Structure From Translational Observer Motion

Ambjörn Naeve
Computational Vision and Active Perception Laboratory
Department of Numerical Analysis and Computing Science
Royal Institute of Technology, S-100 44 Stockholm, Sweden

ABSTRACT:

This work presents a unified, globally based geometric framework, using congruence geometry, for the description and computation of structure from motion. It is based on projectively invariant tangent information in a sequence of monocular images, i.e. occluding contours under general perspective. The strength of the framework is demonstrated by applying it to the case of *translational* observer motion, a type of motion that is of great practical importance since it can be easily implemented with the help of various gyroscopic devices.

Starting with a brief overview of congruence geometry, we show how it can be applied to vision - leading to a classification of the global line-geometric structure of different target-observer configurations. Then we illustrate how this approach facilitates explicit book-keeping of what information that is available in any specific situation, making it possible to parametrise what is "knowable" and what is "unknowable" from any given observational sequence of images. It also allows a consistent treatment of degenerate target shape and observer motion by supporting globally based discrimination of point-like, curve-like and surface-like targets, as well as detection of rectilinear observer motion.

We then introduce a simple technique for the computation of the direction of motion ("Focus Of Expansion") as a function of time. In this way we reduce the recovery of translational observer motion to a problem of determining its speed. From such speed information, we show how to reconstruct the observer motion - as well as a set of silhouette curves on the observed target - and illustrate with a few simulated examples. Applying this FOE-reconstruction techinque to a stationary curve target, we present an algorithm for matching the individual points on the curve between the consecutive time frames of its recorded silhouette, *without any speed information at all*.

The FOE-reconstruction method is then generalized from the real to the complex domain, showing how to combine conjugate complex geometric elements in order to obtain real geometric information concerning the direction of observer motion. We conclude by applying this method to real image data.

KEYWORDS: Translational observer motion, Line congruence , Focal surface, Computing the Focus Of Expansion, Point matching on curves, Global structure from motion.

1 Introduction

The problem of reconstructing geometric information in a scene from a sequence of mobile observations has been intensely studied within the computer vision community. Such information - notably target structure and observer motion - can be used for the purpose of object recognition and for various forms of geometric reasoning, e.g. motion planning and obstacle avoidance. Within this context, structure and motion from the outlines of space-curves or curved surfaces constitutes a difficult problem that has attracted increased attention. Among the early contributions to this field, Koenderink [(1), (2)] has related the curvature of an apparent contour to the Gaussian curvature of the observed surface; Giblin and Weiss [(3)] have worked on the positioning of objects from their occluding contours, assuming planar camera motion - a restriction that was removed by Vaillant in [(4)]. Among the later work we mention Cipolla [(5)] and Cipolla and Blake [(6)], who showed how to recover the geometry of the target surface from the deformation of its apparent contours. They assumed that the observer motion is known and used the so called epipolar parametrization, which is based on the classical conjugacy condition between a contour generator (rim) of a surface and the corresponding tangential ray of sight. Related work has been done by Vaillant and Faugeras [(7)], who used three closely related images of the surface rim - taken under known observer motion - in order to recover the local curvature properties of the surface up to second order. Also, for the case of rigid curve targets, Faugeras and Papadopoulo [(8)] have shown how to use the spatio-temporal image of a space curve in order to derive a constraint on the viewer motion from second order spatio-temporal derivatives. Giblin and Weiss [(9)] have given criteria for the breakdown of the epipolar parametrization and Cipolla, Åström and Giblin [(10)] have used such points (which they call frontier points) in order to reconstruct the observer's motion from the occluding surface contours. Kutulakos and Dyer [(11)] have formulated rules to control observer motion so as to connect local surface patch reconstructions in a global reconstruction process of the surface. Zhao and Mohr [(12)] have used B-spline surface patches, reconstructed with the aid of the epipolar parametrization, in order to achieve global reconstruction under known observer motion.

The underlying theme of these developments is to use locally based (i.e. differential) methods in order to achieve reconstruction under more and more general forms of observer motion. In this paper we advocate a different approach. Restricting the observer motion to *translations*, we use a classical line geometric framework (notably congruence geometry) as a means to describe the geometry of the various target-observer configurations. Our framework is based on the *<Target|Observer> congruence*, which was introduced by Naeve in [(13)]. This is a type of congruence (i.e. two-parametric family of lines) that appears naturally in scene reconstruction problems in vision. It corresponds to a sequence of observations of a curve- or surface-like target by a moving, monocular observer. This congruence can be decomposed into two families of *developables* that carry information concerning the global geometry of the target surface as well as the motion of the observer. A great advantage of this approach is that it does not make any a priori assumptions about the shape of the observed target, thereby allowing a uniform treatment of various types of targets - such as lines and curves, as well as polyhedral and smoothly curved surfaces.

The restriction to translational motion is natural since it is performed by biological observers in many situations - notably by hunting predators. When e.g. a cheetah is hunting, it is easy to observe how every muscle in its body cooperates in order to keep the motion of its head translational - a fact which greatly facilitates the computation of the relative direction of motion of its prey. Moreover, such translationally balanced motion can be easily implemented in machine vision systems by the help of gyroscopic devices, as demonstrated e.g. by the so called "steadycam" cameras systems that are becoming increasingly popular within the television industry.

2 Background

In this chapter we give a brief presentation of some important concepts of classical line geometry - notably the *complex*, the *congruence* with its *focal surface*, the *ruled surface* and the *developable* (*surface*). These concepts were once well known and much used within the mathematical community, and they form the geometrical basis upon which our theory rests. The reader is referred to Naeve [(13)] or Eisenhart [(14)] for more information on this subject.

A smoothly parametrized family of lines in 3d-space is called a *ruled surface*, a *congruence* or a *complex*, if it depends respectively on 1, 2, or 3 parameters. Examples of these three types are given e.g. by the generator lines of a *1-sheeted hyperboloid*, the *normals to an arbitrary surface* and the *tangents to an arbitrary surface*. These examples are all special in the sence that they are highly non-generic: A 1-sheeted hyperboloid can be considered as a ruled surface in *two* different ways, which is characteristic of *quadric* surfaces, the normals to a surface form a special type of congruence, called a *normal congruence*, and the tangents to a surface form a special type of complex, called a *tangent complex*. The complex of lines that intersect a fixed curve is called a *secant complex* - unless the curve itself is a line. In this case, the complex of intersecting lines is known as a *singular complex*.

A generic ruled surface is called *skew*. This term refers to the fact that, in the case of a skew surface, the magnitude of the closest distance between neighboring generators is of the first order, i.e. the same order as the parametric difference itself. We say that *neighboring generators of a skew surface do not intersec*t. Now, it is an interesting - and somewhat surprising - fact, that when this order changes, it suddenly jumps from one to (at least) three. When the closest distance between neighboring generators is of the third order (or higher), one says that *neighboring generators intersect*, and in this case the ruled surface is called *developable*. This terminology can be somewhat confusing, since the term developable in general refers to a (*smooth*) *1-parameter family of planes*. These two modes of usage are, however, non-conflicting, since on the one hand, the envelope of each 1-parameter family of planes can be considered as generated by the 1-parameter family of lines that form their *characteristics*, i.e. *intersections of neighboring elements*, while on the other hand, the lines of a 1-parameter family with intersecting neighbors are all tangent to a certain curve (see below), and therefore must generate the same surface as the 1-parameter family of osculating planes to this curve.

As a surface, a developable consists of two differentiable parts, joined together by a sharp *edge curve*, along which the surface does *not* possess a tangent plane. The lines of the developable are all tangent to the curve describing this edge. It is called the *edge of regression* of the devel-

opable. It can be easily shown [(14)], that if the closest distance between neighboring generators is of order *higher* than three, then the edge of regression of the corresponding developable is a *plane* curve.

A generic (= skew) ruled surface is (locally) smooth without any sharp, non-differentiable edge curves. However, on each ruled surface there is a *striction curve*, which is the geometric locus of the footpoint on each generator, corresponding to its closest distance to its neighbors. If a skew surface is continously deformed until it becomes developable, then its striction curve is transformed into the edge of regression of the final result.

Two important special types of developables are given by the *cone* and the *cylinder*. A cone is a developable whose edge of regression has contracted to a point, called the *vertex*, and a cylinder is a cone whose vertex is located at *infinity*. Hence the difference between a cone and a cylinder is of a purely *metric* nature.

A *congruence* is a smooth *2-parameter family of lines*. In general, these lines envelope two surface-patches, S_1 and S_2, known as the two *focal sheets* of the congruence. Taken together they constitute its *focal surface*. A congruence can be represented by its two focal sheets, in which case we denote the congruence by $<S_1|S_2>$. The lines of the congruence $<S_1|S_2>$ are precisely the lines that are tangents to both S_1 and S_2. The focal sheets of a congruence can degenerate in various ways. If we consider "dimensional degenerations" of the surface as a *point manifold*, each *surface* patch can shrink and degenerate into either a *curve* or a *point*. It is easy to see that if one of the focal sheets has become a point, then the other focal sheet is forced to degenerate into the same point, if the lines of the degenerated configuration are to remain a congruence. This creates a <*Point*|*Point*> congruence better known as a *star*, i.e. the set of lines on a point. If one focal sheet is non-degenerated, while the other has degenerated into a curve, we have a <*Surface*|*Curve*> congruence, which is call (*singly*) *directed*. In the case of a <*Curve*|*Curve*> congruence, both focal sheets are curves, and the congruence is called *doubly directed*.

Below, we will have reason to consider some of the *dual* types of degenerations to the ones presented above, i.e. dimensional degenerations of the focal surface as a *plane manifold*. Consider a degeneration process that transforms a surface (patch) into a curve (segment). In this degeneration process, a 2-parameter family of *points* is collapsed into either a 1-parameter family of *points* (curve) or a single *point*. Hence, the *dual degeneration process* can be described by the following procedure: A 2-parameter family of *planes* is collapsed into either a 1-parameter family of *planes* or a single *plane*. In the latter case, the corresponding <*Plane*|*Plane*> congruence is better known as the set of lines in a plane, or a *plane system* of lines. In the former case, the corresponding focal sheet degenerates into a *developable*. Such a <*Surface*|*Developable* > congruence is called (*singly*) *developed*, while a <*Developable*|*Curve*> congruence is termed *developed directed*. Finally, in agreement with the above established terminology, a <*Developable*|*Developable*> congruence is called *doubly developed*.

As a motivation for these subtypes, we will see below, that an example of a singly directed congruence appears naturally in vision as what we call the <*Target*|*Observer*> *congruence*, corresponding to a sequence of observations of a moving, surface-like target by a moving, point-like observer. If the observer is looking at a *stationary curve* instead of at a moving curve or a sur-

face patch, then the corresponding <Target|Observer> congruence is in fact *doubly directed*. The developed directed type of congruence also enters naturally into the considerations below. It is the subtype of the <Target|Observer> congruence, that corresponds to a moving line-like target, which is recorded by a moving point-like observer.

2.1 Representing a Congruence by its Developables

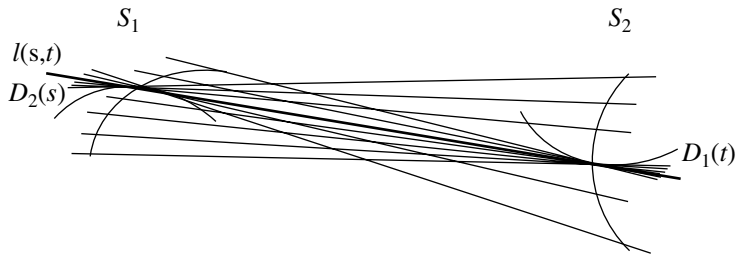


Figure 1

Consider a congruence $<S_1|S_2>$ which is neither a star, nor a plane system[1]. Then it is a fact of fundamental importance [see ...] that its lines can be *uniquely* partitioned into two 1-parameter families of ruled surfaces $\{D_1(t) : t_1<t<t_2\}$ and $\{D_2(s) : s_1<s<s_2\}$ in such a way that each member surface is *developable*. Hence, each line $l(s, t)$ of the given congruence is the intersection of one developable from each of the two families:

$$l(s,t) = D_1(t) \cap D_2(s) \ . \tag{1}$$

This configuration is depicted in Figure (1). Geometrically, a *developable* is the locus of a *1-parameter family of planes*. Hence, each developable D_i has a natural representation given by

$$Developable = \bigcup_x Plane(x) \tag{2}$$

Applying (2) to the developables (1) of a congruence, we get

$$\begin{aligned} D_1(t) &= \bigcup_s D_1(s,t) \ , \\ D_2(s) &= \bigcup_t D_2(s,t) \ . \end{aligned} \tag{3}$$

The two 2-parameter families of planes $D_1(s, t)$ and $D_2(s, t)$ are called the *focal planes* of the congruence. Since $D_1(s, t)$ is tangent to $D_1(t)$, and $D_2(s, t)$ is tangent to $D_2(s)$, it follows from (1) that these two planes must intersect in the line $l(s, t)$, i.e. we can write

$$l(s,t) = D_1(s,t) \cap D_2(s,t) \tag{4}$$

1. In the sequel, we will always exclude stars and plane systems from the congruences under study, unless we explicitly state otherwise.

In the case of a *normal* congruence, the two families of developables always consist entirely of *real* surfaces. This is not always true for a general congruence, where the developables of one family may well be *complex conjugates* of the developables of the other[1].

Each developable of the same family has its edge curve on the same focal sheet. Moreover, this developable is tangent to the other focal sheet along a curve which is *conjugate* to the directions given by the edge curves of the developables of the other family. Denoting the corresponding curve net by the term *edge-conjugate*, we can summarize this configuration the following fundamental:

Fact 1: To each non-degenerated focal sheet of a given congruence there is associated an *edge-conjugate* curve net, corresponding to the two families of its developables.

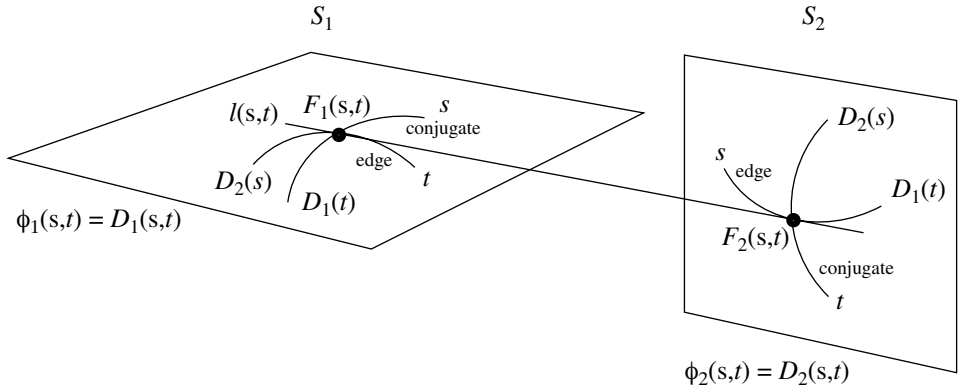


Figure 2

Figure (2) illustrates the focal sheet representation $<S_1|S_2>$ of a given congruence. Starting from the two families of developables, there are two geometrically natural representations of each focal sheet as the locus of a 1-parameter family of curves. We will call these representations the *edge* representation respectively the *conjugate* representation of the focal sheets, and we write:

$$\langle S_1|S_2\rangle^{edge} = \langle \bigcup_s D_2(s) | \bigcup_t D_1(t) \rangle = \bigcup_{s,t} \langle D_2(s)|D_1(t)\rangle \quad \textbf{(5)}$$

respectively

$$\langle S_1|S_2\rangle^{conjugate} = \langle \bigcup_t D_1(t) | \bigcup_s D_2(s) \rangle = \bigcup_{s,t} \langle D_1(t)|D_2(s)\rangle \quad \textbf{(6)}$$

1. However, for the < Target | Observer > congruence that we will define below, the two families of developables will always be real.

3 Applying Congruence Geometry to Scene Analysis

In this chapter we describe the observational configuration that forms the geometric foundation for the application of congruence geometry to vision. Introducing some appropriate notation, we arrive at the classification (15) of the global line-geometric structure of the various target-observer configurations that can appear in the monocular observation of lines, curves and surfaces under translational motion.

3.1 The <Target | Observer> Configuration

Consider a point like observer O, that is moving along an (unknown) curve Γ. While it is occupying the successive positions $O(t) = \Gamma(t)$, the observer is recording a motion picture of the silhouette $S(t)$ of a (possibly) moving target $T(t)$, using a monocular ("pin-hole") camera that is pointing in a *fixed* direction. We can think of the silhouette $S(t)$ as being recorded either on a spherical screen with center $O(t)$ - e.g. the gaussian sphere $\mathbf{S}^2_{O(t)}$ - or on some plane screen $\pi(t)$ that is rigidly attached to $O(t)$. In the discussion below, we will take $\pi(t)$ to be the tangent plane to the gaussian sphere at its north pole, as illustrated in Figure (3).

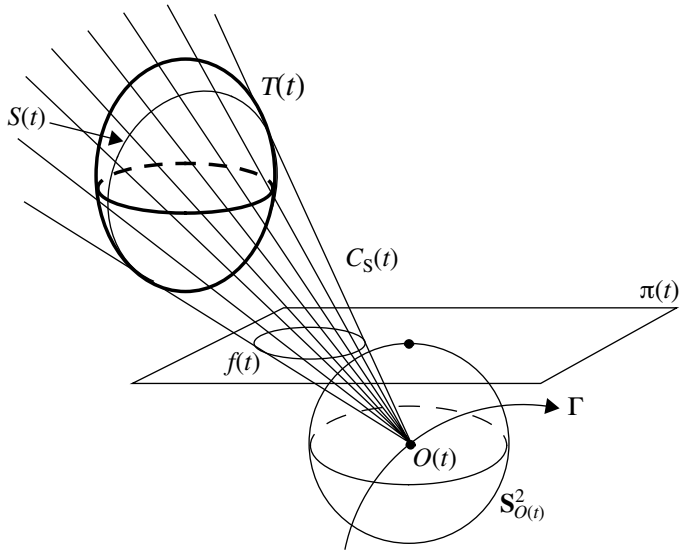


Figure 3

Keeping the direction of the camera fixed during the recording of the film is equivalent to the following geometric condition: During the motion of O, both $\mathbf{S}^2$ and π are *translated* through space without any accompanying rotation, i.e. their directions remains fixed in the surrounding 3d-space. In practice, this can be achieved e.g. by the aid of an inertial navigational system based on the kind of gyroscopic compass employed by modern airplanes.

We will assume that the moving target $T(t)$ consists of points, lines, curves or surfaces. The target curves and target surfaces are assumed to be piecewise smooth, i.e. each target curve must possess a unique tangent line, except possibly at a finite number of points, and each target surface must possess a unique tangent plane, except possibly at a finite number of points and along a finite number of curves.

In the target-observer interaction process described above, let us consider the observation event at time t. In this observation event, each point of the observed target silhouette $S(t)$ corresponds to a line through the point of observation $O(t)$. The totality of these lines will be denoted by $\langle T(t)|O(t)\rangle$.

Unless we explicitely state otherwise, we will always assume that each target $T(t)$ is either a line, a curve, or a surface patch. In this case, the observed silhouette $S(t)$ is a piecewise smooth curve, corresponding to an *observed silhouette cone* $C_S(t)$,

$$C_S(t) = \langle T(t)|O(t)\rangle , \tag{7}$$

with its vertex on the point $O(t)$. Therefore, to each moment of time t, the corresponding observation event associates a 1-parameter family of lines $l_S(s,t)$ that generate the observed silhouette cone $C_S(t)$, that is

$$C_S(t) = \bigcup_s l_S(s,t) . \tag{8}$$

However, what the observer actually *records* is only the *direction*s of these lines, forming the *recorded silhouette cone* $C_R(t)$,

$$C_R(t) = \bigcup_s direction(l_S(s,t)) , \tag{9}$$

which is registered either as a curve on the gaussian sphere $\mathbf{S}^2_{O(t)}$, as in Figure (4), or as a curve in the plane $\pi(t)$, as in Figure (3).

Hence, to each observation sequence there is naturally associated a *2-parameter family of lines*, forming a *1-parameter family of observed silhouette cones*,

Definition 1: The 1-parameter family of observed silhouette cones $C_S(t)$ will be called the *target observer configuration* and will be denoted by $\langle T|O\rangle$.

Introducing the notation $\langle T(t)\rangle$ and $\langle O(t)\rangle$ for the set of lines that are tangent to the target respectively the observer at time t, we thus obtain

$$\langle T|O\rangle = \bigcup_t C_S(t) = \bigcup_t \langle T(t)|O(t)\rangle = \bigcup_t \langle T(t)\rangle \cap \langle O(t)\rangle . \tag{10}$$

Note: In general, the <*T*|*O*> configuration is a 2-parametric line family. However, if the target consists of a *point*, then the observed silhouette cone reduces to a *single* line, and the <*T*|*O*> configuration reduces to a *1-parameter family of lines*.

As described above, a 2-parameter family of lines is known in geometry as a *congruence*, while a 1-parameter family of lines is called a *ruled surface*. Hence, we can summarize in the following:

Fact 2: If the observed target is a *line*, *curve* or *surface*, then the <*T*|*O*> configuration is a *congruence*, and if the observed target is a *point*, the <*T*|*O*> configuration is a *ruled surface*.

It is clear that all the lines of a <*T*|*O*> congruence must pass through the orbit curve Γ of the observer *O*. Hence, a <*T*|*O*> congruence is always *directed* with the curve like focal sheet identical to the observer's orbit curve Γ. A schematic picture of the developables of a <*T*|*O*> congruence is shown in Figure (4).

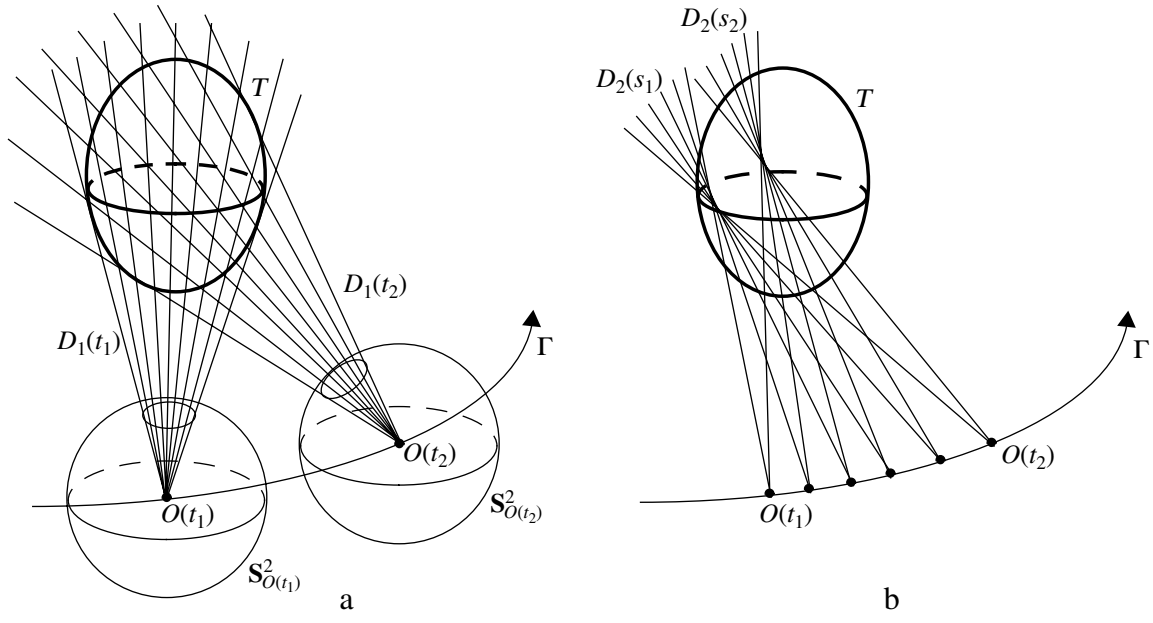


Figure 4

Definition 2: For a given <*T*|*O*>-congruence, the family of conical developables $D_1(t)$ will be called its *target developables*, and the other family $D_2(s)$ will be called its *observer developables*.

3.2 Specializing a General Congruence into a <*T*|*O*> Congruence

We will now consider the question of how a <*T*|*O*> congruence is related to a general <$S_1|S_2$> congruence. The best way to understand this is to consider the general congruence discussed above and observe what happens when one of its focal sheets (say S_2) "shrinks down" into a curve. Referring to Figure (2) and Figure (4), it is clear that we can describe the effects of this transformation as follows:

$$\begin{aligned} &D_1(t) \rightarrow C_S(t) \\ &conjD_2(s) \rightarrow \Gamma \\ &edgeD_1(t) \rightarrow O(t) \end{aligned} \quad . \tag{11}$$

This surface --> curve "shrink down process" results in the directed <*T*|*O*> congruence, and it follows from (11) that the degenerated focal sheet, i.e. the observer's orbit curve Γ can be expressed in the following two ways:

$$\begin{aligned} \Gamma &= \bigcap_{s_0 < s < s_1} D_2(s) \quad , \\ \Gamma &= \bigcup_t edgeD_1(t) = \bigcup_t O(t) \quad . \end{aligned} \tag{12}$$

Summarizing the above discussion, we can conclude the following:

Fact 3: The <*T*|*O*> congruence given by (10) corresponds to the *conjugate-edge* representation (6) of a general congruence, where the second focal sheet has degenerated into the observer's orbit curve.

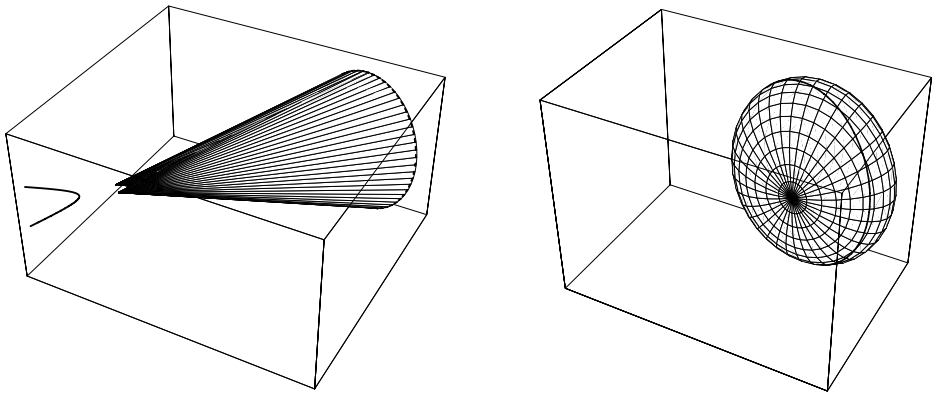
Now, the silhouette curve $S(t)$ is geometrically defined as the part of the intersection between the neighboring silhouette cones $C_S(t)$ and $C_S(t+\delta)$ which does *not* contain the observation point $O(t)$.[1] Hence, denoting by $A\backslash B$ the set of members of A that are not members of B, we obtain

$$S(t) = \{C_S(t) \cap C_S(t+\delta)\} \backslash \langle O(t) \rangle \, . \tag{13}$$

A finite approximation to this configuration for a quadric surface target is shown in Figure (5). The hyperbola shown in the left part of Figure (5a) gives two lines on $O(t)$ in the limit. The silhouette curve $S(t)$ is shown in Figure (5b). Summarizing the above discussion, we can conclude the following

Fact 4: For a <*T*|*O*> congruence, each target developable $D_1(t)$ is tangent to the target $T(t)$ and has its vertex on the observation point $O(t)$. Each observer developable $D_2(s)$ contains the orbit curve Γ of the observer, and has its edge on the target $T(t)$ if the latter is *stationary*, i.e. if $T(t) \equiv T(t_0)$.

1. It is clear that this intersection also contains a finite number of lines that intersect in the point $O(t)$.



a b

Figure 5

The relation between a moving taget and the edges of the corresponding observer developables is of course not so simple and depends on the shape as well as the form of motion of the target.

3.3 The Global Structure of the Various <*T*|*O*> Configurations

Let us introduce the notation $\langle T \rangle$ for the geometric locus of lines the are tangent to the target $T(t)$ for some moment t during the time interval of observation, that is

$$\langle T \rangle = \bigcup_t \langle T(t) \rangle \quad . \tag{14}$$

In (15) we present a summary of the global line-geometric structure of the various forms of <*T*> and <*T*|*O*> configurations that can appear in the process of observing a point-, line-, curve- or surface target, separating between the case of a stationary and a moving target. The deduction of these configurations is straight-forward and follows directly from the previous discussion of the <*T*|*O*> configuration. The notation used in the case of a stationary line target requires an explanation. The target focal sheet of the congruence $\langle FixedLine|O \rangle$ is identical to the $FixedLine$ itself. Since a line can be regarded both as a 1-parameter family of points (ray) and as a 1-parameter family of planes (axis) it follows that the congruence $\langle FixedLine|O \rangle$ is both $DevelopedDirected$ and $DoublyDirected$ - a geometric structure that we have denoted by $DevelopedDirected^2$.

Target	$\langle$**Target**$\rangle$	$\langle$**Target**$\vert$**Observer**$\rangle$
FixedPoint	*StarCongruence*	*Cone*
MovingPoint	$\bigcup_t StarCongruence(t)$	*RuledSurface*
FixedLine	*SingularComplex*	*DevelopedDirected*^{2}*Congruence*
MovingLine	$\bigcup_t SingularComplex(t)$	*DevelopedDirectedCongruence*
FixedCurve	*SecantComplex*	*DoublyDirectedCongruence*
MovingCurve	$\bigcup_t SecantComplex(t)$	*DirectedCongruence*
FixedSurface	*TangentComplex*	*DirectedCongruence*
MovingSurface	$\bigcup_t TangentComplex(t)$	*DirectedCongruence*

(15)

In (15) we have regarded each configuration as a *line manifold*. If, instead, we consider the geometric locus of the target as a *point mainfold*, as well as emphazise the point manifold structure of the focal sheets of the <*T*|*O*> congruence, we get the following modification of (15) :

Target	$\langle$**Target**$\rangle_{point}$	$\langle$**Target**$\vert$**Observer**$\rangle$
FixedPoint	*Point*	
MovingPoint	*Curve*	
FixedLine	*Line*	$\langle Line \vert Curve \rangle Congruence$
MovingLine	*RuledSurface*	$\langle Developable \vert Curve \rangle Congruence$
FixedCurve	*Curve*	$\langle Curve \vert Curve \rangle Congruence$
MovingCurve	*Surface*	$\langle Surface \vert Curve \rangle Congruence$
FixedSurface	*Surface*	$\langle Surface \vert Curve \rangle Congruence$
MovingSurface	*Solid*	$\langle Surface \vert Curve \rangle Congruence$

(16)

4 Geometric Properties of the <*T*|*O*> Congruence

Let us denote the target- and observer focal sheets of a <*T*|*O*> congruence by <*T*| and |*O*> respectively. From the information summarized in (16) we can draw some interesting conclusions. We have already pointed out that if the target is a point, we do not get a <*T*|*O*> congruence at all. Since we will make heavy use of the geometric properties of congruences in the reconstruction process presented below, we exclude point targets from further consideration.Now, even if we were somehow able to completely reconstruct the <*T*|*O*> congruence, we would still only get our hands on the target focal sheet. This is all that is "knowable" about a target from an observation sequence of it, while the manifold that it actually generates in space is beyond recovery (= unknowable) without some form of additional information. It is clear that all the targets that generate the same target focal sheet are indistinguishable (= equivalent) with respect to the given observational sequence, since they generate the same <*T*|*O*> congruence. We express this as the following

Fact 5: Two targets T_1 and T_2 are indistinguishable (i.e. equivalent) with respect to a given observational sequence if and only if they generate the same target focal sheet.

4.1 Line Target

We mentioned in connection with (15) that if the target is a *stationary line*, then the target focal sheet is identical to the target line itself, that is $T = \langle T|$. If, in addition, the observer is subject to *linear motion*, then the resulting <*T*|*O*> congruence has two lines as focal sheets, and therefore both the target- and observer developables reduce to *plane pencils* (Figure (6)). Such a congruence is termed *linear*, since its lines can be described by two linear equations in projective line coordinates.

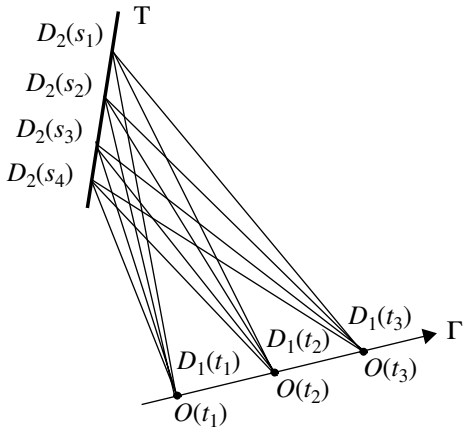


Figure 6

A more interesting situation occurs when the target is a *moving line*. From (16) we see that this line generates a *ruled surface*, while the target focal sheet of the <*T*|*O*> congruence is a *developable*. This situation is depicted in Figure (7), which shows two different views of the same configuration

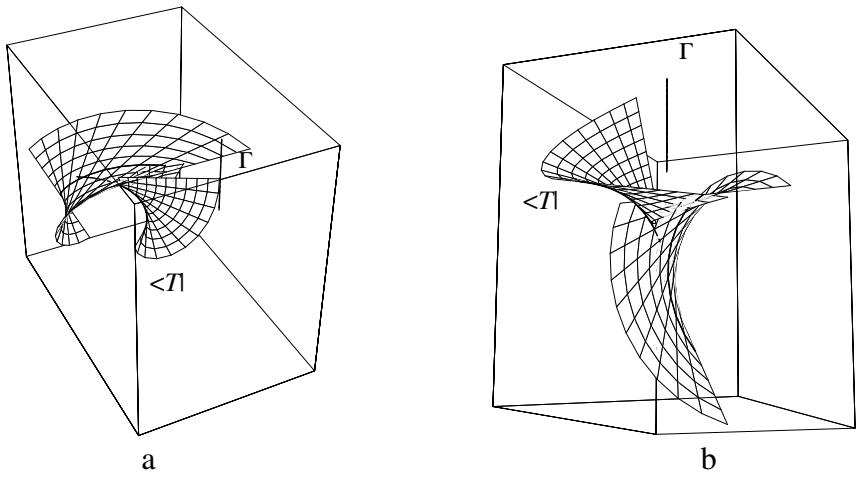


Figure 7

By Fact (5), this developable is all that is knowable about a moving line target from a given observation sequence of it, while the ruled surface that it actually generates in space is unknow-

able. All the moving lines $l(t)$ that generate the same developable are indistinguishable with respect to the given observational sequence. It is clear that two moving lines $l(t)$ and $l'(t)$ generate the same developable if and only if they form the same "plane of observation" together with the point $O(t)$, i.e. if and only if

$$l(t) \vee O(t) = l'(t) \vee O(t) \tag{17}$$

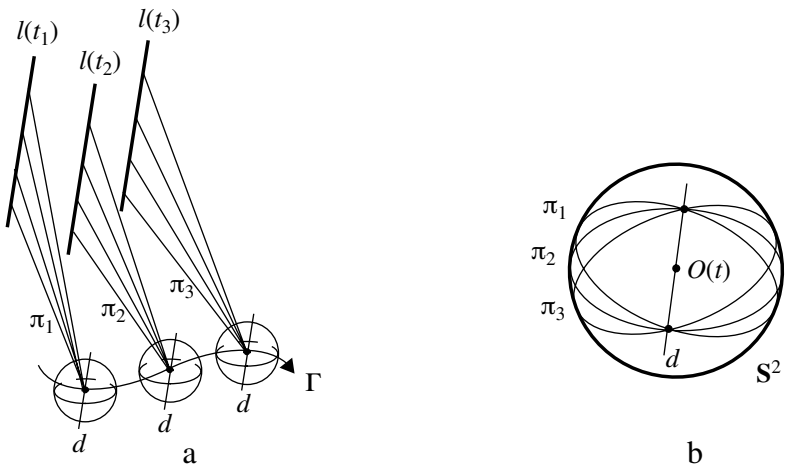


Figure 8

An important special case occurs when the motion of the target line $l(t)$ is a pure *translation*, i.e. when the target line is moving without changing its direction *d* (Figure (8)). In this case all the planes of observation $\pi_i = l(t_i) \vee O(t_i)$ must contain the direction *d*, and therefore they will be recorded on the gaussian sphere as a *pencil of planes with axis d*. Moreover, the developable target focal sheet will take the shape of a *cylinder*, whose generators correspond to the successive positions of the moving line, i.e.

$$\langle T| = \bigcup_t l(t) \tag{18}$$

Hence, in this case, the successive positions of the target line in space are completely recoverable from the <*T*|*O*> congruence. We summarize this configuration in the following

Fact 6: Under translational observer motion, a target line that is subject to some other translational motion will be recorded on the gaussian sphere as a pencil of planes with the axis identical to the direction of the moving target line. The target focal sheet is a cylinder, whose line generators correspond to the successive positions of the moving target line in space.

4.2 Curve or Surface Target

Having discussed the case of point- and line targets above, we now turn to the case of target curves or surfaces. We first consider the case when the observer motion is *linear* - i.e. the observer is translated along a straight line. Since each of the observer developables $D_2(s)$ consists of lines intersecting their neighbors as well as another fixed line (the observer's orbit curve Γ), it follows that all these lines must lie in the same plane. Hence, each observer developable consists of the tangents to a plane curve, and from the first equation of (12) it follows that the corresponding planes form a pencil with the observer's orbit line as axis. We summarize these observations as

Fact 7: Under linear translational observer motion, each observer developable is a tangent surface of a plane curve. The totality of these planes form a pencil with its axis on the observer's orbit line[1].

This fact can be used to discriminate between linear and non-linear observer motion under observation of general targets, provided that we can somehow reconstruct the spherical representation (= gaussian image) of the observer developables. We will see below how this can be achieved in the case of a stationary non-linear target, i.e. a stationary target curve or surface.

5 Reconstructing the <*T*|*O*> Congruence

We will now consider the fundamental problem of reconstructing the <*T*|*O*> congruence, given the spherical representation of its target developables. We recall that the target developable $D_1(t)$ is identical to the observed silhouette cone $C_S(t)$, which is the cone on the silhouette curve $S(t)$ with vertex $O(t)$, given by (7) or (8), and that the spherical representation of $C_S(t)$ is the recorded silhouette cone $C_R(t)$ given by (9). It follows that $C_S(t)$ is identical to the translation of $C_R(t)$ from the origin to the point of observation $O(t)$, i.e.

$$C_S(t) = Translate(C_R(t), O(t)) \tag{19}$$

Let us assume that we have somehow aquired knowledge of the point of observation $O(t)$. Then, from (19) we can reconstruct $C_S(t)$, and thus obtain the <*T*|*O*> congruence as a 1-parameter family of cones. Hence, from (13), we can reconstruct the silhouette curves $S(t)$, and since, by definition of $S(t)$,

$$\begin{aligned} C_S(t) &= \langle T(t)|O(t)\rangle = \langle T(t)\rangle \cap \langle O(t)\rangle = \\ &= \langle S(t)|O(t)\rangle = \langle S(t)\rangle \cap \langle O(t)\rangle \ , \end{aligned} \tag{20}$$

we know that $S(t)$ must be a curve on the target $T(t)$. Hence, if the target is stationary, we have obtained a 1-parameter family of curves on the target - thereby gaining some information

1. This is of course true also for a line target, as illustrated in Figure (6).

about its shape. If the target is moving, we have instead a 1-parameter family of curves on the envelope that the moving target sweeps out over time. This information indicates what region of space to stay away from at a certain time in order to avoid colliding with the moving target. It does not in itself tell us much about the shape of the target, but it could be combined with other observations in order to generate such target shape information.

Now, recall from (8) and (11) that

$$C_S(t) = \bigcup_s l_S(s,t) = D_1(t) . \tag{21}$$

Hence, by reconstructing the observed silhouette cones $C_S(t)$, we have gained information about the target developables $D_1(t)$. What can we say about the observer developables $D_2(s)$?

Denote by $l_\Gamma(t)$ the tangent line to the observer's orbit curve at the point $O(t)$. Knowing the latter point as a function of *t*, the line $l_\Gamma(t)$ can be reconstructed as

$$l_\Gamma(t) = \langle O(t) \rangle \cap \langle O(t+\delta) \rangle . \tag{22}$$

Recall from (3) the definition of the focal planes $D_i(s,t)$. Since $D_2(s,t)$ is tangent to the observer developable $D_2(s)$, this focal plane must in fact contain the line $l_\Gamma(t)$. Hence, it follows from (4) that

$$D_2(s,t) = l_S(s,t)^{plane} \cap l_\Gamma(t)^{plane} , \tag{23}$$

and therefore we can reconstruct the observer developable $D_2(s)$ as

$$D_2(s) = \bigcup_t D_2(s,t) . \tag{24}$$

Moreover, from Fact (1) and Fact (3), we know that the edges of the observer developables form a 1-parameter family of curves that are conjugate to the silhouette curves $S(t)$ on the target focal sheet. Hence, if the target is stationary, we obtain from (21) and (24) a conjugate net of curves on the target[1]. This is the epipolar parametrization introduced by Cipolla in [(5)], where he makes use of it in order to reconstruct the geometry of the target surface.

Let us summarize what has been achieved so far. Starting from a knowledge of the point of observation $O(t)$, we have aquired complete information about the two families of developables of the <*T*|*O*> congruence. We will now address the question of how we can deduce the necessary information about the location of the point $O(t)$. Writing $\boldsymbol{O}(t)$ for the position vector of the point $O(t)$, we have

$$\boldsymbol{O}(t) = \boldsymbol{O}(t_0) + \int_{t_0}^{t} \boldsymbol{O}'(\tau)\, d\tau , \tag{25}$$

1. If the target is moving, we obtain instead a net of curves on the target envelope.

where

$$\boldsymbol{O}'(\tau) = velocity(\tau) = speed(\tau)direction(\tau) \quad \textbf{(26)}$$

is the velocity vector of the observer at time τ.

Note: In the following we will always assume that all targets are finitely extended and totally visible, and that the target and the observer are *translated* relative to each other. Moreover, in the term *envelope* of a family of curve segments we will include the singular part, i.e. the locus of singular points such as cusps or endpoints of the curve segments.

Before proceeding further we will introduce some useful terminology. At a certain moment *t*, the relative motion of target and observer will be called *transversal*, if the recorded silhouette curves at times *t* and *t*+Δ have at least two common (real) tangents for all small enough positive Δ, i.e. if the envelope of the recorded silhouette curves has at least two real tangents at time *t*.

Such a situation, for a stationary target, is depicted in Figure (9). The notation used in this figure corresponds to an infinitesimal time difference, $\Delta = \delta$, but - for the moment - let us consider the two observed silhouette cones $C_S(t)$ and $C_S(t+\delta)$ as being finitely separated in space. Now, imagine that we take a plane ψ through the line that joins the points $O(t)$ and $O(t+\delta)$ and turn the plane ψ around this line until it touches the target $T(t_0)$ at the point P_1. Denoting the corresponding plane by ψ_1, we obtain

$$\psi_1 = O(t) \vee O(t+\delta) \vee P_1 . \quad \textbf{(27)}$$

Since the line $O(t) \vee P_1$ is tangent to $T(t_0)$ and passes through $O(t)$, this line must belong to $C_S(t)$, and for the analogous reason the line $O(t+\delta) \vee P_1$ must belong to $C_S(t+\delta)$. It follows that the plane ψ_1 must be tangent to both of the cones $C_S(t)$ and $C_S(t+\delta)$. Moreover, this construction can be repeated by turning the plane ψ in the opposite direction until it touches the target in another point P_2. For the same reason, the corresponding plane ψ_2 must also be tangent to both of the cones $C_S(t)$ and $C_S(t+\delta)$. Hence we can summarize in the following:

Fact 8: If the target and observer are subject to transversal relative motion, then for small enough positive Δ, the observed silhouette cones $C_S(t)$ and $C_S(t+\Delta)$ have at least two common tangent planes.

We are now in a position to prove the following[1]:

Fact 9: Under transversal observer motion relative to a stationary target, the direction of the observer's motion (the so called *Focus Of Expansion*) at time *t* is identical to the limiting position, as $\Delta \to 0$, of the point of intersection between the common tangent lines to the two recorded silhouette curves at time *t* and time *t*+Δ.

1. The author has recently discovered that Fact 9 is implicit in a paper by Cipolla , Åström and Giblin [(10), Property 1].

Proof: The situation is illustrated in Figure (9). Let the recording in the plane $\pi(t)$ of the silhouette curve $S(t)$ be denoted by $f(t)$, and let ψ_1 and ψ_2 denote the two common tangent planes to the observed silhouette cones $C_S(t)$ and $C_S(t+\delta)$ constructed in the proof of Fact (8). (These planes exist, since the relative motion is assumed to be transversal.) Moreover, let $C_S^{\delta}(t)$ denote the translation of $C_S(t)$ along the observer's orbit curve from $O(t)$ to $O(t+\delta)$, and let $f_\delta(t) = C_S^{\delta}(t) \cap \pi(t+\delta)$.

Now, since $C_S^{\delta}(t)$ is the translation of $C_S(t)$ along the observer's orbit curve, it must be identical to the recorded silhouette cone at time *t*, i.e. $C_S^{\delta}(t) = C_R(t)$. Therefore, the two curves $f_\delta(t)$ and $f(t+\delta)$ are identical to the two recorded silhouette curves at times *t* respectively *t*+δ. Moreover, since the translation takes place along the line $l_\Gamma(t) = \psi_1 \wedge \psi_2$, it follows that the two planes ψ_1 and ψ_2 must remain tangential to the translated cone $C_S^{\delta}(t)$. Hence, ψ_1 and ψ_2 are tangent planes to both $C_S(t)$ and $C_S^{\delta}(t)$, and, by construction, also to $C_S(t+\delta)$.

Consider now the two lines $l_1 = \psi_1 \wedge \pi(t+\delta)$ and $l_2 = \psi_2 \wedge \pi(t+\delta)$. Since they are located in ψ_1 respectively ψ_2, it follows that they must intersect on the line $l_\Gamma(t)$, and also that they must be tangent to the cones $C_S(t+\delta)$ and $C_S^{\delta}(t)$. Furthermore, since l_1 and l_2 are located in the plane $\pi(t+\delta)$, it follows that they must be tangent to both of the recorded silhouette curves $f(t+\delta)$ and $f_\delta(t)$. Therefore, we can conclude that *the directions of the lines l_1 and l_2 are constructible from the recorded image data alone, as common tangents to two successively recorded silhouette curves.* Since the direction of the observer's motion, i.e. the Focus Of Expansion, is given by

$$FOE(t) = l_\Gamma(t) \wedge \pi(t+\delta) = l_1 \wedge l_2, \qquad \textbf{(28)}$$

we have obtained a proof of Fact (9). <<>>

Note: For surface targets with a simple enough geometry, such as e.g. convex surfaces, the transversality condition on the relative motion of target and observer discriminates between colliding and non-colliding motion. Transversal motion corresponds to non-collision (moving "around" the target, while non-transversal motion corresponds to collision (moving "towards" or "away from" the target. This can be seen from the fact that motion along a generator of the observed silhouette cone corresponds to tangency of the recorded silhouette curves. Hence this "grazing" motion, which separates colliding from non-colliding configurations corresponds to the boundary between transversal and non-transversal motion. For more complicated surface shapes the correspondence between transversality- and collision conditions is an interesting problem that merits further study.

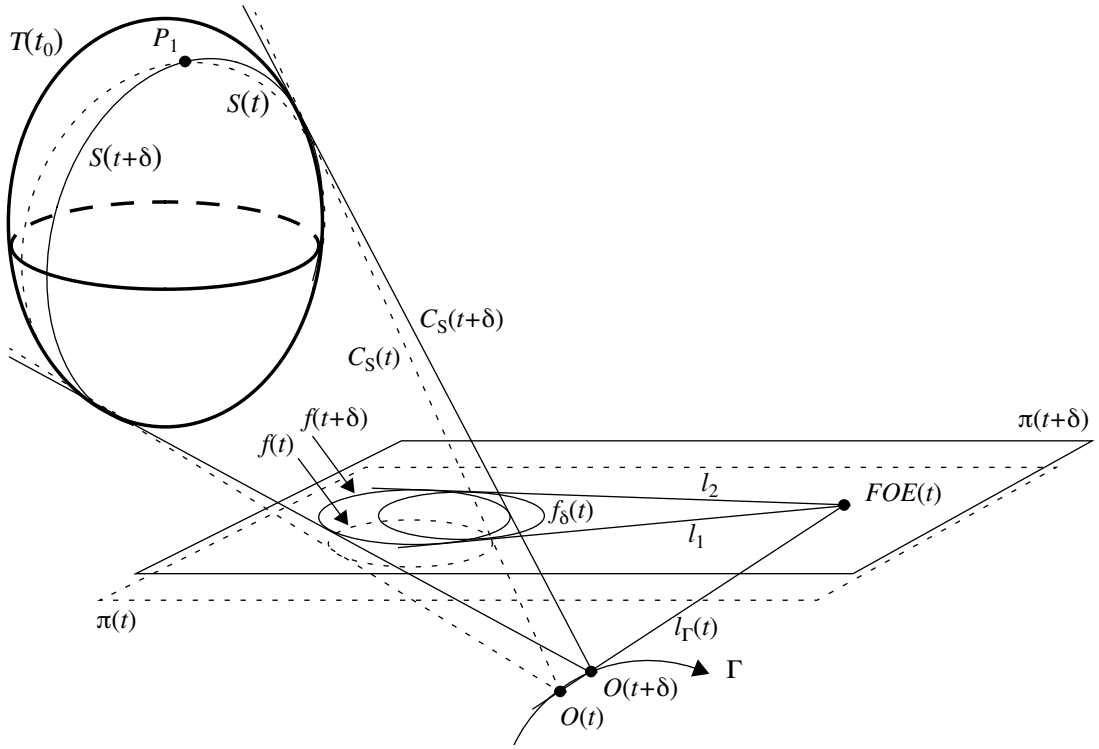


Figure 9

Combining Fact (9) with (25) and (26) we can draw an interesting conclusion. Any hypothesis regarding the speed of the observer's motion and the location of the observation start point $O(t_0)$ constitutes the necessary input to a reconstruction process of the observer's position $O(t)$ as a function of time. Moreover, from the fundamental theorem of calculus, we know that variation of the constant vector $\boldsymbol{O}(t_0)$ gives rise to congruent observer orbit curves, and therefore, by (19) and (24), to congruent reconstructions of the target and observer developables of the <*T*|*O*> congruence. Hence we can state the following

Fact 10: Under transversal motion of an observer relative to a stationary target, the shape of the observer's orbit curve - as well as the shape of a net of silhouette curves and their conjugates - are reconstructible, knowing only the speed of the observer as a function of time. Knowing also the position of the starting point of observation, makes it possible to reconstruct the position of these curves in space.

6 Simulated Examples of <*T*|*O*> Congruence Reconstruction

Following the theoretical discussion presented above, we will now take a look at some examples examples of <*T*|*O*> congruence reconstruction. In this section, the examples will be computer simulated in *Mathematica*®, while in the next section we will work with real image data. For reasons of simplicity, the target will always be chosen as a stationary quadric surface, in which case the intersection of tangent cones can be handled analytically.

6.1 Transversal Observer Orbit Curve

In the first example, shown in Figure (10a), the stationary quadric surface target is observed transversally by an observer moving with constant speed along a circular helix orbit curve Γ. Figure (10b) shows a finite sample of the recorded silhouette curves, where the dots represent the intersections of common real tangents to neighboring curves. Hence, by Fact (9), the dotted circle represents the direction of motion as a (discretely sampled) function of time. The circular shape of the dotted curve in Figure (10b) is due to the fact that the circular helix has its axis in the vertical direction, and therefore the unit tangent to this curve must trace a circle on the Gaussian sphere and therefore also in the (horisontal) recording plane. Moreover, the equal spacing of the dots is a consequence of the constant speed of the observer. Starting from a correct "guess" of the observer's speed and starting position $O(t_0)$, the result of reconstructing 40 points of the observer's orbit curve is shown as the dotted curve of Figure (10c). A factor 4 subsample of these points, as well as the corresponding reconstructed silhouette curves are shown in Figure (10d). Since the reconstruction algorithm assumes that the observer's velocity is constant between each of the sample points, the error is accumulative.

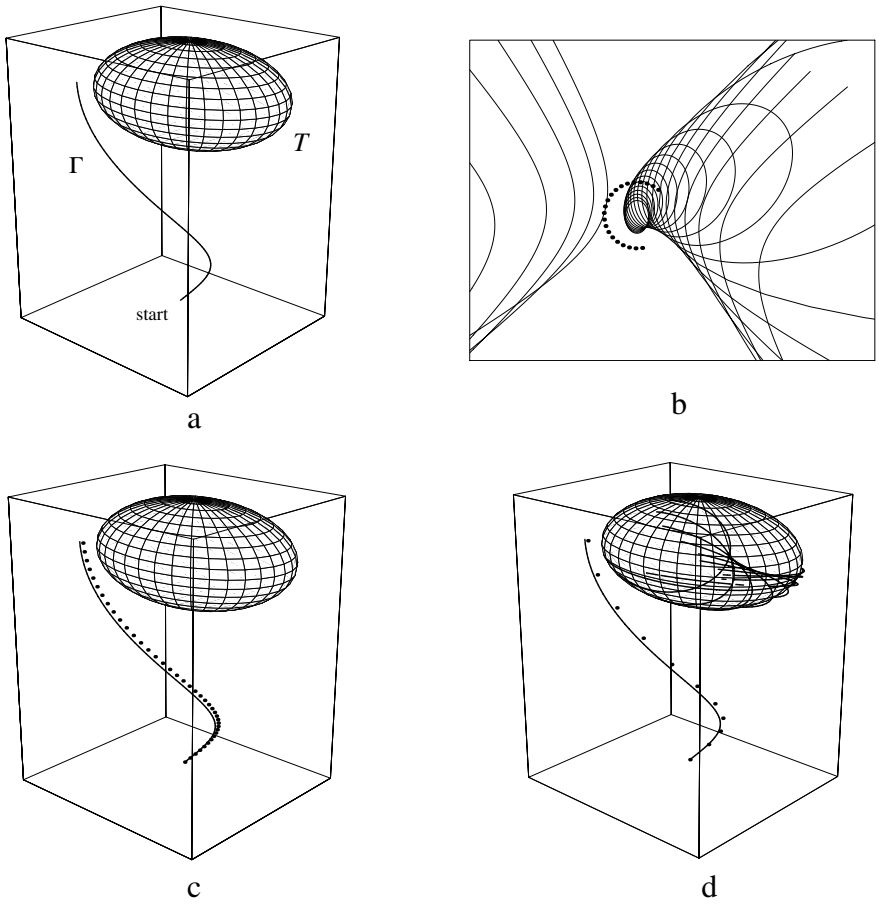


Figure 10

Figure (11) illustrates the sensitivity of the reconstruction process with respect to variation of observer speed. Reconstruction with 60 points is shown with correct speed in Figure (11a), and with a speed error of respectively +2.5% in Figure (11b), -5% in Figure (11c) and -10% in Figure (11d).

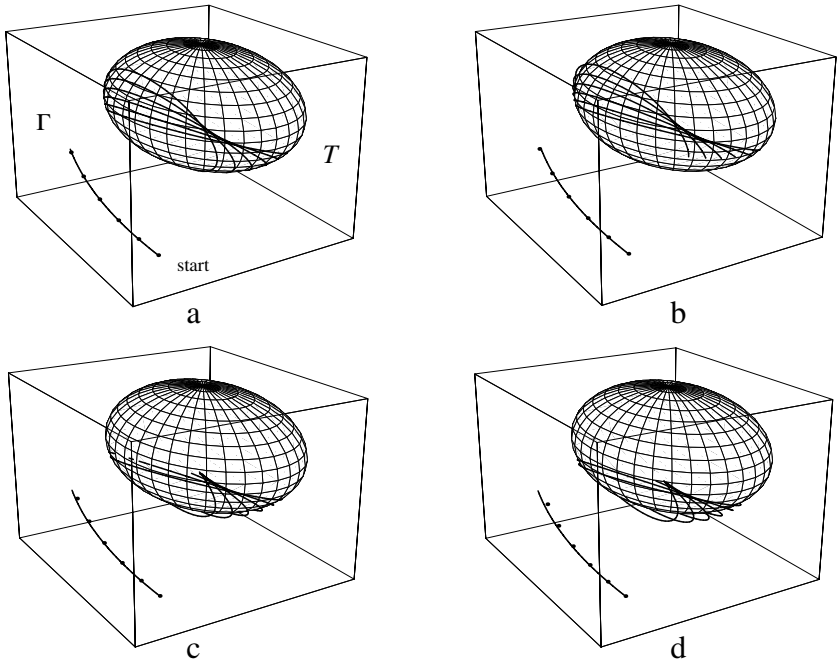
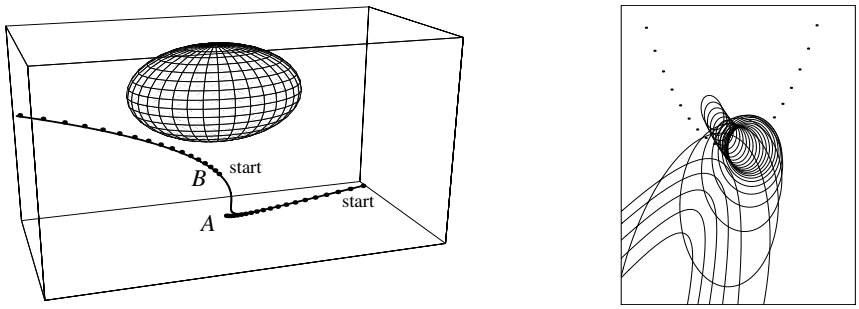


Figure 11

6.2 Non-transversal Observer Orbit Curve

Figure (12) illustrates the complications that are introduced by a non-transversal observer motion. Here the observer is moving along a twisted cubic which is located so that the initial motion is transversal. However, at the point *A* the transversality condition is violated, and the observer starts moving towards the target. Therefore, the common tangent planes ψ_1 and ψ_2 in the proof of Fact (9) are no longer real, and the algorithm for reconstructing the direction of motion breaks down at this point. At a later moment (point *B*) the motion again becomes transversal, and the reconstruction process can be started again, given a correct input of speed and initial observer position.

This illustrates the serious limitations of the transversality condition to the reconstruction algorithm as presented above. In order to be practically useful, we should be able to allow *generic* observer motion, and extend the algorithm to reconstruct also the non-transversal part of the observer's orbit curve. In the case of *quadric* surface targets, a method for achieving this will be presented below. This method can be naturally extended to *polynomial* surface targets, but the details of this extension will not be discussed here.



a b

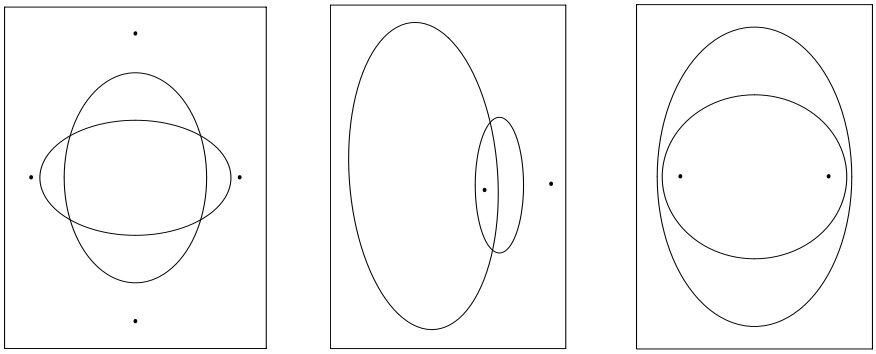
Figure 12

Our method is based on the fact that, for a quadric surface target, the tangent planes ψ_1 and ψ_2 will be complex-conjugates of each other, and so will the lines l_1 and l_2. Hence the line of intersection of ψ_1 and ψ_2 will still be real, and since we have made no special reality assumptions in the proof of Fact (9), the conclusion is still valid, provided only that we modify it to account for this complex-conjugacy. We formulate this as

Fact 11: Under translational observer motion relative to a *quadric* target, the direction of the observer's motion at time *t* is identical to the limiting position, as $\Delta \to 0$, of the point of intersection between one of the two pairs of complex-conjugate common tangent lines to the two recorded silhouette *conics* at time *t* and time *t*+Δ. If there is a *real* pair of such common tangents, then their point of intersection represents the correct direction, and in this case the corresponding observer motion is *transversal*.

Fact (11) illustrates the additional problem. Due to the fact that the motion is non-transversal, the recorded silhouette curves will be conics with no real points of intersection, and since two such conics have (in general) four common tangent lines that are complex-conjugates in pairs, we no longer have a unique reconstruction point for the direction of observer motion. Instead, we have two candidates, one of which represents the correct direction, whereas the other is a "phantom solution" introduced by the complexity of the situation. However, as will be seen in the next section, by considering the global configuration of recorded silhouette curves, we can exclude the false solution for the following reason: The direction of motion must be closest to the part of the contours where the curves are most densely packed (See Figure (14)).

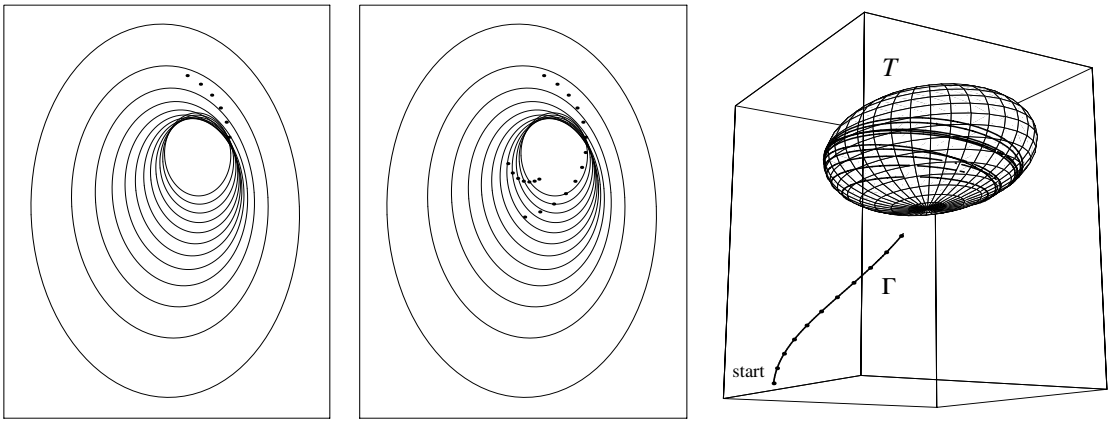
Figure (13) illustrates the three different ways that two generically situated conics can intersect, corresponding respectively to four, two and zero real common tangents. It is clear that only the last two cases can appear as neighboring recorded silhouette curves in the process of observation of a stationary quadric surface target, Figure (13b) corresponding to transversal and Figure (13c) to non-transversal observer motion. Figure (13) also shows the real points of intersection between pairs of common tangent lines, corresponding to pairs of real or complex-conjugate common tangent lines of the two conics. The real tangent pairs are distinguished by the fact that they intersect in points that are located outside of both conics, while the intersection points of complex-conjugate tangent pairs are located on the inside of both conics.



a b c

Figure 13

Assume that the observer motion has been started in such a way that it is inititally transversal. Hence we have (qualitatively) the situation of Figure (13b). When the motion changes to non-transversal, the unique FOE-point will move across the border of both conics, from the outside to the inside [Figure (13c)]. Hence, by keeping track of the FOE-point in the previous reconstruction step, we can discriminate the true solution from the false by a simple continuity argument. Clearly, if the time step is small enough, the correct FOE-point in the next step is the one that is closest to the previous one. Hence, always keeping track of the previous FOE-point, this continuity argument can be applied in the same way during the entire non-transversal part of the motion. If ever the motion becomes transveral again, we do not need to remember the previous FOE-point anymore.



a b c

Figure 14

Figure (14) shows the result of applying this modified type of reconstruction algorithm to the case of a circular helix observer orbit curve, traversed with constant speed, where the configuration is such that the motion changes from transversal to non-transversal. The FOE-points corresponding to the transversal region of motion, with real silhouette envelope, is shown in Figure (14a), and the added FOE-points, corresponding to the intersection of conjugate-com-

plex tangents to the silhouette envelope is shown in Figure (14b). This clearly demonstrates how the continuity condition of the direction of motion helps to discriminate between the two direction candidates. The result of the corresponding silhouette reconstruction process, using 40 points is shown in Figure (14c) with a subsampling factor of 4.

6.3 Reconstructing the Image of the Observer Developables

In this section we will illustrate the reconstruction of the spherical image of the observer developables, and show how this leads to an algorithm for matching points on a stationary target curve over different recorded images. From (15) we recall that if the target is a *stationary* curve, then the *<T|O>* congruence is *doubly directed*. Hence, in this case, both the target- and the observer developables consist of cones. In particular, each observer developable consists of the lines joining a fixed point on the target curve with the successive points of the observer's orbit curve. Thus we arrive at the following

Fact 12: Finding the image of the observer developables $D_2(s)$ in the recording plane $\pi(t)$ is equivalent to constructing a matching between the individual points on a stationary target curve over the different recorded silhouette frames.

Let us now address the question of how the image of the observer developables can be found. From (23) we know that the focal plane $D_2(s,t)$ must contain the tangent line $l_\Gamma(t)$ to the observer's motion curve at time *t*. Hence, by joining the point FOE(*t*), representing the current direction of motion, to the successive points on the recorded silhouette curve *f*(*t*) (Figure (15)), we get a direction field on the latter curve, representing the direction of the corresponding focal plane $D_2(s,t)$. From (24) it follows that integrating this field gives a family of curves representing the image of the family of observer developables, and hence, by Fact (12), the required matching of the points on the target curve over time. We summarize this as:

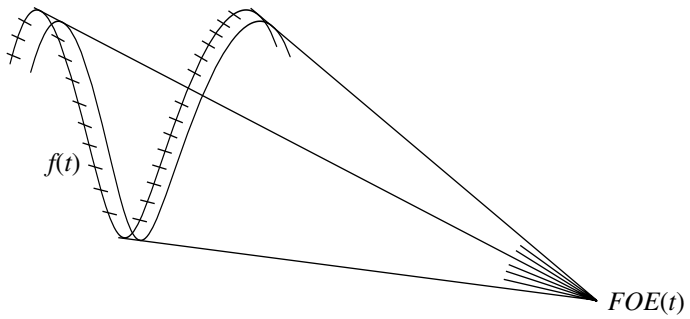


Figure 15

Fact 13: Matching the points on a stationary target curve from a sequence of recorded silhouettes can be achieved by integrating the direction field obtained by associating each point on the recorded silhouette curve with the direction to the corresponding focus of expansion.

Figure (16) shows the results of this process in the case of a circular helix target curve observed under motion along another circular helix. Figure (16a) shows the target-observer configuration, Figure (16b) the recorded silhouette curves, i.e. the image of the target developables, and Figure (16c) shows the result of the integration process described above, matching the points on the target curve over time.

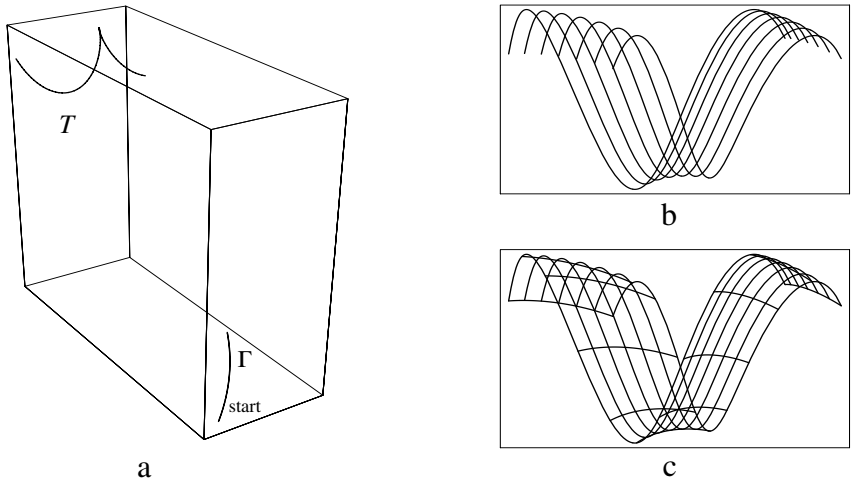


Figure 16

7 Application to Real Image Data

Having presented a few reconstruction examples on synthetic data, we will now apply the FOE-reconstruction technique to a sequence of real images. Since we do not presently possess equipment for controlled general 3d-motion, we have limited the experiments to detecting the motion direction under linear motion.

7.1 Linear Motion with Conjugate Complex Silhouette Envelope

Figure (17) shows a sequence of images of a plastic ellipsoid target[1] taken under linear motion of the observer. The direction of motion was chosen to be in the "colliding" region , thus assuring that the envelope of the recorded silhouette curves would be conjugate complex. For practical reasons we have chosen to translate the target along a straight line instead of doing the same thing with the camera. These two configurations are of course equivalent. Figures (17a) and (17b) show the starting and finishing positions of the target, and Figure (17c) shows the recorded silhouettes overlayed in the same image. In Figure (17d) we see the result of applying an edge detector to each successive image frame and then overlaying the results in a single image. Finally, Figure (17e) shows the overlay of Figure (17a) and Figure (17d).

1. This target was manufactured by T. Henderson at the University of Utah and made available by the courtesy of O.D. Faugeras at INRIA.

Figure 17

To arrive at an estimation of the direction of motion, we now proceed as follows: First, we perform a least square fit of a quadratic polynomial to each extracted edge-curve (i.e. recorded silhouette curve) of Figure (17d). Since we have been using a quadratic target, these fitted conics can be expected to give a good analytic approximation of the recorded silhouette curves. Second, when we have determined these approximating conics, we use them just as in the previous chapter, and determine the points of intersection to double tangents of neighboring conics in the family. This corresponds to using the time frames t and $t + \Delta t$, where Δt is the time difference between successive images. The result is shown in Figure (18a). As can be seen, the presence of noise in our data introduces a problem. Computing instead the points of intersection to double tangents of the conics corresponding to times t and $t + 2\Delta t$, we get the result of Figure (18b). For the sake of simplicity, we will refer to such curves as being related by step-2. Figures (18c) and (18d) show the corresponding computation for conics related by step-3 respectively step-4. As might be expected, we infer from Figures 18a, b, c, d that increasing the step greatly reduces the noise. If we have reason to trust one of the silhouette curves significantly more than the others, we can use it repeatedly and compute the points of intersection to double tangents of it and the other curves in the family[1]. The result of using the initial silhouette curve (smallest of the conics) in this way is shown in Figure (18e). It is clear that this further increases the precision of the computation of the FOE-point.

Finally, to get an idea of the "correct" result, an idealized configuration is shown in Figure (18f). The conics here are formed from a dualized linear combination of the first and the last conic in the family, which means that all pairs of conics from this new family have the same common tangents.

Note: We can exclude the false direction candidate on account of the global structure of the family of curves in Figure (18). Since the recorded silhouettes are much closer to each other at the upper left corner of the picture, the true direction of motion must be the point closest to this area. If the curves would have all touched each other in one point, then the camera would in fact have been moving towards that point - tangentially to the target object.

8 Conclusions and Future Work

In this paper we have discussed the scene reconstruction problem under translational observer motion. We have introduced the <Target|Observer> congruence - with its corresponding target developables and observer developables - as a framework for a line-geometric approach to such problems [Definitions (1) and (2)]; we have demonstrated how these concepts give a way to describe the global structure of the various geometric configurations that arise from the observation of points, lines, curves and surfaces [(15)]; we have analyzed under what conditions two targets are indistinguishable in a given sequence of monocular observations [Fact (5)]; we have characterized linear observer motion in terms of the global geometric structure of the corre-

1. We might, for instance, have kept the camera at rest for a long time, and therefore have been able to determine the corresponding silhouette curve with a significantly better precision than the others.

sponding observer developables [Fact (7)]. Under known observer motion we have shown how to reconstruct both the target- [(19), (21)] and the observer-developables [(22), (23), (24)].

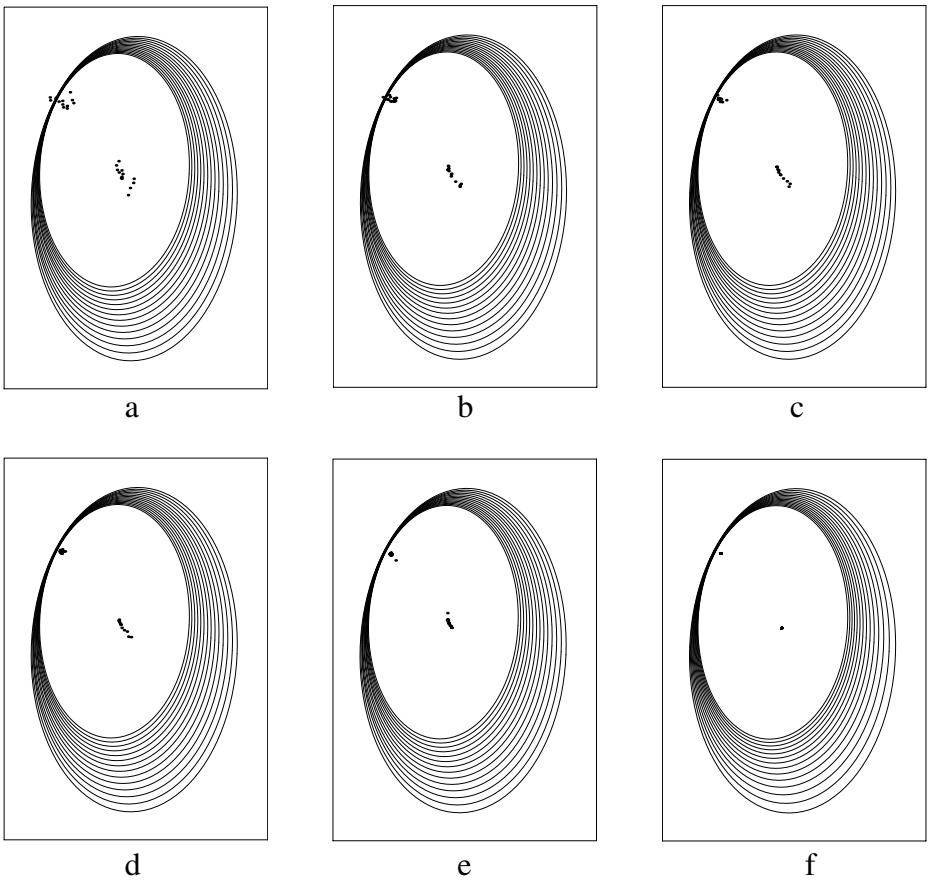


Figure 18

Moreover, we have introduced a direct way to compute the direction of motion (FOE) as the intersection of two double tangents to two consecutively recorded silhouette curves. As pointed out above, when these common tangents are real [Fact (9)] this method has been independently described by Cipolla, Åström and Giblin in [(10) Property 1]. Here we have extended this FOE reconstruction technique to include conjugate complex double tangents [Fact (11)], thereby widening its scope to include observer motion “towards” and “away from” the target. Knowing only the observer’s speed as a function of time, we have shown how to recover the observer´s motion as well as a set of silhouettes on the target and illustrated this technique on synthetic images [Figure (10), Figure (14)] as well as on real image data [Figure (17)]. For the sake of simplicity we have used quadric surface targets, where the computation of silhouette curves can be handled analytically. We stress that any hypothesis regarding observer speed provides the necessary input to a reconstruction process of the scene, the results of which can be compared with any independently available scene information, in order to verify or falsify the initial speed hypothesis. The sensitivity of the reconstruction process with respect to observer speed has been illustrated in Figure (11). Finally we have presented an algo-

rithm for the matching of points on a curve-like target from a sequence of images taken under translational observer motion - without any speed information at all [Fact (13)].

Although we have presented some experimental results, the emphasis of this paper is on the theoretical framework and its ability to provide a unifying description of structure from motion problems that emphasizes the global geometry of the configuration. The practical applicability of this framework in recovering scene structure from real images is therefore still an open question. In the future we plan to investigate this field, aiming to demonstrate how the congruence-geometric framework can enable the construction of globally based methods for determining target shape, verifying and stabilizing observer motion, discriminating between stationary targets and targets moving relative to the background, grouping targets with the same motion relative to the moving observer and discriminating beteen curve edges and occluding surface contours.

9 Acknowledgements

The author is grateful to Meng Xiang Li for his help with the practical experiments, and to Jan-Olof Eklundh for many interesting discussions related to these subjects, as well as for creating and maintaining a highly stimulating research environment at CVAP. This work has been performed within the project CVAP93, funded by The Swedish Research Counsil for Engineering Sciences, whose support is gratefully acknowledged.

10 References

[1] J. J. Koenderink, *What does the occluding contour tell us about surface shape?*, Perception 13, 1984, pp. 321-330.

[2] J.J. Koenderink, *Solid Shape*, MIT Press, Cambridge, Massachussetts, 1990.

[3] P.J. Giblin and R. Weiss, *Reconstruction of Surfaces from Profiles*, Proceedings of the 1st International Conference of Computer Vision, London 1987, pp. 136-144.

[4] R. Vaillant, *Using Occluding Contours for 3D Object Modeling*, O. Faugeras, ed., Computer Vision-ECCV '90 (First European Conference on Computer Vision, Antibes, France, April 23-27, 1990), Proceedings, Springer, Berlin, 1990, pp 454-464.

[5] R. Cipolla, *Active Visual Inference of Surface Shape*, PhD thesis, Univ. of Oxford, 1991.

[6] R. Cipolla and A. Blake, *Surface Shape from the Deformation of Apparent Contours*, International Journal of Computer Vision, Vol 9:2, 1992, pp 83-112

[7] R. Valliant and O.D. Faugeras, *Using extremal boundaries for 3D object modeling*, IEEE Trans. Pattern Analysis and Machine Intelligence, 14(2), 1992, pp. 157-173.

[8] O.D. Faugeras and T. Papadopoulo, *A Theory of the Motion Fields of Curves*, International Journal of Computer Vision, Vol 10:2, 1993, pp. 125-156.

[9] P.J. Giblin and R. Weiss, *Epipolar Fields on Surfaces*, J.O. Eklundh, ed., Computer Vision-ECCV '94 (Third European Conference on Computer Vision, Stockholm, Sweden, May 2-6, 1994), Volume A, Springer, Berlin, 1994, pp 14-23.

[10] R. Cipolla, K.E. Åström and P.J Giblin, *Motion from the frontier of curved surfaces*, Fifth International Conference on Computer Vision, MIT, Cambridge, Massachussetts, June 20-23, 1995, pp. 269-275.

[11] K.N. Kutulakos and C.R. Dyer, *Global surface reconstruction by purposive control of observer motion*, Artificial Intelligence, vol. 78, no. 1-2, 1995, 147-177. surfaces, contours

[12] ChangSheng Zhao and Roger Mohr, *Global 3-D surface reconstruction from occluding contours*, Computer Vision and Image Understanding, vol. 64, no.1, 1996, 62-96.

[13] A. Naeve, *Focal Shape Geometry of Surfaces in Euclidean Space*, Dissertation, TRITA-NA-P9319, Computational Vision and Active Perception Laboratory, KTH, 1993.

[14] L. Eisenhart, *A Treatise on the Differential Geometry of Curves and Surfaces*, Ginn & Co, The Athaenum Press, Boston, Massachussetts, 1909.