (\psi_{j1}(t), \psi_{j2}(t), \psi_{j3}(t))$ be parametrised using scaled image arclength $t \in I$, so that $\psi_j(0)$ and $\psi_j(1)$ are the endpoints, and such that $\sum_{k=1}^3 \{\psi'_{jk}(t)\}^2$, $j = 0, \dots, m-1$, is constant. Initially, let the depths be $\alpha_j(t) \equiv 1$ for all points in all curves, and let $d(\psi_j) = \text{linhull}(\psi_{i1}, \psi_{i2}, \psi_{i3})$.

3.2 Step II. Computation of $d(\phi)$ given $d(\psi_j)$

Let $\{\psi_{j1}, \psi_{j2}, \psi_{j3}\}$ be an orthonormal basis for the 3-dimensional linear space $d(\psi_j)$. By (3), the 4-dimensional linear space $d(\phi)$, corresponding to the 3D-curve to be reconstructed, is the linear span of all basis functions $\psi_{j,k}$, $j = 0, \dots, m-1$, $k = 1, 2, 3$, i.e. $d(\phi) = \text{linhull}\{\psi_{j,k}\}$. An estimate of $d(\phi)$ is obtained by solving

$$\min_{\dim d(\phi)=4} \mu = \min_{\dim d(\phi)=4} \sum_{j=0}^{m-1} \sum_{k=1}^3 \|\mathbf{P}_\phi \psi_{j,k}\|^2 . \quad (5)$$

This optimization problem can be solved using singular value decomposition. Form the symmetric matrix $M_1 = \{v_l | v_m\}_{l,m}$, where

$$v = (\psi_{0,1}, \psi_{0,2}, \psi_{0,3}, \psi_{1,1}, \dots, \psi_{m-1,3}) .$$

Compute a singular valued decomposition $M_1 = USV^T$. In the case of exact data, the matrix M_1 has rank 4. In the case of measured data, the matrix which is closest in Frobenius norm to a matrix of rank 4, is $\hat{M} = US_4V^T$, where S_4 is obtained by setting all but the four largest diagonal elements in S to zero. An orthonormal basis for $d(\phi)$ is then $\phi_k = S_{k,k}^{-1/2} \sum_l v_l V_{l,k}$, $k = 1, 2, 3, 4$. This $d(\phi)$ solves the optimization problem (5).

3.3 Step III. Computation of depths

Let $\{\psi_{j1}, \psi_{j2}, \psi_{j3}\}$ be an orthonormal basis for $d(\psi_j)$ and let $\{\phi_1, \phi_2, \phi_3, \phi_4\}$ be an orthonormal basis for the 4-dimensional linear space $d(\phi)$, corresponding to

the curve to be reconstructed. Recall the projection operator $\mathbf{P}_\phi : L^2(I) \rightarrow s(\phi)$ in (4). We want to solve

$$\min_{\|\alpha_j\|=1} \sum_{k=1}^3 \|\mathbf{P}_\phi \psi_{jk} \alpha_j\|^2, \quad j = 0, \dots, m-1. \quad (6)$$

For convenience, the index j is dropped below, since each image is treated separately. Parametrise α using a finite orthonormal basis f_l according to

$$I \times \mathbb{R}^n \ni (t, x) \rightarrow \alpha(t, x) = \sum_{l=1}^n x_l f_l(t) \quad (7)$$

and let $\vartheta_{k,l} = \mathbf{P}_\phi(\psi_k f_l)$. In the case of non-exact data, the solution of (6) can be found by singular value decomposition USV^T of the matrix $M_2 = \sum_{k=1}^3 \{\langle \vartheta_{k,l} | \vartheta_{k,m} \rangle\}_{l,m}^n$. By taking x as the last column of V we obtain the vector x of unit length which inserted in (7) gives α_j . Set $d(\psi_j) = \text{linhull}\{\alpha_j \psi_{j1}, \alpha_j \psi_{j2}, \alpha_j \psi_{j3}\}$.

3.4 Step IV. Global reparametrisation

This step is divided into two parts. The first is used for closed curves and consists of finding a cyclic translation and the second part, that is used for parts of curves, consists of finding an affine transformation of the curve parametrisation. A closed curve can loosely be characterized as having start point equal to end point. For open curves, where the endpoints are detectable in all images, this step IV can be omitted.

Closed curves:

Let $\{\psi_{j1}, \psi_{j2}, \psi_{j3}\}$ be an orthonormal basis for $d(\psi_j)$ and $\{\phi_1, \phi_2, \phi_3, \phi_4\}$ an orthonormal basis for $d(\phi)$. The objective is to find a cyclic translation τ_{t_j} , where $\tau_t \circ f(x) = f(x-t)$, such that

$$f_j(t) = \sum_{k=1}^3 \|\mathbf{P}_\phi \tau_t \psi_{j,k}\|^2$$

is minimized for each $j = 0, \dots, m-1$. Since $\mathbf{P}_\phi = I - \mathbf{Q}_\phi$ and \mathbf{Q}_ϕ is a projection, it follows that

$$f_j = \sum_{k=1}^3 \|\tau_t \psi_{j,k} - \mathbf{Q}_\phi \tau_t \psi_{j,k}\|^2 = 3 - \sum_{\substack{1 \leq k \leq 3 \\ 1 \leq l \leq 4}} \langle \phi_l | \tau_t \psi_{j,k} \rangle^2.$$

Let $\check{\psi}(x) = \psi(-x)$, then $\langle \phi_l | \tau_t \psi_{j,k} \rangle = \phi_l * \check{\psi}_{j,k}(t)$, where $*$ denotes cyclic convolution. If ψ_{jk} and ϕ_l are sampled, with the number of samples being 2^n for some integer n , then f_j can be computed very fast by using the fast Fourier transform. Step IV can also be used in Step I in order to get a better initial parametrisation of the curves. In this case one compares the shape of each image curve with the first image curve.

Parts of curves:

Let $\{1, \psi_{j2}, \psi_{j3}\}$ be functions on \mathbb{R} that form a basis for $d(\psi_j)$, under the measure

$$d\mu(t) = \begin{cases} dt, & t \in [0, 1], \\ 0, & \text{otherwise} . \end{cases}$$

Note that the first function is a constant function and that the basis is not necessarily orthogonal. As above let $\{\phi_1, \phi_2, \phi_3, \phi_4\}$ be an orthonormal basis for $d(\phi)$. The objective is to find an affine transformation $a_{\alpha,t} : \mathbb{R} \rightarrow \mathbb{R}$, defined by $a_{\alpha,t}(x) = \alpha x + t$, such that

$$f_j(\alpha, t) = \|\mathbf{P}_\phi \mathbf{P}_{d(\psi_j) \circ a_{\alpha,t}}^\perp\|^2$$

is minimized for each $j = 0, \dots, m-1$. A problem is that even if $\{1, \psi_{j2}, \psi_{j3}\}$ happens to be an orthonormal basis for $d(\psi_j)$, the composed functions $\{1, \psi_{j2} \circ a_{\alpha,t}, \psi_{j3} \circ a_{\alpha,t}\}$ may be far from being an orthonormal basis. However, a Hilbert-Schmidt like orthogonalization process gives that $e_1 = 1$,

$$e_2 = \frac{\psi_1 - \langle \psi_1, 1 \rangle}{\sqrt{\|\psi_1\|^2 - \langle \psi_1, 1 \rangle^2}} - \frac{\psi_2 - \langle \psi_2, 1 \rangle}{\sqrt{\|\psi_2\|^2 - \langle \psi_2, 1 \rangle^2}}$$

and

$$e_3 = \frac{\psi_1 - \langle \psi_1, 1 \rangle}{\sqrt{\|\psi_1\|^2 - \langle \psi_1, 1 \rangle^2}} + \frac{\psi_2 - \langle \psi_2, 1 \rangle}{\sqrt{\|\psi_2\|^2 - \langle \psi_2, 1 \rangle^2}}$$

form an orthogonal basis for $d(\psi)$, with

$$\|e_2\|^2 = 2 - 2 \frac{\langle \psi_1, \psi_2 \rangle - \langle \psi_1, 1 \rangle \langle \psi_2, 1 \rangle}{\sqrt{\|\psi_1\|^2 - \langle \psi_1, 1 \rangle^2} \sqrt{\|\psi_2\|^2 - \langle \psi_2, 1 \rangle^2}}$$

and

$$\|e_3\|^2 = 2 + 2 \frac{\langle \psi_1, \psi_2 \rangle - \langle \psi_1, 1 \rangle \langle \psi_2, 1 \rangle}{\sqrt{\|\psi_1\|^2 - \langle \psi_1, 1 \rangle^2} \sqrt{\|\psi_2\|^2 - \langle \psi_2, 1 \rangle^2}}.$$

It follows that

$$\begin{aligned} f_j(\alpha, t) &= \sum_{k=1}^3 \left\| \frac{e_{j,k} \circ a_{\alpha,t}}{\|e_{j,k} \circ a_{\alpha,t}\|} - \mathbf{Q}_\phi \frac{e_{j,k} \circ a_{\alpha,t}}{\|e_{j,k} \circ a_{\alpha,t}\|} \right\|^2 \\ &= 3 - \sum_{\substack{1 \leq k \leq 3 \\ 1 \leq l \leq 4}} \left\langle \phi_l \left| \frac{e_{j,k} \circ a_{\alpha,t}}{\|e_{j,k} \circ a_{\alpha,t}\|} \right. \right\rangle^2, \end{aligned}$$

where $\{e_{j,k}\}_{k=1}^3$ is an orthogonal basis for $d(\psi_j)$ as above. All the involved scalar products can be computed as a function of t , by using FFT, which makes the computation of f_j of complexity $O(MN \log(N))$, if the α :s are sampled at M values and the t :s at N values.

3.5 Step V. Local reparametrisation

Let $\{\psi_{j1}, \psi_{j2}, \psi_{j3}\}$ be an orthonormal basis for $d(\psi_j)$, and let $\{\phi_1, \phi_2, \phi_3, \phi_4\}$ be an orthonormal basis for the 4-dimensional linear space $d(\phi)$, corresponding to the curve to be reconstructed. We want to find a reparametrisation γ_j in each image $j = 0, \dots, m-1$, such that

$$\sum_{k=1}^3 \|\mathbf{P}_\phi(\psi_{jk} \circ \gamma_j)\|^2$$

is minimized over some set of reparametrisations. Again, we drop the index j for convenience. Parametrise γ by using a finite basis g_l , according to

$$I \times \mathbb{R}^n \ni (t, x) \rightarrow \gamma(t, x) = t + \sum_{l=1}^n x_l g_l(t) , \quad (8)$$

where the basis function fulfill $g_l(0) = 0$ and $g_l(1) = 0$. g_l can for example be a translated and dilated Gaussian function multiplied $\sin(2\pi x)$ in order to fulfill $g_l(0) = g_l(1) = 0$. The function $\gamma(t, x)$ is monotonic for small x , that is

$$\frac{\partial \gamma}{\partial t} = 1 + \sum_{l=1}^n x_l g'_l(t) > 0, \quad t \in I ,$$

if $|x|$ is sufficiently small. Set $\Theta_k(x) = \mathbf{P}_\phi(\psi_k \circ \gamma(x))$ around $x = 0$. Then $\Theta_k(x) \approx \Theta_k(0) + \nabla_x \Theta_k(0)x$, where ∇_x is the gradient operator in the x -variables. We use this approximation and Gauss-Newton iterations to solve the minimization problem, i.e. finding x in (8). We set $d(\psi_j) = \text{linhull}\{\psi_{j1} \circ \gamma_j(x), \psi_{j2} \circ \gamma_j(x), \psi_{j3} \circ \gamma_j(x)\}$.

4 Bundle adjustment for curves

A bundle adjustment algorithm was developed for estimating all unknown parameters. This will briefly be described here. Let m denote the number of images and n the number of points estimated on the reconstructed curve. Let $\mathcal{Y}_i(s)$ denote the estimated curve in image i with curve parameter s and let s_{ij} denote the curve parameter corresponding to the projection of point j in image i . Denote by \mathbf{m} the bundle of all unknown parameters, $\mathbf{m} = \{P_1, \dots, P_m, \mathbf{X}_1, \dots, \mathbf{X}_n, s_{11}, \dots, s_{mn}\}$. Each such element belongs to a non-linear manifold, \mathcal{M} . The goal of the bundle adjustment is to find the parameters \mathbf{m} that minimizes

$$f(\mathbf{m}) = \sum_{i=1}^m \sum_{j=1}^n \|P_i \mathbf{X}_j - \mathcal{Y}(s_{ij})\|^2$$

This will be done by an iterative method. One such iteration will now be described. Assume that an initial estimate \mathbf{m}_0 is given. Introduce a *local parametrisation* $\mathbf{m}(\Delta \mathbf{x})$, around $\mathbf{m}_0 \in \mathcal{M}$ according to

$$\mathcal{M} \times \mathbb{R}^N \ni (\mathbf{m}_0, \Delta \mathbf{x}) \mapsto \mathbf{m}(\mathbf{m}_0, \Delta \mathbf{x}) \in \mathcal{M} ,$$

where $N = 9m + 3n + m * n$. (11 parameters in each camera matrix, 3 parameters for the coordinates of each reconstructed point and 1 parameter for each curve parameter s_{ij}). Let $\Delta \mathbf{x} = [\Delta a_1, \dots, \Delta a_m, \Delta b_1, \dots, \Delta b_n, \Delta s_{11}, \dots, \Delta s_{mn}]^T$, so that Δa_i parametrise changes in camera matrix P_i and Δb_j parametrise changes in reconstructed point X_j and Δs_{ij} parametrise changes in curve parameter s_{ij} . Changes in each camera matrix, P_i , are parametrised by

$$P_i(\mathbf{m}_0, \Delta \mathbf{x}) = P_i + \begin{pmatrix} \Delta a_i(1) & \Delta a_i(2) & \Delta a_i(3) & \Delta a_i(4) \\ \Delta a_i(5) & \Delta a_i(6) & \Delta a_i(7) & \Delta a_i(8) \\ \Delta a_i(9) & \Delta a_i(10) & \Delta a_i(11) & 0 \end{pmatrix} ,$$

changes in each object point, \mathbf{X}_j , by

$$\mathbf{X}_j(\mathbf{m}_0, \Delta \mathbf{x}) = \begin{pmatrix} X_j + \Delta b_j(1) \\ Y_j + \Delta b_j(2) \\ Z_j + \Delta b_j(3) \\ 1 \end{pmatrix} ,$$

and changes in each curve parameter, s_{ij} , by

$$s_{ij}(\mathbf{m}_0, \Delta \mathbf{x}) = s_{ij} + \Delta s_{ij} .$$

Introduce a *residual vector* \mathbf{Y} , formed by putting all reprojected errors weighted by estimated standard deviations in a column vector

$$\mathbf{Y} = [e_{11} \dots e_{mn}]^T ,$$

where

$$e_{ij} = \frac{\lambda_{ij} P_i \mathbf{X}_j - \mathcal{Y}(s_{ij})}{\sigma_n} ,$$

where as usual λ_{ij} is chosen to normalize the length of the image vector $P_i \mathbf{X}_j$. These residuals depend on our measured image curves \mathcal{Y}_i and on our estimated parameters \mathbf{m} . The residual vector $\mathbf{Y}(\Delta \mathbf{x})$ is a non-linear function of the local parametrisation vector $\Delta \mathbf{x}$. The sum of squared residuals $f = \mathbf{Y}^T \mathbf{Y}$ was minimized with respect to the unknown parameters $\Delta \mathbf{x}$, using the Gauss-Newton method. Let

$$A = \frac{\partial \mathbf{Y}}{\partial \Delta \mathbf{x}}(0), \quad b = \mathbf{Y}(0) .$$

Instead of taking

$$\Delta \mathbf{x} = -(A^T A)^{-1} A^T b , \tag{9}$$

which might be numerically sensitive if $(A^T A)$ has small singular values one uses the update

$$\Delta \mathbf{x} = -(A^T A + \epsilon I)^{-1} A^T b ,$$

where ϵ is a ‘small’ positive number.

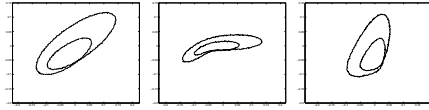


Figure 2: From left: Image 1,2, and 3, with added white normal noise with standard deviation $\sigma = 10^{-3}$.

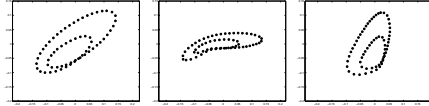


Figure 3: Image curves together with backprojected curve.

5 Reconstruction using bundle adjustment for curves

5.1 Simulated curves

In the first experiment, closed 3D curves were generated by Monte-Carlo simulations together with 3 camera matrices. This data was used to obtain 3 corresponding image curves for each 3D-curve. The image curves were parametrised by arc-length with random starting point on the curve. This gives that there are no known point correspondences beforehand. Furthermore, white normally distributed noise was added to the image curves with varying degrees of standard deviation σ . Figure 2 shows 3 corresponding image curves when $\sigma = 1.0 \cdot 10^{-3}$. This noise corresponds to approximately one pixel i standard deviation for a 350×400 pixel image. The image curves were represented by b-splines which also gives a little smoothing. Figure 3 shows the b-spline represented curves (o) together with the reprojection of the reconstructed curve (+). The points appear to coincide. Figure 4 shows a close up of the left image in Figure 3. Note that the image points (o) and reprojected points (+) still appears to coincide, indicating the precision and robustness of the method. Figure 4 right shows a projective reconstruction of the object curve. Figure 5 shows the standard deviation of the residuals from all the experiments as a function of the standard deviation of the added noise to the images curves. Finally, the algorithm has a record of converging 100 % of the time for this kind of experiments, independently of the added noise, ranging from $\sigma = 0$ to $\sigma = 3.0 \cdot 10^{-3}$.

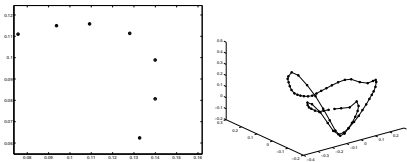


Figure 4: From left: Close up of first image, and a projective reconstruction.

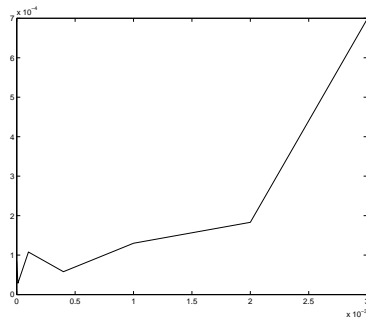


Figure 5: Standard deviation of residuals as a function of the standard deviation of added white normal noise to the image curves.

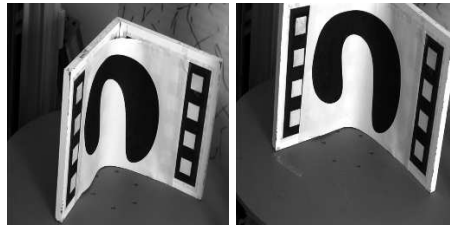


Figure 6: The first and last image in a controlled experiment of 100 images where the curve was rotated 0.3 degrees per frame and translated 1 mm per frame.

5.2 Images of real curves

We have also checked our algorithms on real data in a carefully controlled situation. A three dimensional curve was placed on a table which can be rotated and translated in a well defined manner. One hundred images were then taken as the table was rotated at 0.3 degrees per frame with a velocity of 1 mm per frame. The first and last of these images are shown in figure 6. The curves in these one hundred images were extracted with sub-pixel accuracy. These were then fed to the shape based factorization algorithm. After 20 iterations the difference between measured and reprojected curves were in the order of X pixels. This intermediate result was then fed into the bundle adjustment routine assuming general projective cameras. Then followed autocalibration (assuming constant intrinsic parameters) and finally bundle adjustment (again assuming constant intrinsic parameters). In the final result the difference between measured and reprojected curves were in the order of X pixels. In figure 7 is shown the reconstructed 3D curve as well as the focal point of the camera relative to the curve during all of these 100 images. In order to validate this approach the estimated rotational velocity was compared to the ground truth (0.3 degrees per frame). This comparison is shown in figure 8.

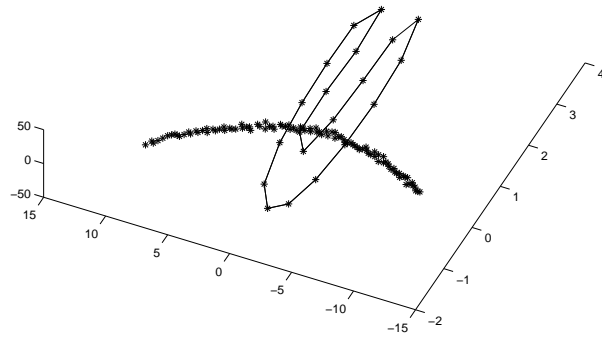


Figure 7: The reconstruction of the space curve and the focal points of the cameras using the 100 images of the rotating table.

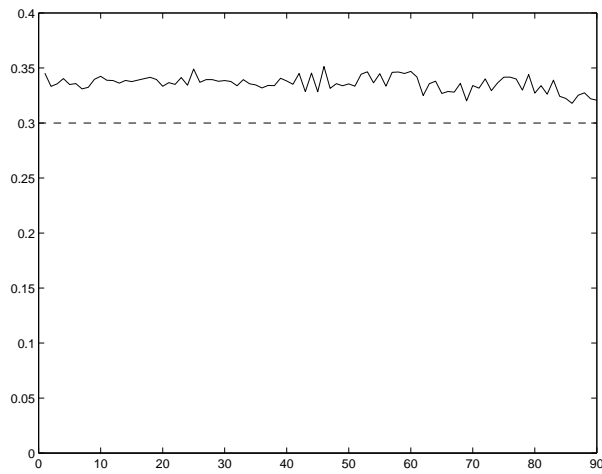


Figure 8: The estimated rotational velocity compared to the nominal 0.3 degrees per frame.

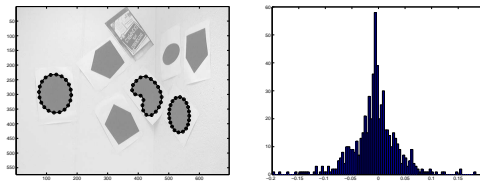


Figure 9: One of five images used in experiment. The extracted curves and 20 reprojected points on each space curve are shown. To the right is shown a histogram of the residuals (in pixels) between extracted image curves and reprojected points of the reconstructed curve. The standard deviation of these residuals is 0.03 pixels.

5.3 Verification with point based methods

Experiments with real data and verification of the method using traditional point based methods was performed as follows. Five images containing curves and easily extracted corner points were taken. One such image is shown in Figure 9. Fifteen points were extracted with high precision from each image and point based methods were used to calculate camera motion (and 3D structure of the points) using bundle adjustment. Euclidean reconstruction was obtained using the assumption of constant intrinsic parameters.

Independent of this, three curves were extracted in each image. Initial estimates, auto-calibration and iterative refinement using bundle adjustment for curves was computed as described in the paper. Figure 9 also show the extracted curves in one of the images as well as 20 reprojected points from each of the three 3D curves. A histogram of the errors between these reprojected points and the measured image curves is also shown in Figure 9. The standard deviation of these errors is approximately 0.03 pixels.

The resulting 3D structure of the curves as well as the focal points of the camera for the five images is shown in figure 10. In the same figure is shown the focal points of the five cameras as computed independently using the point features. As can be seen in the figure, these two set of focal points overlap.

References

- [1] R. Berthilsson. Extension of affine shape. Technical report, Dept of Mathematics, Lund Institute of Technology, 1997. Licentiate Thesis, www.maths.lth.se/matematiklth/personal/rikard/index.html.
- [2] R. Berthilsson and K. Åström. Reconstruction of 3d-curves from 2d-images using affine shape methods for curves. In *Computer Vision and Pattern Recognition*, 1997.
- [3] R. Berthilsson and K. Åström. Reconstruction of 3d-curves from 2d-images using affine shape methods for curves. *Journal of Mathematical Imaging and Vision*, 1999. To appear.

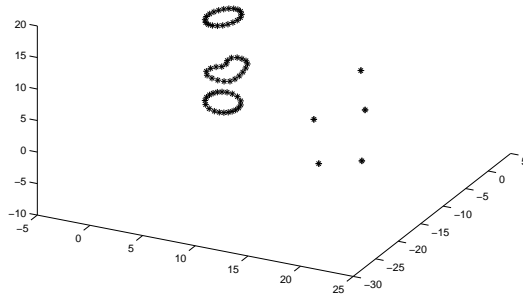


Figure 10: Reconstruction of curves and the 5 camera positions from curve reconstruction + and point based reconstruction *. The five camera centers overlap in the image.

- [4] R. Berthilsson, K. Åström, and A. Heyden. Projective reconstruction of 3d-curves from its 2d-images using error models and bundle adjustments. In *Scandinavian Conference on Image Analysis*, pages 581–588, 1997.
- [5] R. Berthilsson, A. Heyden, and G. Sparr. Recursive structure and motion from image sequences using shape and depth spaces. In *Proc. Conf. Computer Vision and Pattern Recognition*, pages 444–449. IEEE Computer Society Press, 1997.
- [6] N. Dunford and J. T. Schwarz. *Interscience. Linear operators. Spectral theory*, 2, 1963.
- [7] O. D. Faugeras and T. Papadopoulos. A theory of the motion fields of curves. *Int. Journal of Computer Vision*, 10(2):125–156, 1993.
- [8] A. Heyden. Projective structure and motion from image sequences using subspace methods. In *Proc. 10th Scandinavian Conference on Image Analysis*, pages 963–968, 1997.
- [9] A. Heyden, R. Berthilsson, and G. Sparr. An iterative factorization method for projective structure and motion from image sequences. *Image and Vision Computing*, 1999.
- [10] T. Papadopoulos 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.
- [11] G. Sparr. Structure and motion from kinetic depth. Technical Report ISRN LUTFD2/TFMA-95/7016-SE, Dept of Mathematics, Lund Institute of Technology, 1995. To appear in Proc of the Sophus Lie International Workshop on Computer Vision and Applied Geometry, Nordfjordeid, Norway, 1995.
- [12] G. Sparr. Simultaneous reconstruction of scene structure and camera locations from uncalibrated image sequences. In *Proc. International Conference on Pattern Recognition, Vienna, Austria*, 1996.