International Journal of Computer Vision

The geometry and matching of lines and curves over multiple views

Cordelia Schmid[1] and Andrew Zisserman[2]
[1] INRIA Rhône-Alpes, 655 av. de l'Europe, 38330 Montbonnot, France
Cordelia.Schmid@inrialpes.fr
[2] Dept. of Engineering Science, 19 Parks Rd, Oxford OX13PJ, UK
az@robots.ox.ac.uk

Abstract

This paper describes the geometry of imaged curves in two and three views. Multi-view relationships are developed for lines, conics and non-algebraic curves. The new relationships focus on determining the plane of the curve in a projective reconstruction, and in particular using the homography induced by this plane for transfer from one image to another. It is shown that given the fundamental matrix between two views, and images of the curve in each view, then the plane of a conic may be determined up to a two fold ambiguity, but local curvature of a curve uniquely determines the plane. It is then shown that given the trifocal tensor between three views, this plane defines a homography map which may be used to transfer a conic or the curvature from two views to a third. Simple expressions are developed for the plane and homography in each case.

A set of algorithms are then described for automatically matching individual line segments and curves between images. The algorithms use both photometric information and the multiple view geometric relationships. For image pairs the homography facilitates the computation of a neighbourhood cross-correlation based matching score for putative line/curve correspondences. For image triplets cross-correlation matching scores are used in conjunction with line/curve transfer based on the trifocal geometry to disambiguate matches. Algorithms are developed for both short and wide baselines. The algorithms are robust to deficiencies in the segment extraction and partial occlusion.

Experimental results are given for image pairs and triplets, for varying motions between views, and for different scene types. The methods are applicable to line/curve matching in stereo and trinocular rigs, and as a starting point for line/curve matching through monocular image sequences.

1 Introduction

This paper has two strands: the first is a set of novel results for the geometry of curves imaged in two and three views; the second is an application of these results to the automatic matching of lines and curves over multiple views.

In the geometry strand simple relationships are developed for conics and general non-algebraic curves imaged in two or three views. The key unifying idea is that the plane of the geometric entity can be computed from two views given the fundamental matrix [17]. In particular it is shown that given the image of a conic in 3-space in two views its support plane is determined up to a two fold ambiguity; and that given the imaged curvature and tangent of any curve at corresponding points in two views its plane is determined uniquely. The plane can then be used to transfer that entity into a third view via the trifocal tensor [25, 53, 54]. The transfer of points and lines via the trifocal tensor is well known and explored. This paper extends transfer to conics and curvature. In particular it is shown that the image of a conic in a third view is determined up to a two fold ambiguity from its image in two other views, and that the imaged curvature of any curve is determined uniquely in a third view from imaged curvatures and tangents at corresponding points in two other views. These relationships apply to multiple views of a rigid scene and require only the fundamental matrix and trifocal tensor, but not the camera calibration. The relationships are derived in section 2 and evaluated in section 3.

In the matching strand the objective is to match lines and curves over two or three views. If matching is attempted using only geometric constraints then there is considerable ambiguity. In particular for matching lines over two views there is no equivalent of the epipolar constraint for points, where a point in one view restricts the search for a corresponding point to a line in the other view. The key contribution here is to augment the geometric matching constraints with a photometric constraint, based on the invariance of the surface brightness in the neighbourhood of the curve. This photometric constraint is shown to be a powerful disambiguation measure for line and curve matches over two and three views. Its computation requires the geometric relationships developed in section 2.

An efficient algorithm based on these ideas has been tested over many scene types and over significant changes in viewpoint. It achieves a performance generally of over 90% correct matches for two views, and usually 100% correct over three. Of particular interest is that these scores can be achieved by matching individual curves, without requiring curves to be grouped prior to matching. The matching algorithm is applicable to the 3D reconstruction of rigid scenes from stereo and trinocular camera rigs, and as a starting point for line/curve matching through monocular image sequences. The photometric constraint, matching algorithms and many examples are given in section 4.

This paper can justly be seen as extending two papers by Faugeras and Roberts. The first, [21], dealt with the transfer of conics and curvature from two views to a third via epipolar transfer; in this paper the transfer is via the trifocal tensor and so does not suffer from the failings of epipolar transfer (as described in section 2.7). The second, [45], dealt with the trinocular matching of curves using only geometric constraints; in this paper these constraints are now augmented by the photometric constraint.

1.1 Related literature on line and curve matching

Line and curve matching is a recurring topic in the stereo and motion literature, with the main application being the recovery of 3D geometry. Other applications include qualitatively distinguishing surface curves from apparent contours [7, 58], estimating camera motion [12, 20, 32], and object recognition [39]. We will be concerned with matching surface curves, rather than apparent contours, and will consider views which are not closely spaced, so that curve matching via snake tracking is not viable.

One of the reasons that matching lines and curves between images is difficult is because the entities can not be perfectly retrieved from images: there are segmentation deficiencies, and in particular, the topological connections between segments are often lost [18, 46]. Consequently, segments may be matched individually, or an attempt may first be made to recover the topological connections and then groups of segments are matched.

Line matching. Individual line segments are generally matched on their geometric attributes — orientation, length, extent of overlap [1, 37, 63]. Some such as [15, 16, 31] use a nearest line strategy which is better suited to image tracking where the images and extracted segments are similar. Additional views may be used to verify potential matches [3, 56].

There is an advantage in matching *groups* of line segments in that more geometric information is available for disambiguation, the disadvantage is an increase in algorithmic complexity. A number of methods have been developed around the idea of graph-matching [2, 23, 30, 61]. The graph captures relationships such as left of, right of, cycles, collinear with etc, as well as topological connectedness. Although such methods can cope with more significant camera motion, their complexity can be prohibitive.

An alternative, developed largely in the object recognition literature, is to compute invariants of the groupings, such as affine or projective planar algebraic invariants [47].

Curve matching. Basic criteria for curve matching in stereo pairs were established by Pollard *et al* in the PMF Stereo Algorithm [40]. Their primary criteria were the epipolar and ordering constraints, but these were supplemented by figural continuity to overcome problems where curves coincide with epipolar lines (since if the curve coincides with an epipolar line, point correspondences cannot be determined). Zhang and Gerbrands [62] again used epipolar geometry as their primary constraint, but followed this with criteria on the variation in disparity along the putatively matched curves. Brint and Brady [9] matched curves in trinocular views. Their primary matching constraint was a similarity measure based on the deformation between a curve and its putative match in another view. This measure was used to eliminate many potential mismatches. Trinocular consistency constraints were only used as a final verification.

As in the case of line matching, there are advantages in first carrying out a monocular grouping of curve segments, and then matching the groups. This is the approach adopted by Chung and Nevatia [11], and continued in work at USC by Havaldar and Medioni [29]. Invariants may also be used for curve matching. Van Gool *et al* [60] for example use semi-differential invariants. Carlsson *et al* [10] compare a variety of such methods for planar curves. A very thorough review of curve matching techniques is given in [22].

Photometric constraints. The above methods primarily use only geometry to disambiguate curve matches. This ignores the many photometric constraints that can often be employed for matching over images of the same scene. Bignone *et al* [6] augment purely geometric line matching with photometric attributes: the median luminance and chrominance of the regions adjacent to the line are used to rank geometrically valid matches. However, these photometric attributes are less discriminating than the neighbourhood correlation measure developed in this paper (section 4.1.2), and consequently more missmatches are obtained in their case. Nayar and Bolle [38] developed a photometric matching constraint based on the ratio of intensities across region boundaries. This ratio has good invariance to lighting conditions, but again is less discriminating than the measure developed here.

2 Geometry

The idea that is used in the following sections is that a plane in the world defines a planar homography (plane projective transformation) between two perspective views. This plane may be determined, at least partially, by the correspondence of geometric entities such as lines or conics. By determining the plane, the homography is also determined up to the same ambiguity. We begin with algebraic entities, lines and conics, and then continue with non-algebraic curves. Simple formulae are derived for the homography and plane given the fundamental matrix for the view pair.

The idea is then developed further: given the homography and plane determined in this manner from two views, the homography to a third view is determined using the trifocal tensor for the view triplet. The geometric entity can then be transferred to the third view by this homography. These geometric results are used in the wide baseline matching algorithm of section 4.

2.1 Background material

Notation. Image points are represented by homogeneous 3-vectors $\mathbf{x} = (x, y, 1)^\top$, and 3-space points by homogeneous 4-vectors $\mathbf{X} = (\mathrm{X}, \mathrm{Y}, \mathrm{Z}, 1)^\top$. A bold-face symbol such as $\mathbf{x}$ always represents a column vector, and its transpose $\mathbf{x}^\top$ is a row vector.

A scene point $\mathbf{X}$ is mapped to an image point $\mathbf{x}$ by perspective projection. This map is represented by a 3×4 camera matrix $\mathtt{P}$, as $\mathbf{x} = \mathtt{P}\mathbf{X}$, where $=$ indicates equality up to scale. In general corresponding entities in various views are indicated by primes, so that the image of $\mathbf{X}$ is $\mathbf{x}, \mathbf{x}'$ and $\mathbf{x}''$, in the first, second and third views respectively. The image of $\mathbf{X}$ in the second and third views is given by $\mathbf{x}' = \mathtt{P}'\mathbf{X}$ and $\mathbf{x}'' = \mathtt{P}''\mathbf{X}$ respectively. Image lines are represented by homogeneous 3-vectors $\mathbf{l}$, and image conics by 3×3 symmetric matrices $\mathtt{C}$.

The epipolar geometry for the first two views is represented by the fundamental matrix $\mathtt{F}$: if $\mathbf{x}$ and $\mathbf{x}'$ are corresponding image points then $\mathbf{x}'^\top \mathtt{F}\mathbf{x} = 0$. The epipole in the second view is $\mathbf{e}'$, i.e. $\mathbf{F}^\top \mathbf{e}' = \mathbf{0}$. The notation $[\mathbf{x}]_\times$ means the 3×3 skew matrix with null space $\mathbf{x}$, representing the vector cross product: $[\mathbf{x}]_\times \mathbf{y} = \mathbf{x} \times \mathbf{y}$.

Multiple view geometry of planes. It is now well known [17, 27, 28] that given two views and the fundamental matrix $\mathtt{F}$ between them, a projective reconstruction can be determined from corresponding image points. In particular the camera matrices may be retrieved from $\mathtt{F}$ as $\mathtt{P} = [\mathtt{I}|\mathbf{0}]$ and $\mathtt{P}' = [\mathtt{A}|\mathbf{e}']$, where $\mathtt{F} = [\mathbf{e}']_\times \mathtt{A}$.

Suppose there is a plane $\boldsymbol{\pi}$ in the scene, then it defines a map between points in two views by the following procedure: the ray corresponding to a point $\mathbf{x}$ is extended to meet the plane $\boldsymbol{\pi}$ in a point $\mathbf{X}_\pi$; this point is projected to a point $\mathbf{x}'$ in the other image. The map from $\mathbf{x}$ to $\mathbf{x}'$ is the homography induced by the plane $\boldsymbol{\pi}$, and is represented as $\mathbf{x}' = \mathtt{H}\mathbf{x}$, where $\mathtt{H}$ is a homogeneous 3×3 matrix of rank 3.

It is shown in [34] that the homographies induced by a plane $\boldsymbol{\pi}$ may be parametrized as

$$\mathtt{H}(\mathbf{v}) = \mathtt{A} + \mathbf{e}'\mathbf{v}^\top \tag{1}$$

where the plane is represented by the 4-vector $\boldsymbol{\pi} = (-\mathbf{v}^\top, 1)^\top$ in the projective reconstruction with cameras $\mathtt{P} = [\mathtt{I}|\mathbf{0}]$, $\mathtt{P}' = [\mathtt{A}|\mathbf{e}']$. The inhomogeneous 3-vector $\mathbf{v}$ parametrizes the 3-parameter family of planes in 3-space.

Equation (1) shows that given the plane $\boldsymbol{\pi}$ in a projective reconstruction the homography it induces is determined uniquely. Conversely, given the homography $\mathtt{H}$ the plane which induces it is also determined uniquely as $\boldsymbol{\pi} = (-\mathbf{v}^\top, 1)^\top$ with

$$\mathbf{v} = (\mathtt{H}/\lambda - \mathtt{A})^\top \mathbf{e}'/||\mathbf{e}'||^2 \tag{2}$$

where λ is obtained from

$$\lambda[\mathbf{e}']_\times \mathtt{A} = [\mathbf{e}']_\times \mathtt{H}$$

This computation is given in detail in appendix A.1.

Transformation rules. Points map as $\mathbf{x}' = \mathtt{H}\mathbf{x}$ under a homography. Transformation rules [52] may also be given for lines:

$$\mathbf{l}' = \mathtt{H}^{-\top}\mathbf{l} \tag{3}$$

and conics:

$$\mathtt{C}' = \mathtt{H}^{-\top}\mathtt{C}\mathtt{H}^{-1} \tag{4}$$

2.2 Corresponding lines

The correspondence of two image lines determines a line in 3-space, and a line in 3-space lies on a one parameter family (a pencil) of planes, see figure 1. This pencil of planes induces a pencil of homographies between the two images which map the corresponding lines to each other. Algebraically, it is shown in appendix A.2 that:

Proposition 1. *The homography for the pencil of planes defined by a line correspondence* $\mathbf{l} \leftrightarrow \mathbf{l}'$ *is given by*

$$\mathtt{H}(\mu) = [\mathbf{l}']_\times \mathtt{F} + \mu \mathbf{e}'\mathbf{l}^\top \tag{5}$$

provided $\mathbf{l}'.\mathbf{e}' \neq 0$*, where* μ *parametrizes the pencil. In a projective reconstruction with cameras* $\mathtt{P} = [\mathtt{I} \mid \mathbf{0}]$, $\mathtt{P}' = [\mathtt{A} \mid \mathbf{e}']$*, and* $\mathtt{F}$ *computed as* $\mathtt{F} = [\mathbf{e}']_\times \mathtt{A}$ *(which determines the relative scales), the plane has coordinates*

$$\boldsymbol{\pi}(\mu) = \left(\mathbf{e}'^\top\left(\mathtt{H}(\mu) + (\mathbf{e}'.\mathbf{l}')\mathtt{A}\right), (\mathbf{e}'.\mathbf{l}')||\mathbf{e}'||^2\right)^\top$$

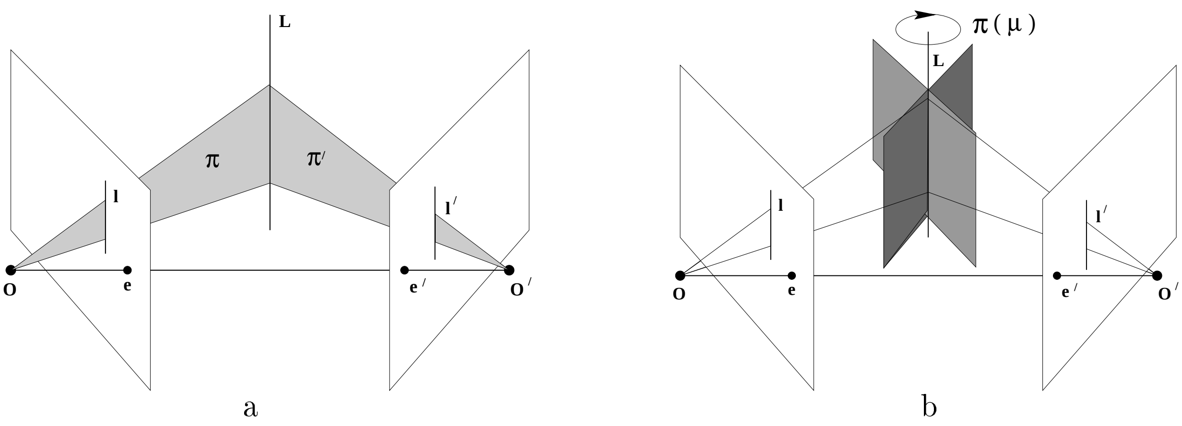


Figure 1: (a) Image lines $\mathbf{l}$ and $\mathbf{l}'$ determine planes $\boldsymbol{\pi}$ and $\boldsymbol{\pi}'$ respectively. The intersection of these planes defines the line $\mathbf{L}$ in 3-space. (b) The line $\mathbf{L}$ in 3-space is contained in a one parameter family of planes $\boldsymbol{\pi}(\mu)$. This family of planes induces a one parameter family of homographies between the images.

The condition $\mathbf{l}'.\mathbf{e}' \neq 0$ arises because if the line in 3-space intersects the baseline between the cameras then its image intersects the epipole, i.e. $\mathbf{l}.\mathbf{e} = \mathbf{l}'.\mathbf{e}' = 0$. In this case the line in 3-space cannot be determined uniquely from its images since it lies in an epipolar plane, and only this plane can be determined from the imaged lines. It is worth remarking that there are degenerate

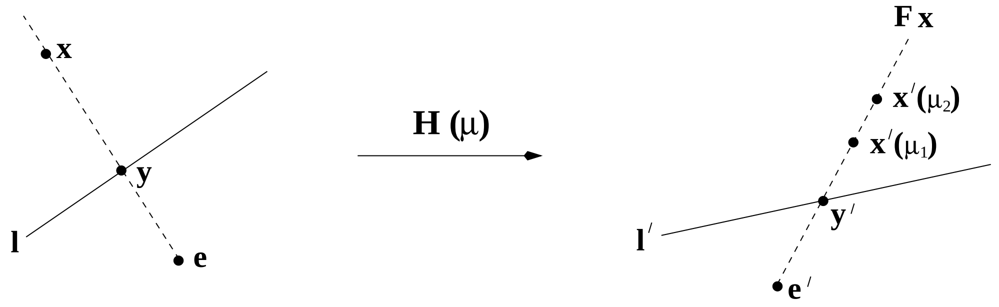


Figure 2: As μ varies, a point $\mathbf{x}$, which is not on the line $\mathbf{l}$, is mapped by $\mathtt{H}(\mu)$ to a point $\mathbf{x}'(\mu)$ which moves along the epipolar line ($\mathtt{F}\mathbf{x}$) corresponding to $\mathbf{x}$. However, the point $\mathbf{y}$, which lies on $\mathbf{l}$, is mapped to the same point $\mathbf{y}'$ for all values of μ.

homographies, i.e. the matrix $\mathtt{H}$ is singular, in this pencil at $\mu = \infty$ and $\mu = 0$. These correspond to planes through the first and second camera centres respectively.

Since $\mathtt{H}(\mu)$ is compatible with the epipolar geometry [34], all members of the pencil map the epipoles to each other, i.e. $\mathbf{e}' = \mathtt{H}(\mu)\mathbf{e}$, and corresponding epipolar lines are mapped to each other, i.e. if $\mathbf{x}$ and $\mathbf{x}'$ are corresponding points, not necessarily images of a point on the plane $\boldsymbol{\pi}(\mu)$, and $\mathbf{l}'^e = \mathtt{F}\mathbf{x}, \mathbf{l}^e = \mathtt{F}^\top\mathbf{x}'$ are their corresponding epipolar lines, then $\mathbf{l}^e = \mathtt{H}^\top(\mu)\mathbf{l}'^e$.

The action of $\mathtt{H}(\mu)$ on points as μ varies is illustrated in figure 2. A point $\mathbf{y}$ which is on $\mathbf{l}$ satisfies $\mathbf{y}.\mathbf{l} = 0$ and is mapped as

$$\begin{aligned} \mathbf{y}' &= \mathtt{H}(\mu)\mathbf{y} = \left([\mathbf{l}']_\times\mathtt{F} + \mu\mathbf{e}'\mathbf{l}^\top\right)\mathbf{y} \\ &= [\mathbf{l}']_\times\mathtt{F}\mathbf{y} = \mathbf{l}' \times (\mathtt{F}\mathbf{y}) \end{aligned}$$

The point $\mathbf{y}'$ is on $\mathbf{l}'$ at the intersection with the epipolar line $\mathtt{F}\mathbf{y}$ of $\mathbf{y}$. Thus for points on $\mathbf{l}$, $\mathtt{H}(\mu)$ is a line to line map (a 1D homography) between $\mathbf{l}$ and $\mathbf{l}'$ which does not depend on μ, since all dependence is removed when $\mathbf{l}.\mathbf{y} = 0$. A point $\mathbf{x}$ which is *not* on $\mathbf{l}$ is mapped to a point $\mathbf{x}'$ on the epipolar line $\mathtt{F}\mathbf{x}$ of $\mathbf{x}$, but in general $\mathbf{x}'$ is not on $\mathbf{l}'$ and does depend on μ. If $\mu = 0$ then $\mathbf{x}'$ lies on $\mathbf{l}'$ for any $\mathbf{x}$, since in this case $\mathbf{l}'.\mathbf{x}' = \mathbf{l}'^\top[\mathbf{l}']_\times\mathtt{F}\mathbf{x} = 0$.

The parameter μ is not a very natural way to parametrize the homography. In practice it is more convenient to choose a point $\mathbf{x}$ in the first view, and a point $\mathbf{x}'$ on the epipolar line $\mathtt{F}\mathbf{x}$ in the second view. Then μ may be determined from the equation $\mathbf{x}' = \mathtt{H}(\mu)\mathbf{x} = \left([\mathbf{l}']_\times\mathtt{F} + \mu\mathbf{e}'\mathbf{l}^\top\right)\mathbf{x}$. A solution is obtained by multiplying by $[\mathbf{x}']_\times$, which gives μ as the ratio of vectors in $\mu(\mathbf{l}.\mathbf{x})[\mathbf{x}']_\times\mathbf{e}' = -[\mathbf{x}']_\times[\mathbf{l}']_\times\mathtt{F}\mathbf{x}$. This parametrization is used in the line matching implementation.

2.3 Corresponding conics

In this section we consider the reconstruction of a conic in 3-space given its image in two views and the associated fundamental matrix $\mathtt{F}$. An explicit and simple expression is given for the plane of the conic and the homography it induces. The geometry is illustrated in figure 3.

Several previous authors have investigated this geometry [35, 43, 49]. The most elegant is the approach of Quan [43] who shows, by reasoning on the back projected cones from each image conic, that there is a two fold ambiguity in the plane of the space conic. However, Quan's expressions are over complicated because the camera matrices are not written explicitly in terms of the epipolar geometry. Instead correspondence conditions for the conics are derived which are actually equivalent to the condition that the epipolar tangents to the conic must correspond (as in the Kruppa equations [36]). Full details are given in appendix A.3.

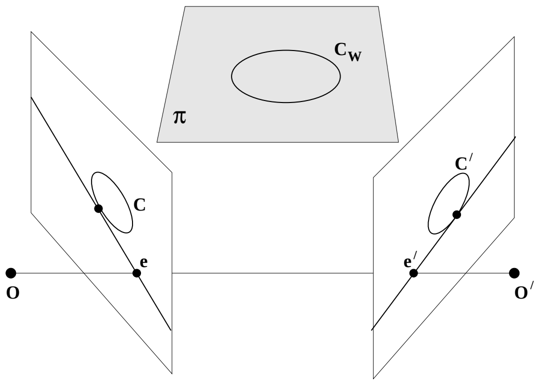


Figure 3: A conic defines a unique plane in 3-space. This plane is determined up to a two-fold ambiguity from the image of the conic in two views. One of the epipolar tangents is also shown.

We now give the main result of this section:

Proposition 2. *The homography induced by the plane of a conic defined by the correspondence* $\mathtt{C} \leftrightarrow \mathtt{C}'$ *is determined up to a two fold ambiguity as*

$$\mathtt{H}(\mu) = [\mathtt{C}'\mathbf{e}']_\times \mathtt{F} + \mu \mathbf{e}'(\mathtt{C}\mathbf{e})^\top \tag{6}$$

where the two values of μ *are obtained from*

$$\mu^2 \left[(\mathtt{C}\mathbf{e})(\mathtt{C}\mathbf{e})^\top - (\mathbf{e}^\top \mathtt{C}\mathbf{e})\mathtt{C}\right] (\mathbf{e}'^\top \mathtt{C}'\mathbf{e}') = \mathtt{F}^\top [\mathtt{C}'\mathbf{e}']_\times \mathtt{C}' [\mathtt{C}'\mathbf{e}']_\times \mathtt{F}$$

provided that $\mathbf{e}'^\top \mathtt{C}'\mathbf{e}' \neq 0$ *and* $\mathbf{e}^\top \mathtt{C}\mathbf{e} \neq 0$*. In a projective reconstruction with cameras* $\mathtt{P} = [\mathtt{I} \,|\, \mathbf{0}]$*,* $\mathtt{P}' = [\mathtt{A} \,|\, \mathbf{e}']$*, and* $\mathtt{F}$ *computed as* $\mathtt{F} = [\mathbf{e}']_\times \mathtt{A}$*, the plane has coordinates*

$$\boldsymbol{\pi}(\mu) = (\mathbf{e}'^\top \left(\mathtt{H}(\mu) + (\mathbf{e}'\mathtt{C}'\mathbf{e}')\mathtt{A}\right), (\mathbf{e}'\mathtt{C}'\mathbf{e}')||\mathbf{e}'||^2)^\top$$

The proof is given in appendix A.3. Note that when $\mathbf{e}'^\top \mathtt{C}'\mathbf{e}' = \mathbf{e}^\top \mathtt{C}\mathbf{e} = 0$ the epipole is on the conic in each image, and the baseline intersects the world conic. Although the algebraic expressions are not valid if the epipole lies on the conic, the plane can still be computed in this case, i.e. it is a problem with the expressions not the geometry. In this case the plane is determined uniquely.

It is somewhat unsatisfactory that the case of the epipole on the conic must be excluded in Proposition 2. An alternative solution, for which this degeneracy is avoided, is to proceed as follows: determine the one parameter family (pencil) of dual quadrics from the corresponding image conics as described in [14]; a conic is represented by a dual quadric of rank 3 (see [57]), and so may be determined as the degenerate members of this pencil; the 4-vector representing the plane of the conic is the null-vector of the degenerate dual quadric.

2.4 Curve osculating plane

In this section we show first how Euclidean curvature is mapped under a homography and then, based on this result, determine the osculating plane of a space curve from two views.

Proposition 3. *Suppose a plane curve is mapped under the homography* $\mathtt{H}$ *so that a point* $\mathbf{x}$ *is mapped to* $\mathbf{x}' = \mathtt{H}\mathbf{x}$*, and* κ, κ' *are the Euclidean curvatures of the curve at* $\mathbf{x}$ *and* $\mathbf{x}'$ *respectively, then*

$$\kappa' = \frac{\mathrm{sign}(|\mathtt{H}|)}{|\mathtt{H}|^2} \frac{x_3'^3}{(l_1'^2 + l_2'^2)^{3/2}} \kappa \tag{7}$$

where $|\mathtt{H}|$ *is the determinant of* $\mathtt{H}$*, and (with exact equality, not just to scale)* $x_3' = \mathbf{h}^3.\mathbf{x}$ *and* $\mathbf{l}' = \mathtt{H}^{-\top}\mathbf{l}$*, with the normalization* $l_1^2 + l_2^2 = 1, x_3 = 1$*. The notation* $\mathbf{h}^{i\top}$ *indicates the i'th row of* $\mathtt{H}$*, i.e.* $\mathtt{H} = [\mathbf{h}^1, \mathbf{h}^2, \mathbf{h}^3]^\top$*, and the line* $\mathbf{l} = (l_1, l_2, l_3)^\top$ *is tangent to the curve at* $\mathbf{x}$*.*

A proof is given in appendix A.4. Note, that the map depends both on the position ($\mathbf{x}$) and orientation ($\mathbf{l}$) of the curve. In the case that the homography $\mathtt{H}$ is an affine transformation (i.e. the third row of the matrix is $(0, 0, 1)$) there is no dependence on position.

Suppose a *plane* curve is imaged in two views, then the image curves are related by a homography induced by the curve plane. If the curve is not planar (it has non-zero torsion) then the image curves are not related globally by a homography. However, the osculating plane of the space curve may be used to define a homography. The geometry is illustrated in figure 4. We now show that the osculating plane and induced homography may be computed from the curvatures at corresponding image points on the imaged curve.

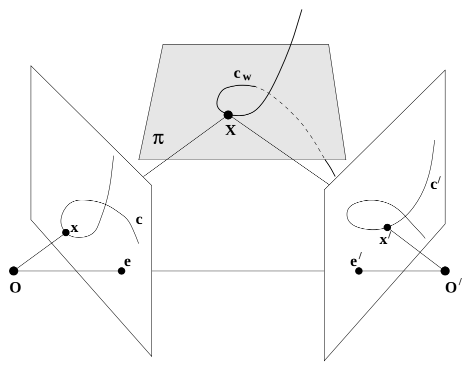


Figure 4: The osculating plane of a (non-planar) curve varies, but is always defined in 3-space provided the curvature is not zero. This plane is determined uniquely from the image of the curve in two views. The plane induces a homography between the images.

Suppose a point on a space curve $\mathbf{X}$ is imaged at $\mathbf{x}$ and $\mathbf{x}'$ in two views, and that the tangent lines on the imaged curves at $\mathbf{x}$, $\mathbf{x}'$, are $\mathbf{l}$ and $\mathbf{l}'$ respectively. The objective is to determine the homography $\mathtt{H}$ induced by the osculating plane at $\mathbf{X}$. Since tangency is preserved under projective transformations, the tangent lines are related as $\mathbf{l}' = \mathtt{H}^{-\top}\mathbf{l}$. From (5) this line correspondence determines $\mathtt{H}$ up to a one parameter family $\mathtt{H}(\mu) = [\mathbf{l}']_\times \mathtt{F} + \mu \mathbf{e}'\mathbf{l}^\top$ parametrized by μ. The remaining degree of freedom, μ, is determined *uniquely* using the relation (7) between image curvatures κ and κ' at $\mathbf{x}$ and $\mathbf{x}'$ respectively.

Proposition 4. *Given two perspective images of a* plane *curve, and the fundamental matrix for the view pair, then the plane of the curve (and consequently the homography induced by this plane) is defined uniquely by the corresponding tangent lines and curvatures at one point.*

Similarly, given two perspective images of a space *curve, and the fundamental matrix for the view pair, then the osculating plane at a point* $\mathbf{X}$ *on the space curve is determined uniquely from the tangent lines and curvatures at the corresponding points* $\mathbf{x}$ *and* $\mathbf{x}'$ *on the imaged curves.*

Writing $\mathtt{A} = [\mathbf{l}']_\times \mathtt{F}$*, the homography is* $\mathtt{H}(\mu) = \mathtt{A} + \mu \mathbf{e}'\mathbf{l}^\top$*, where* μ *is obtained from*

$$\mu = \frac{(b_1^2 + b_2^2)^{3/2}\kappa'}{|\mathtt{A} + \mathbf{e}'\mathbf{l}^\top|(\mathbf{a}^3.\mathbf{x})^3\kappa}$$

where $\mathbf{a}^{i\top}$ *is the i'th row of* $\mathtt{A}$*, i.e.* $\mathtt{A} = [\mathbf{a}^1\mathbf{a}^2\mathbf{a}^3]^\top$*,* $\mathbf{b} = \mathrm{adj}\,(\mathtt{A})^\top\mathbf{l}$ *with* $\mathrm{adj}\,(\mathtt{A})$ *the adjoint of* $\mathtt{A}$ *which is given by* $\mathrm{adj}\,(\mathtt{A}) = [\mathbf{a}^2 \times \mathbf{a}^3 \quad \mathbf{a}^3 \times \mathbf{a}^1 \quad \mathbf{a}^1 \times \mathbf{a}^2]$*, and* $\mathbf{l}$*,* $\mathbf{x}$ *are as defined in Proposition 3.*

The proof is given in appendix A.4. The result builds on previous work on the perspective imaging of the curvature of a space curve in a single view [13], and on invariants of space curves determined from two perspective images [20, 59].

There are two types of degenerate points at which the plane/homography cannot be uniquely determined by this method. The first is at epipolar tangents, where the tangent lines contain the epipole (i.e. if $\mathbf{l}.\mathbf{e} = \mathbf{l}'.\mathbf{e}' = 0$), since in this case the line in 3-space which projects to $\mathbf{l}$ and $\mathbf{l}'$ can not be determined as it lies in an epipolar plane. The second is at zero's of curvature (inflections) since then the curvature cannot be used to determine μ.

2.5 Rejection of the incorrect solution for corresponding conics

If the algebraic conic method of Proposition 2 is used to compute the homography (or plane) for corresponding conics, then two solutions result. The homography may also be computed using the local curvature method of Proposition 4, by selecting a point on the conic in the first view and obtaining the corresponding point in the second view by intersecting the corresponding conic with the point's epipolar line. Since there are two intersections of this line with the conic, again there are two solutions for the homography. However, the fallacious solutions originate for different reasons in the two methods. Consequently, the fallacious algebraic solution can be eliminated by also using the curvature method: the correct solutions from both methods will be the same, but in general the fallacious solutions differ.

Figure 5 shows the reprojection into a new view of the two solutions obtained using the local curvature. Depending on which point in the first image is used, the second (incorrect) solution varies.

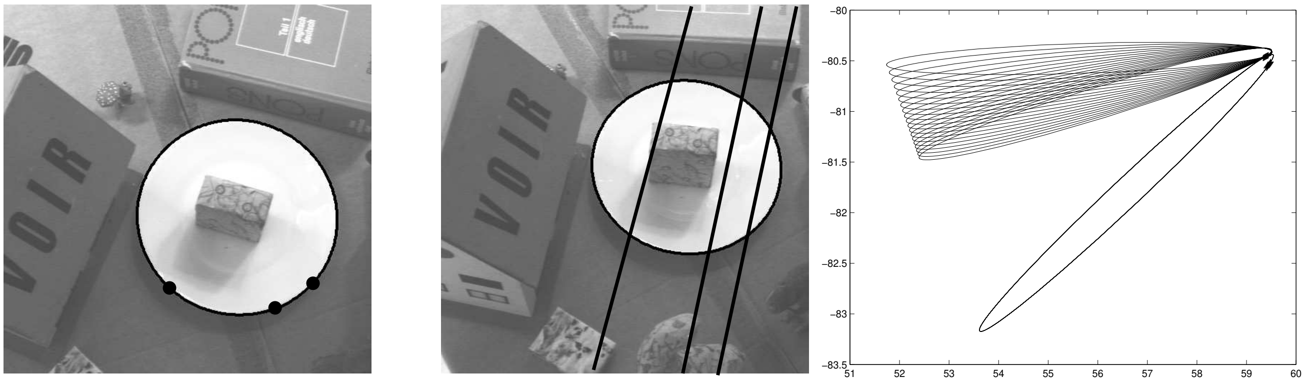
2.6 Transfer to a third view

The previous sections have shown that given corresponding images of entities (such as a conic) on a plane in the first two views, and the fundamental matrix between those views, then the homography induced by the plane may be determined. This enables other entities on the plane (such as points) to be transferred between views. For example given the image of a point on the plane in the first view, its image in the second view may be computed.

In this section we extend this construction to a third view. It is shown that given the trifocal tensor for three views, and the homography between two of the views induced by a plane, then the homography induced by this plane may be computed between any of the other pairs of views.

As an illustration suppose the homography is between the first two views, and we require the homography between the first and third. Then:

Proposition 5. *Given the trifocal tensor for three views, and a homography* $\mathtt{H}_{12}$ *induced by a plane between the first and second view, (so that points map as* $\mathbf{x}' = \mathtt{H}_{12}\mathbf{x}$*), then the homography*



a b c

Figure 5: (a) Three points on the conic in the first view, and (b) their corresponding epipolar lines in the second. For each of the two intersections of an epipolar line with the conic the plane of the conic is computed using the curvature method of section 2.4. For these three points (and others on the conic), figure (c) shows the two space conic solutions reprojected into a new view. The correct solution (the lower conic), arising from the correct point correspondence, is the same for each point in the first view, but the conic transferred by the incorrect solution varies. The envelope of these conics is the image outline of the cone of rays from the first camera centre to the world conic.

between the first and third views $\mathtt{H}_{13}$ *(so that points map as* $\mathbf{x}'' = \mathtt{H}_{13}\mathbf{x}$*), is* $\mathtt{H}_{13} = \mathtt{B} + \mathbf{e}'\mathbf{v}^\top$*; where the camera matrices*

$$\mathtt{P} = [\mathtt{I} \mid \mathbf{0}] \qquad \mathtt{P}' = [\mathtt{A} \mid \mathbf{e}'] \qquad \mathtt{P}'' = [\mathtt{B} \mid \mathbf{e}'']$$

are obtained from the trifocal tensor as described by Hartley [26]; the 3-vector $\mathbf{v}$ *is obtained from* $\mathtt{P}, \mathtt{P}'$ *and* $\mathtt{H}_{12}$ *using (2); and the expression for* $\mathtt{H}_{13}$ *follows from (1) for the cameras* $\mathtt{P}, \mathtt{P}''$*.*

This construction may clearly be varied to compute the homography between any two views given the homography between another pair.

A useful application of this result is to determine the image of an entity in a third view, given its image in two others. In the case of a conic:

Proposition 6. *Given the trifocal tensor for three views, and the perspective images of a conic in the first two, then the image of the conic in the third view may be computed up to a two fold ambiguity.*

As in Proposition 5 the camera matrices are determined from the trifocal tensor, and then the two solutions for the plane of the conic $\mathbf{v}$ are determined from the corresponding image conics $\mathtt{C} \leftrightarrow \mathtt{C}'$ with $\mathtt{P} = [\mathtt{I} \mid \mathbf{0}]$, $\mathtt{P}' = [\mathtt{A} \mid \mathbf{e}']$, as described in Proposition 2. The homography between the first and third views is $\mathtt{H}_{13} = \mathtt{B} + \mathbf{e}''\mathbf{v}^\top$, and from (4) the conic $\mathtt{C}$ is transferred as $\mathtt{C}'' = \mathtt{H}_{13}{}^{-\top}\mathtt{C}\,\mathtt{H}_{13}^{-1}$. The transferred conic $\mathtt{C}''$ is the sought image of the world conic in the third view. In general the transferred conics for each solution of $\mathbf{v}$ will differ.

In a similar manner, the curvature of a curve in a third view can be computed given its curvature and tangent at corresponding points in the two other views and the trifocal tensor for the image triplet. The plane of the curve is computed using Proposition 4 and the curvature transferred using Proposition 3.

In the case of corresponding lines in two views, the image in a third view can also be computed in the same manner: from Proposition 1 there is a pencil of planes determined between the first two views; and correspondingly a pencil of planes and induced homographies determined between the first and third views using Proposition 5. Provided particular degenerate homographies are avoided (at $\mu = \infty$, where the plane contains the first camera centre) then, as is evident from figure 1, the line may be transferred from the first to third views for any value of μ. However, it is not necessary to proceed in this way as the line can be transferred directly from two views to a third by the trifocal tensor [26]. In the case that a line is transferred from the first view to the third by the trifocal tensor, this is identical to transfer by a homography induced by a plane defined by the second camera centre and $\mathbf{l}'$ (i.e. the homography for $\mu = 0$).

Incidentally, Faugeras and Robert [21] derive a very nice, simple, relationship between the imaged curvatures at corresponding points in three views. They show that $a_1\kappa + a_2\kappa' + a_3\kappa'' = 0$, i.e. that the three imaged curvatures are linearly related, with no constant term. However, their derivation is "quite" long. In appendix A.5 it is shown that this result may be derived in a few lines from Proposition 4 and Proposition 5.

2.7 Relation to epipolar transfer

In [21] Faugeras and Robert developed a method for transferring conics and curvature from two views to a third based on epipolar transfer of points. In epipolar point transfer the objective is to compute the image $\mathbf{x}''$ in the third view of a 3D point given its images $\mathbf{x}, \mathbf{x}'$ in the first and second views. The point $\mathbf{x}''$ lies on the epipolar line $\mathbf{l}''^{e}_{\mathbf{x}} = \mathtt{F}_{31}\mathbf{x}$ corresponding to the point $\mathbf{x}$ in the first view. Also, $\mathbf{x}''$ lies on the epipolar line $\mathbf{l}''^{e}_{\mathbf{x}'} = \mathtt{F}_{32}\mathbf{x}'$ corresponding to the point $\mathbf{x}'$ in the second view. Thus, the point $\mathbf{x}''$ may be computed as the intersection of these two epipolar lines $\mathbf{l}''^{e}_{\mathbf{x}}$ and $\mathbf{l}''^{e}_{\mathbf{x}'}$.

However, epipolar transfer has a serious degeneracy [28, 65]. If points lie in the trifocal plane (the plane defined by the three camera centres) then the epipolar lines $\mathbf{l}''^{e}_{\mathbf{x}}$ and $\mathbf{l}''^{e}_{\mathbf{x}'}$ are coincident and the point $\mathbf{x}''$ cannot be computed from their intersection. Furthermore, in the case that the three camera centres are collinear epipolar transfer is degenerate for all points (the trifocal plane degenerates to a family of planes in this case which covers 3-space). Even for situations which are not exactly degenerate the computation of the intersection becomes increasingly ill-conditioned as the epipolar lines become less "transverse".

Epipolar transfer underpins the transfer methods of [21]. For instance, in [21] a line is transferred from the first two views to a third by selecting two points on the line $\mathbf{l}$ in the first view, computing the corresponding points in the second view by intersecting $\mathbf{l}'$ with the epipolar lines from the points in the first (similar to figure 12), and then transferring these points from the first and second views to the third. This defines two points on the line in the third view, and hence transfers the line. Clearly if epipolar point transfer fails, then the line cannot be transferred by this method.

The transfer method of Proposition 5 using a planar homography does not have the degeneracy of epipolar transfer (nor does transfer using the trifocal tensor).

3 Geometry Assessment

In this section we assess the performance of the methods of section 2 for an implementation on real images. The accuracy is evaluated from three views of a curve as follows: the plane of the curve is estimated from views one and two, and the curve and curvature are then transferred by the induced homography into the third view. The *transfer error* is measured by comparing the transferred curve/curvature with that of the actual curve imaged in the third view. The sensitivity of the different methods to noise in the contour extraction is evaluated in section 3.4.

Assessment Criteria. The transfer is assessed by two error measures. The first, ϵ_{x}, is the average Euclidean distance between the curves

$$\epsilon_{\mathrm{x}} = < d_{\perp}(\mathbf{c}(s), \mathbf{c}_{\mathrm{t}}(s)) >$$

where $d_{\perp}$ is the distance between the curves in the normal direction from the imaged curve $\mathbf{c}$ to the transferred curve $\mathbf{c}_{\mathrm{t}}$, and the average is performed over an arc length curve parameter s for the curve $\mathbf{c}$. The correspondence between the curves is established by intersecting the curve $\mathbf{c}_{\mathrm{t}}$ with the normal of $\mathbf{c}(s)$. The second measure, ϵ_{κ}, is the average relative curvature difference of the curves

$$\epsilon_{\kappa} = < \frac{|\kappa(s) - \kappa_{\mathrm{t}}(s)|}{\kappa(s)} >$$

where κ is the curvature of the imaged curve, and κ_{t} the curvature of the transferred curve. Points at which transfer is not possible, due to degeneracies in determining the osculating plane (see section 2.4), are excluded from the average.

The method is assessed for both conics (i.e. plane curves), and space curves. A conic allows the conic-specific method of section 2.3 to be compared to the general curve method of section 2.4 based on curvature. A conic also allows the curvature method to be assessed in two ways. In the first case a conic is fitted to the extracted conic boundary (edge chain) and intersections, tangents, and curvature are determined from the algebraic curve. This is a *global* method of estimation. In the second case the curve, its tangent, and curvature are determined locally (by spline fitting) using only this *local* information. In general we would expect the global estimates to be superior to the local.

For the plane curves the evaluation images consist of the two image triplets shown in figure 6, with the plate providing a conic. The triplets share the same first two views, but differ in the third. In triplet II the baseline between the second and third views is four times larger than that in triplet I. Space curves are evaluated using the image triplet shown in figure 7. The space curves used for the evaluation are superimposed; one is on the bottle, the second on the mug.

a b c d

Figure 6: Frames 3, 6, 7 and 10 from the "plate" sequence. The extracted conics are superimposed in black. Test triplet I consists of images (a), (b), (c); and test triplet II of images (a), (b), (d).

Implementation details. Conic outlines are extracted using a subpixel Canny edge detector, and the conic estimated by Bookstein's fitting algorithm [8] on normalized edgel coordinates [48]. The local estimate of curvature is obtained by fitting cubic B-splines to the edgel chains. The

Figure 7: The "bottle" sequence (images 11,15 and 19). Two of the extracted contour chains are superimposed.

fitting algorithm [33] takes into account the distance between contour and spline points as well as a regularity measure based on the third derivatives. The cost function minimized in order to determine the B-spline $\mathbf{s}(t)$ is

$$\sum_{i=1}^{n} |\mathbf{s}(t_i) - \mathbf{p}_i|^2 + \tau \int |\frac{\partial^3 \mathbf{s}(t)}{\partial^3 t}|^2 dt$$

where $\mathbf{p}_i$ are the contour points, $\mathbf{s}(t_i)$ the corresponding points on the spline and τ is a positive parameter which controls the regularity of the solution.

The average involved in both error measures, ϵ_{X} and ϵ_{κ}, is computed by sampling the conic/B-spline curve at one pixel arc length intervals.

The fundamental matrix and trifocal tensor for these examples are computed automatically from the images using the methods described in [5, 28, 56, 64].

3.1 Assessing conic transfer

The conic plane is computed from the result of Proposition 2 of section 2.3 using conics from views one and two. In order to have conics consistent with the epipolar geometry, a correction is required to ensure that the epipolar tangents to the conic (the tangent lines from the epipole to the conic) correspond between the images. This correction is achieved by applying the algorithm of [14]. The corrected conic is then transferred into the third image as described in Proposition 5 of section 2.6. The distance error and the relative curvature error are computed between the transferred and the extracted conic in the third image. The second solution for the conic plane is easily ruled out as the distance error is very large.

For triplet I, $\epsilon_{\mathrm{X}} = 0.15$ pixels, and $\epsilon_{\kappa} = 0.0024$. For triplet II, $\epsilon_{\mathrm{X}} = 0.28$ pixels, and $\epsilon_{\kappa} = 0.0074$. Clearly, the transfer is excellent (visually the conics are indistinguishable.).

3.2 Assessing curvature transfer

Global curvature estimation. Tangent lines and curvatures are obtained for the conics in each view by implicit differentiation of the fitted conics. The correspondence between points on the conic in the first and second view is obtained by intersecting the conic in the second view with the epipolar line of the point in the first view. The conic epipolar tangents and their neighbours are excluded from the error measures since curvature cannot be transferred for these points.

For triplet I, $\epsilon_{\mathrm{X}} = 0.15$ pixels, and $\epsilon_{\kappa} = 0.0016$. For triplet II, $\epsilon_{\mathrm{X}} = 0.34$ pixels, and $\epsilon_{\kappa} = 0.0047$. These results are almost identical to those of the previous section where conic transfer is used (Proposition 2), as opposed to conic curvature transfer (Proposition 4). They demonstrate that given almost perfect curvature (from the fitted conic) the curvature transfer is similarly almost perfect. We next investigate the deterioration when the curvature measurement is local and inferior.

Local curvature estimation. The conics in image one and two are here represented by B-splines fitted to the edgel chain. The tangents, curvatures and epipolar line intersections are then computed from the B-splines, as opposed to the conics. The transferred curvature is compared to the curvature computed implicitly from $\mathtt{C}''$. For triplet I, $\epsilon_{\mathrm{X}} = 0.57$ pixels, and $\epsilon_{\kappa} = 0.025$. For triplet II, $\epsilon_{\mathrm{X}} = 0.63$ pixels, and $\epsilon_{\kappa} = 0.035$. For the space curves of figure 7 B-splines are used in all cases and $\epsilon_{\mathrm{X}} = 0.43$ pixels, $\epsilon_{\kappa} = 0.037$ for the curve on the bottle, and $\epsilon_{\mathrm{X}} = 0.71$ pixels, $\epsilon_{\kappa} = 0.097$ for the curve on the mug. This is a very similar performance to the case of B-spline curvature computation for plane curves.

Comparison of local and global estimation. The position error ϵ_{X} has a similar value for both global and local curvature estimation. This is not surprising because the point transfer error only arises because of the difference between the fitted curves (conic or B-spline) and the veridical conic. This will be consistently sub-pixel because of the sub-pixel acuity of the Canny edge detector used.

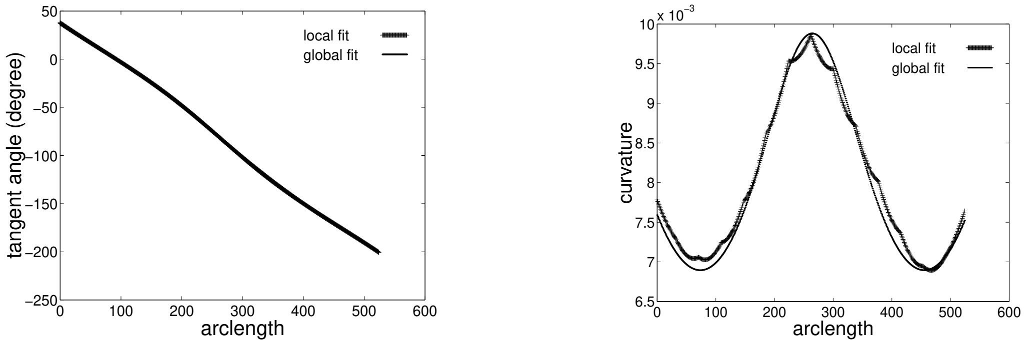
The curvature error ϵ_{κ} is more illuminating. It is an order of magnitude larger when using local estimation of curvatures (via B-splines). This increased error is a direct consequence of the error in curvature estimation as is detailed in the following section.

3.3 The effects of curvature measurement error

In this section we compare in detail the transfer results obtained for triplet I, when using a global fit (via the conic) and a local fit (via the B-spline). The global fit is treated here as ground truth. First the curvature *measurement* is compared. Figure 8 shows that although measured tangent angle is indistinguishable there is a notable difference between measured curvature. This can be traced to the cubic splines used for the local curvature measurement which are C^1 continuous in tangent angle, but C^0 in curvature. Next, figure 9 compares the *transferred* tangent angles and curvatures using the curvatures computed by the two methods. The transferred tangent angles are indistinguishable, but the influence of the curvature measurement error on the transferred curvature is evident. The transfer error is about twice as large as the measurement error in one image, as it results from the curvature measurement errors in both the views used to compute the homography.

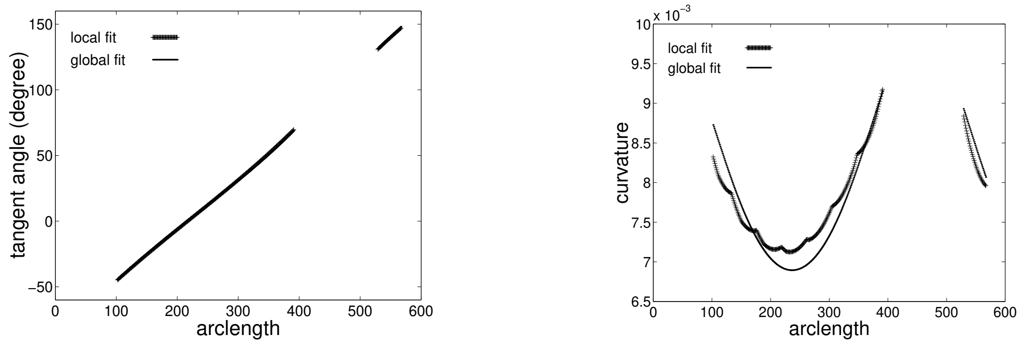
We now investigate the magnitude of the error in the computed homography. To assess this error, we introduce a third error measure $\epsilon_{\mathtt{H},D}$ which, like ϵ_{X}, measures transfer error but now for points *not* lying precisely on the conic. Instead a set of points $\mathbf{x}_D$ are chosen in the first view a distance D along the normal direction from the curve (so that the complete set forms a parallel curve). The points $\mathbf{x}_D$ are then transferred to the second view using the homography $\mathtt{H}_\kappa$ computed from the curvature. The points are also transferred using the homography $\mathtt{H}$ computed from the conic correspondence (Proposition 2), which gives the ground truth. The error is the Euclidean distance between the points transferred by $\mathtt{H}_\kappa$ and $\mathtt{H}$

$$\epsilon_{\mathtt{H},D} = d(\mathtt{H}_\kappa \mathbf{x}_D, \mathtt{H}\mathbf{x}_D)$$



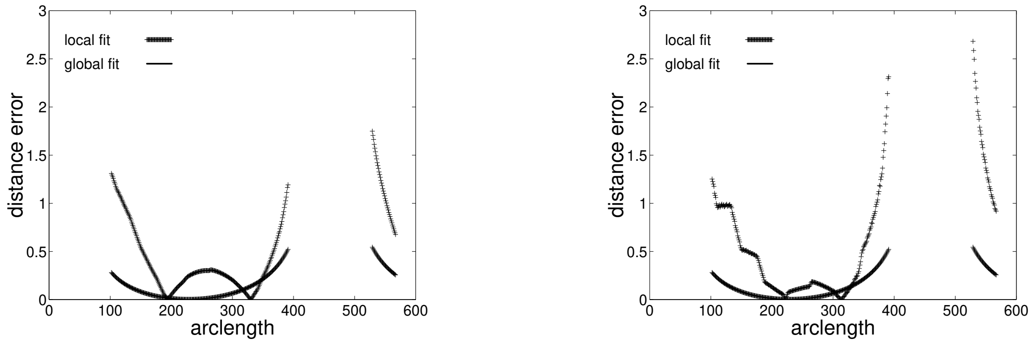
a b

Figure 8: (a) Comparison of tangent angle computed from a local fit (via B-splines), and a global fit (via the conic). (b) Comparison of curvature computed from a local and global fit. Arc length is in pixels. The results are for the conic in the first view of the “plate” sequence of figure 6.



a b

Figure 9: (a) Comparison of the transferred tangent angle using a local fit (via B-splines) and a global fit (via the conic). The curves are indistinguishable. (b) Comparison of the transferred curvature using a local and a global fit. Arc length is in pixels. The results are for triplet I. The missing curve segments are at epipolar tangents where the curvature cannot be transferred.



a b

Figure 10: Comparison of the distance error $\epsilon_{\mathtt{H},D}$ as a function of arclength (pixels) for triplet I. The homographies are computed with curvatures obtained by a local fit (via B-splines) and a global fit (via the conic). The points are at a distance D pixels from the conic in the first view of the triplet with $D = 2$ in (a), and $D = 10$ in (b). The missing curve segments are at epipolar tangents where the curvature cannot be transferred.

Figure 10 shows the error measure $\epsilon_{\mathtt{H},D}$ as a function of the arclength for the homographies computed from the global/local curvature (via conic/B-spline) for $D = 2$ and $D = 10$ pixels. For points which are not in the neighbourhood of the epipolar tangents, results are very good: less than one pixel error for a distance D of 10 pixels. This explains the high quality of the wide-base line results of section 4.3 – a one pixel error doesn't significantly influence cross-correlation results. The errors introduced by the local curvature measurements are also evident (compare points with the same arc length parameter between figures 9 and 10). However, the dominant effect — which introduces far larger errors than the curvature measurements – is the proximity to epipolar tangents.

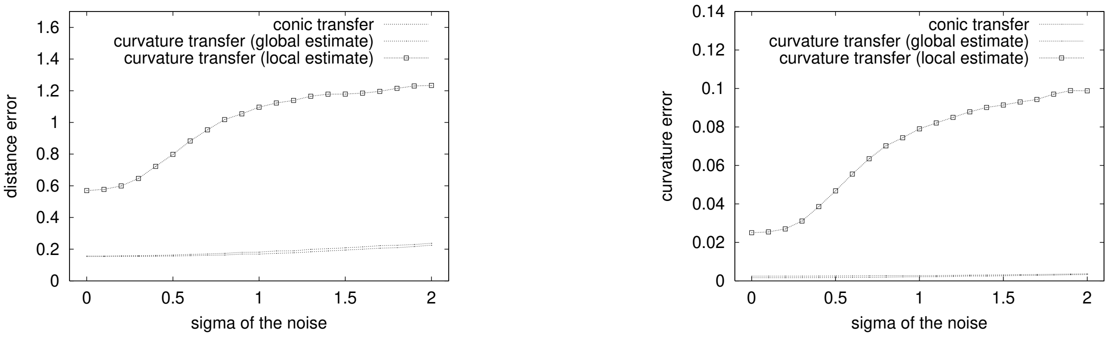
3.4 Sensitivity analysis

To examine the sensitivity of the method, Gaussian noise with varying σ is added to the position of the conic edgels extracted from the real images (triplet I). Position error ϵ_{x} and curvature error ϵ_{κ} are then computed for the three different transfer methods: conic transfer and curvature transfer with global and local curvature estimation. The error measures are averaged over a 1000 trials for each noise level σ. The average distance and curvature errors as a function of σ are displayed in figure 11 for σ varying from 0.0 to 2.0. The two conic based methods are hardly influenced by the noise and are indistinguishable. In the case of a local curvature estimation the results degrade slowly. For $\sigma = 1$ pixel, the distance error is about 1 pixel and the curvature error 8%.

To summarise: the results of this section demonstrate that conic and curvature transfer are viable and not particularly noise sensitive provided care is paid to curvature measurement. The dominant error in curvature transfer is in fact the degeneracy which occurs at epipolar tangents.

4 Matching lines and curves over multiple views

In this section we describe a set of algorithms for matching lines and curves between images of a static scene for which the multi-view relations, such as the epipolar and trifocal geometry, are



a b

Figure 11: Tolerance to noise. (a) Distance error ϵ_{x} and (b) curvature error ϵ_{κ} as a function of the σ of the added Gaussian noise. Errors are computed for conic transfer, curvature transfer with curvature estimated from the conic, and curvature transfer with curvature estimated from B-splines. The two conic based methods are hardly influenced by the noise and are indistinguishable.

available. In the case of a calibrated stereo/trinocular rig these relations are known in advance. The relations may also be computed automatically from the images themselves using interest point correspondences [5, 42, 55, 56, 64].

The problem is the following: lines and curves are detected in each image; this extracted geometry will certainly not be perfect in that there will be gaps and fragmentation, missing segments, erroneous additional segments, etc. The objective is to establish correspondences across images between the extracted lines. The image lines correspond if they are the image of the same line in 3-space. A similar objective applies to curves. Once correspondences are established then the 3D structure of the lines and curves is determined.

4.1 Two view matching

We first review the geometric constraints available, which are now fairly well established, and then describe the photometric constraints and their integration to provide disambiguation in the matching. It is this interplay between the geometric and photometric constraints which is essential for successful matching.

4.1.1 Geometric constraints

When matching lines over two views, there is a weak overlap constraint for line segments of finite length due to epipolar geometry. The two end-points of a segment generate two epipolar lines in the other image. These two lines define a region, called the epipolar "beam", which necessarily intersects or contains the corresponding segment [63]. Figure 12 illustrates this idea.

The epipolar beam reduces the combinatorics of line segment matching since it restricts which line segments need to be considered for matching. The reduction is very much dependent on the scene and motion, but in typical examples (see section 4.2.2) only a third of the line segments need be considered.

For curves over two views, the epipolar beam may be used in a similar manner. For example, by checking bounds on the curve, such as the convex hull, against the epipolar beam. There

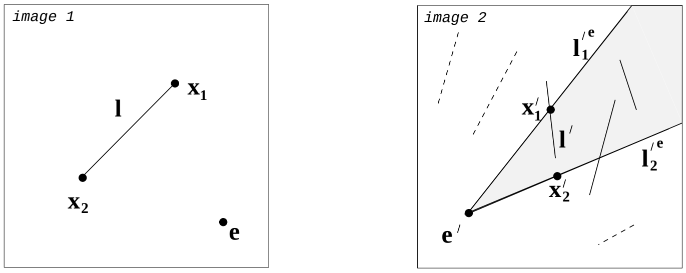


Figure 12: The epipolar geometry reduces the search space for line segments corresponding to $\mathbf{l}$. The segments have to lie in the region of the epipolar beam [63] formed by $\mathbf{l}'^e_1$ and $\mathbf{l}'^e_2$, which are the epipolar lines corresponding to $\mathbf{x}_1$ and $\mathbf{x}_2$ respectively.

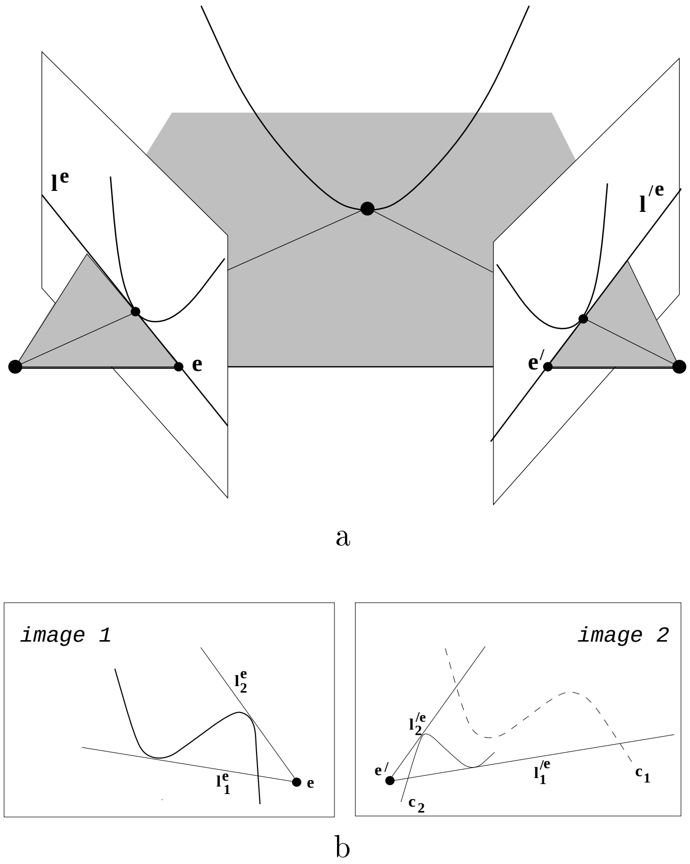


Figure 13: **Epipolar tangencies**. (a) The epipolar plane is tangent to the space curve. The corresponding epipolar lines $\mathbf{l}^e \leftrightarrow \mathbf{l}'^e$ are tangent to the imaged curve. (b) Suppose the epipolar lines $\mathbf{l}^e_1$, $\mathbf{l}^e_2$ are tangent to the imaged curve in the first view, and the lines $\mathbf{l}'^e_1$, $\mathbf{l}'^e_2$ are the corresponding epipolar lines in the second view. Then in the second view the curve $\mathbf{c}_1$ (dashed) is not tangent to these epipolar lines so it cannot match the imaged curve in the first view. However, $\mathbf{c}_2$ (solid) is a possible match.

is also an additional constraint from the epipolar geometry: epipolar tangents correspond [41]. Figure 13 illustrates this constraint.

4.1.2 Photometric constraints

From here on "curves" will cover both curves and lines. Suppose that for each curve in the first image there is a set of possible corresponding curves in the second image, all of which satisfy the weak geometric constraints above. Other than the ordering constraint, which may not apply, there is no further geometric information available to disambiguate the matches.

However, there is photometric information available from the surfaces in the neighbourhood of the curve: consider a strip on the surface on either side of the curve; this is imaged in the two views on either side of the image curve. Under reasonable imaging conditions the intensity of these imaged strips will be 'similar' across viewpoints. This similarity of the intensity neighbourhood for corresponding curves, or conversely the lack of similarity for non-corresponding curves, provides a photometric measure that may be used to rank curve matches, and hence disambiguate matches.

Consider two possibly corresponding curves $\mathbf{c}$ and $\mathbf{c}'$ in the first and second images respectively. If they are corresponding, then a point to point correspondence on the curves may be determined using the epipolar geometry : for an image point $\mathbf{x}$ on the curve $\mathbf{c}$, the epipolar line in the second image is $\mathbf{l}'^e = \mathtt{F}\mathbf{x}$, and this line intersects the curve $\mathbf{c}'$ in the point $\mathbf{x}'$ corresponding to $\mathbf{x}$, i.e. $\mathbf{x}$ and $\mathbf{x}'$ are images of the same 3D point. A similarity score for the *curves* $\mathbf{c}$ and $\mathbf{c}'$ may then be determined by averaging the similarity of neighbourhoods for all corresponding *points* on the curves. The similarity measure should be invariant to a local affine transformation of intensities on the neighbourhood, $I \rightarrow \alpha I + \beta$, as might occur due to automatic gain control, changing illumination conditions etc.

If the viewpoints are sufficiently close then the similarity of neighbourhoods may be measured by normalized cross-correlation. This measure is invariant to an affine intensity change over the spatial neighbourhood used. If the curves are indeed corresponding, then the similarity score for the curves will be high — certainly in general it will be higher than the score for images of two different 3D curves. Figure 15 demonstrates this for a pair of corresponding lines, and a pair of lines which do not correspond: the correct match between lines a in figure 15 has a normalized cross-correlation score, averaged over the pixels of the line, of 0.92, compared to the incorrect match of a to b with a score of -0.67, i.e. a significant difference.

4.1.3 Photometric constraints over a wide baseline

If there is a significant rotation of the camera or a wide baseline between views, then a simple cross-correlation of image intensities will fail as a measure of the similarity of the neighbourhoods of corresponding image points on the curve. Think of a camera motion consisting of a translation parallel to the image x-axis, followed by a 90^o rotation about the camera principal axis (i.e. a rotation axis perpendicular to the image plane). The cross-correlation of the neighbourhoods will be very low if the rotation is not corrected for. More generally, suppose a square neighbourhood in one image back-projects to a planar facet in the scene. The image of this region in the second image is a quadrilateral, but its shape depends entirely on the relative positioning of the cameras and plane. Consequently both the orientation and shape of the corresponding region must be corrected for.

Suppose the 3D curve lies on a surface, then the rotation, and in general all perspective distortion, can be corrected for, if the cross-correlation is computed as follows: for each point in the intensity neighbourhood of $\mathbf{x}$ in the first image, compute the intersection of the back projected ray with the surface, then determine the image of this intersection in the second view. The surface defines a mapping between the neighbourhoods of $\mathbf{x}$ and $\mathbf{x}'$, and the cross-correlation is computed over points related by this map. We don't know the surface, and so don't know this map, but a very good approximation to the map is provided by the homography induced by the tangent plane of the surface at the 3D point of interest. Indeed in piecewise planar environments such as rooms or aerial views of urban scenes the local tangent plane coincides with the planar surface, so no approximation is involved.

In the case of line matching this homography can only be determined up to a one parameter family because a line in 3D only determines the plane inducing the homography up to a one parameter family (figure 1). This means that for lines a one dimensional search over homographies is required. However, in the case of curve matching, the curve osculating plane can provide a homography which is uniquely determined by the curve correspondence (Proposition 4). It will be seen that the osculating plane homography is sufficient in practice to correct for the rotations of wide baseline views, and no search over homographies is required.

The idea of using a point-to-point map to correctly measure the cross-correlation is not new – indeed Gruen proposed this in the Photogrammetry literature in 1985 [24]. However, the map in that case was an affinity, and its computation required a search over six parameters. The benefit of using corresponding lines so that a homography (in general specified by eight parameters) could be computed by a one-parameter search was first described in Schmid and Zisserman [50], and the benefit of corresponding curves (no search) was first described in Schmid and Zisserman [51]. Others have used affine transformations consistent with the epipolar geometry and thus requiring only three parameters to be specified [19, 44].

In the following we describe two matching algorithms, one for short baseline view pairs and the other for wide base line views. The cases are distinguished by the severity of the differing perspective distortions in the images. In both cases a photometric similarity measure between putatively matched curves is obtained from an aggregated neighbourhood cross-correlation score. In the short baseline case simple neighbourhood cross-correlation of corresponding points suffices; in the wide baseline case corresponding points for the cross-correlation are determined by a local planar homography.

4.2 Short baseline matching between two views

If the viewpoints are sufficiently close then the similarity between curves can be measured as described in section 4.1.2. Again here "curves" includes both curves and lines. The four steps of the algorithm are:

1. For each curve in the first image determine the curves in the second image which satisfy the epipolar constraints of section 4.1.1. Those that do are putative curve matches.

2. For each putative curve match determine the similarity score based on normalized cross-correlation as described in section 4.1.2.

3. The matches are decided by a winner takes all scheme based on the similarity scores. Each curve in the first image can only match one curve in the second.

4. The corresponding parts of the line segments/curves are determined.

4.2.1 Implementation details.

Feature extraction. Curve segments are extracted by a local implementation of the Canny edge detector at sub-pixel accuracy. Edgels are then linked into chains, jumping up to a one pixel gap. Tangent discontinuities in the chain are located using a worm, and a line fit between the discontinuities is estimated using orthogonal regression. A very tight threshold is used for the line fitting so that curves are not piecewise linear approximated. Chains which are not fitted as lines are curves.

Similarity score. The similarity score for a pair of curve segments $\mathbf{c}$ and $\mathbf{c}'$ is computed as the average of the individual edgel correlation values, with point correspondences determined by the epipolar geometry. In the case of multiple intersections of a curve and epipolar line the best correlation value is included in the overall score. A correlation window of 15×15 pixels is used. There is a problem at epipolar tangents in determining the corresponding edgel when several edgels of the curve lie along the epipolar line. In these cases figural continuity is used to determine the match.

To be robust to occlusion an individual correlation value is only included if above a threshold. Here this threshold value is 0.6. This value has been found empirically to be suitable for a wide variety of scenes and camera motions. If there are fewer than a minimum number of matched edgels for a putative segment match, then that match is eliminated. Here only segments of length 15 pixels or more are included.

Corresponding parts. For a putative line match, corresponding parts can be determined simply by intersecting with the epipolar lines. For a putative curve match, determining corresponding parts is more complicated for several reasons. First, it is often the case that curves are erroneously joined to each other (e.g. at a 'T'-junction) so that a curve in one image only matches parts of a contiguous curve in the other image. Second, the point to point correspondence is often not unique, as multiple intersections may occur on the same curve. Third, the correspondence between curve points is lost at an epipolar tangent. To determine corresponding parts we first determine point pairs for which the epipolar constraint is valid and for which the correlation score is above a threshold. Points are then ordered into subchains using the ordering constraints and figural continuity. The three longest putatively corresponding subchains (10 pixels or more) are retained.

Algorithm complexity. If n is the average number of line segments in an image and l the average length of a line segment in pixels, the complexity of the 2-view matching algorithm is $\alpha l n^2$, where α is the proportion of lines included by the epipolar beam. Typically, the length l of a line segment is of the same order of magnitude as n, so that the algorithm is $\mathcal{O}(n^3)$ in practice.

The complexity of curve matching is of the same order, that is $\mathcal{O}(n^3)$. The only difference is a multiplicative factor. Instead of using the average length of the line, we have to multiply l by the average number of intersection s of the epipolar line with the curve. As $s << l$, the order of complexity doesn't change.

4.2.2 Matching results

In the following we present examples of four different scenes using the short baseline algorithm for matching line segments and curves. At present the ground-truth matches are assessed by hand.

Line matching. For the first two examples only line segments are matched. Figure 14 shows matching on aerial images of buildings, and figure 15 on images of a model house. In both these examples about 50% of the detected lines are matched, and of these matches over 95% are correct. This is an excellent performance : about 35% of the lines detected in one image are not detected in the other image, so although only 50% of the *detected* lines are matched, 77% of the lines which could be matched are indeed matched by the algorithm.

The epipolar beam reduces the search complexity (i.e. the number of putative matches that must be considered) to about 1/3 for the aerial images and about 1/5 for the house images.

Curve matching. For the bottle example (cf. figure 16) curve segments are matched. The matched segments are displayed in two different ways: first, the complete edgel chain matches are shown; second, only the parts of the matched chains for which there are corresponding edgels in both views are shown. The latter choice excludes the parts of the chain which are detected as edgels in one view but not in the other. Only corresponding parts are shown for the rest of the examples in this paper. Figure 17 shows curve matching for the plate sequence. In the bottle example one of the matched chains is incorrect. In the plate example all the matches are correct.

In general there is a performance trade off between the "quality" of the edgel chains and the number of matches and mismatches. If only long edgel chains with high contrast are detected then there are fewer mismatches (usually none at all) but also fewer curves to match. Conversely, if shorter and lower contrast edgel chains are detected then there are many more curves to be matched, and often 90% of these will be matched correctly. Examples of this trade off are given in [51].

It is worth remarking on the pros and cons of computing separate scores for each side of the curve or using a correlation window which straddles the curve (similarly for lines). Empirically it is found in the short baseline case that it is preferable *not* to distinguish the sides. This is because a correlation window that overlaps the curve includes the edge gradient, and the edge gradient is quite discriminating between curves. However, in the wide baseline case, described next, the two sides are treated separately because occlusions can severely change the similarity score for one side. The non-occluded side then provides disambiguation.

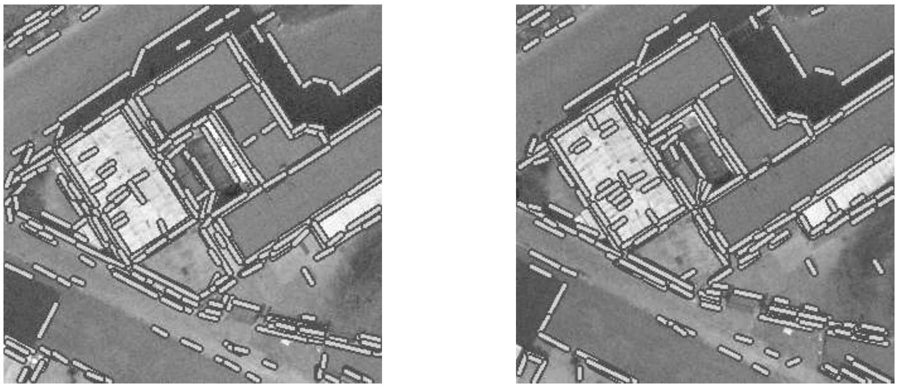
4.3 Wide baseline matching between two views

In the wide baseline case a planar homography is required in order to compute the cross-correlation score on curve neighbourhoods. In other respects the matching algorithm is identical to the small baseline algorithm.

In the case of line matching, as discussed in section 4.1.3, this homography can only be determined up to a one parameter family and a one dimensional search is required. Given a putative line match the aim is to compute a cross-correlation score using $\mathtt{H}(\mu)$ (cf. section 2.2). This means that for lines a single score is obtained by computing the value μ^*, with corresponding homography $\mathtt{H}^*$, for which the cross-correlation is highest over all μ. In the following we first describe the computation of the $\mathtt{H}$-correlation score, and then the estimation of μ^*.

In the case of curve matching a local homography is determined at each point of the curve (in contrast a single homography is determined over the entire extent of a line). However, unlike in the case of a line, no search is required because the homography at each point is determined uniquely from the osculating plane computed from the curvatures, as described in section 2.4.

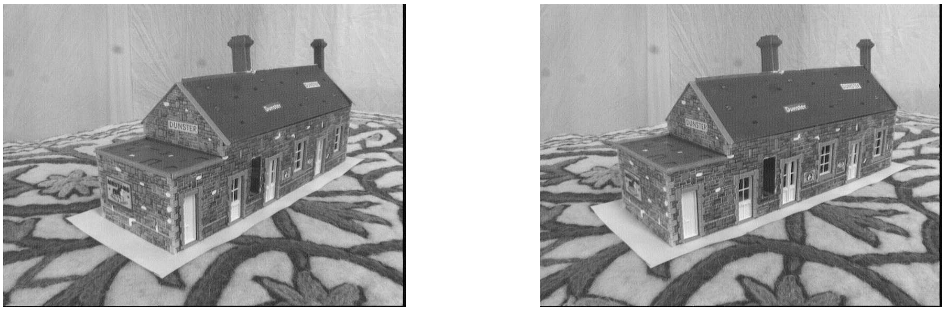
Original image pair



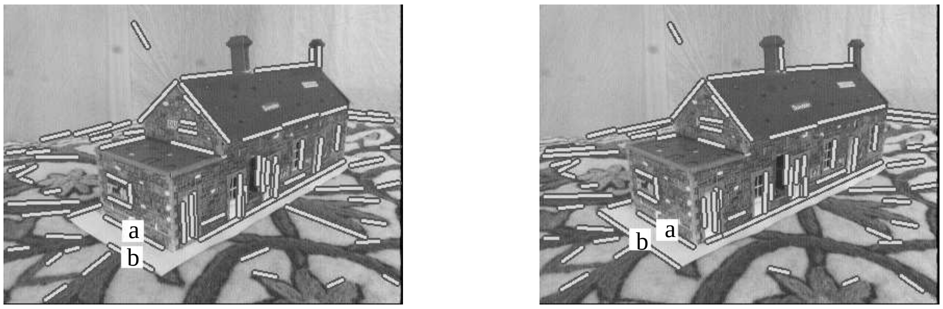
Extracted line segments

Matched line segments

Figure 14: **Short baseline two view line matching I**: The images are 600×600 pixels. There are 248 and 236 line segments detected in the left and right images, respectively, and 122 lines are matched. Of these 122 matches 119 are correct.



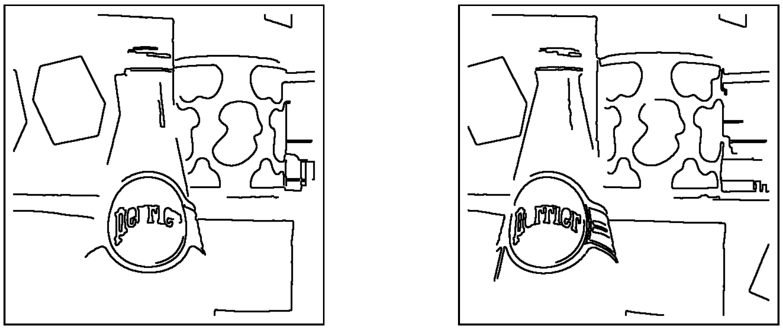
Original image pair



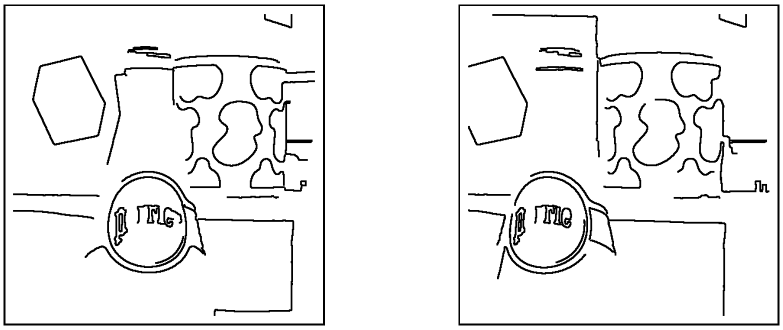
Matched line segments

Figure 15: **Short baseline two view line matching II**: The images are 768 $\times$ 576 pixels. There are 120 and 135 line segments detected in the left and right images, respectively, and 73 lines are matched. Of these 73 matches 69 are correct. Note also that segments b are matched correctly despite their different segmented lengths in the two images.

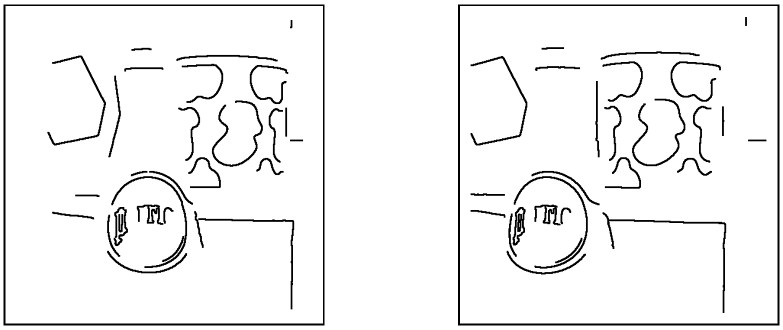
Original images



Extracted curve segments



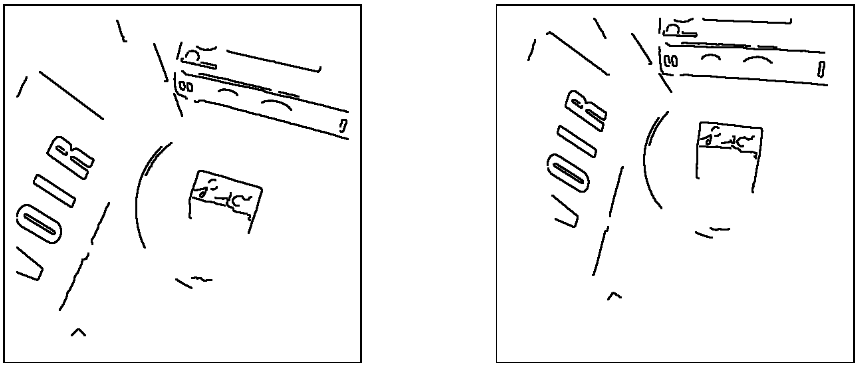
Matched curve segments – complete



Matched curve segments – only corresponding parts

Figure 16: **Short baseline two view curve matching I**: The images are 512×512 pixels. There are 37 and 47 curve segments for the left and right images respectively. Of these 29 are matched, and only one of the matches is incorrect (annoyingly, the left side of the neck of the bottle is incorrectly matched to the right side).

Original images



Matched curve segments

Figure 17: **Short baseline two view curve matching II**: The images are 512×512 pixels. There are 80 and 89 curve segments for the left and right images respectively. Of these 40 are matched with no mis-matches.

Implementation details. We require a homography which maps between the neighbourhoods of the putative line match. A particular homography $\mathtt{H}(\mu)$ may be assessed by its correlation score: the cross-correlation is evaluated for a rectangular strip on one side of the line segment by using $\mathtt{H}(\mu)$ to associate corresponding points. The length of the strip is the common overlap of the lines determined by the epipolar geometry. The area of the strip must be sufficient to include neighbouring texture, otherwise the cross-correlation will not be discriminating. Here a strip of width 14 pixels is used. The computation is carried out to sub-pixel resolution using bilinear interpolation. To be invariant to an affine intensity change over the spatial neighbourhood, normalized cross-correlation is used.

The homography $\mathtt{H}^*$ which maximizes the cross-correlation must then be estimated. This involves estimating μ^*. However μ is a projective parameter and is not directly measurable or meaningful in the image. Instead of μ the homography is parametrized by the mapping of a single point (which in turn determines μ) as described in section 2.2. The corner of the rectangular correlation strip is ideal for this purpose. The set of possible correspondences in the other image $\mathbf{x}'(\mu)$ lie on the epipolar line of $\mathbf{x}$ (cf. figure 2). The value μ^* is obtained by searching the epipolar line for the $\mathbf{x}'(\mu^*)$ which maximizes the cross-correlation. Consider the point $\mathbf{x}'$ which has the same distance to the line $\mathbf{l}'$ as $\mathbf{x}$ has to $\mathbf{l}$. This corresponds to a scale/foreshortening factor of one between the images. To restrict the search we limit the possible scale factors to the range 1/3 to 3. This range defines an interval on the epipolar line in the second image. The correlation score is then evaluated at 10 equi-spaced points on this interval, and the best score determines μ^*.

As straight line segments often occur at the junction of planar facets, the homographies are in general different for the two sides of the line. It is therefore necessary to process the two sides separately, and a correlation strip of width 14 pixels is used on each side. The mean of the two correlation scores is used as the matching score for the line.

In the case of curve matching the curvature is computed at each point of the edgel chain. Curvature is computed from the cubic B-spline fitted to the edgel chain (see section 3 for details). The homography (and thence the correlation score) is computed via the osculating plane. As the homography is based on the osculating plane, left and right sides of the curve are included simultaneously.

If the edgel chain is locally straight, or almost straight, then curvature is close to zero and the osculating plane cannot be determined. In this case the correlation score is determined by the line algorithm above, using the tangent line at that point on the contour chain. A threshold of 0.02 on local curvature is used to decide which case applies. Note, this threshold is determined empirically. It is conservative (i.e. errs towards lines), because if a curve is incorrectly classified as a line then the only loss is that a one-parameter search is carried out that could have been avoided. Conversely, if a line is incorrectly classified as a curve, the resulting homography may be erroneous and this should be avoided.

4.3.1 Wide baseline matching results

Figures 18 – 20 show examples of the matching algorithm for three scenes. In figure 18, 50% of the detected lines are matched, a similar performance to that of the short baseline case. Figure 19 compares line matching with curve matching using the same edgel chains. Line segmentation results in more matches (there are more lines than edgel chains) and more mis-matches. Conversely the direct matching of the chains results in fewer matched edgels, but fewer mis-matches. In figure 20 only the edgel chains are matched as there are few linear features. In all cases (both short and wide baselines) if more lines or curves are detected (by lowering thresholds) there is a higher proportion of mismatches.

The success of the matching algorithm does support the use of homographies as a means of disambiguation.

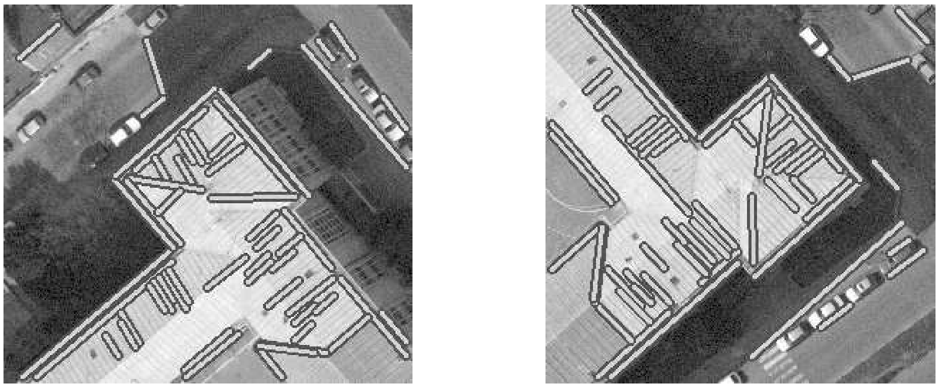


Figure 18: **Wide baseline two view matching I**: The images are 550×550 pixels. There is a significant rotation between the images. 104/96 line segments are detected and matched. Of the 55 line matches shown, 51 are correct.

4.4 Three view matching

In three views there are strong geometric constraints available for matching. For example, given the trifocal tensor and corresponding lines in two images, the corresponding line in the third image is determined. In the case of curves there are three geometric constraints available for matching [45]: the curve point, its tangent line, and its curvature. As shown in section 2.4, tangent lines and curvatures may be transferred to a third view given their correspondence in the first two views. Point transfer is described in [26].

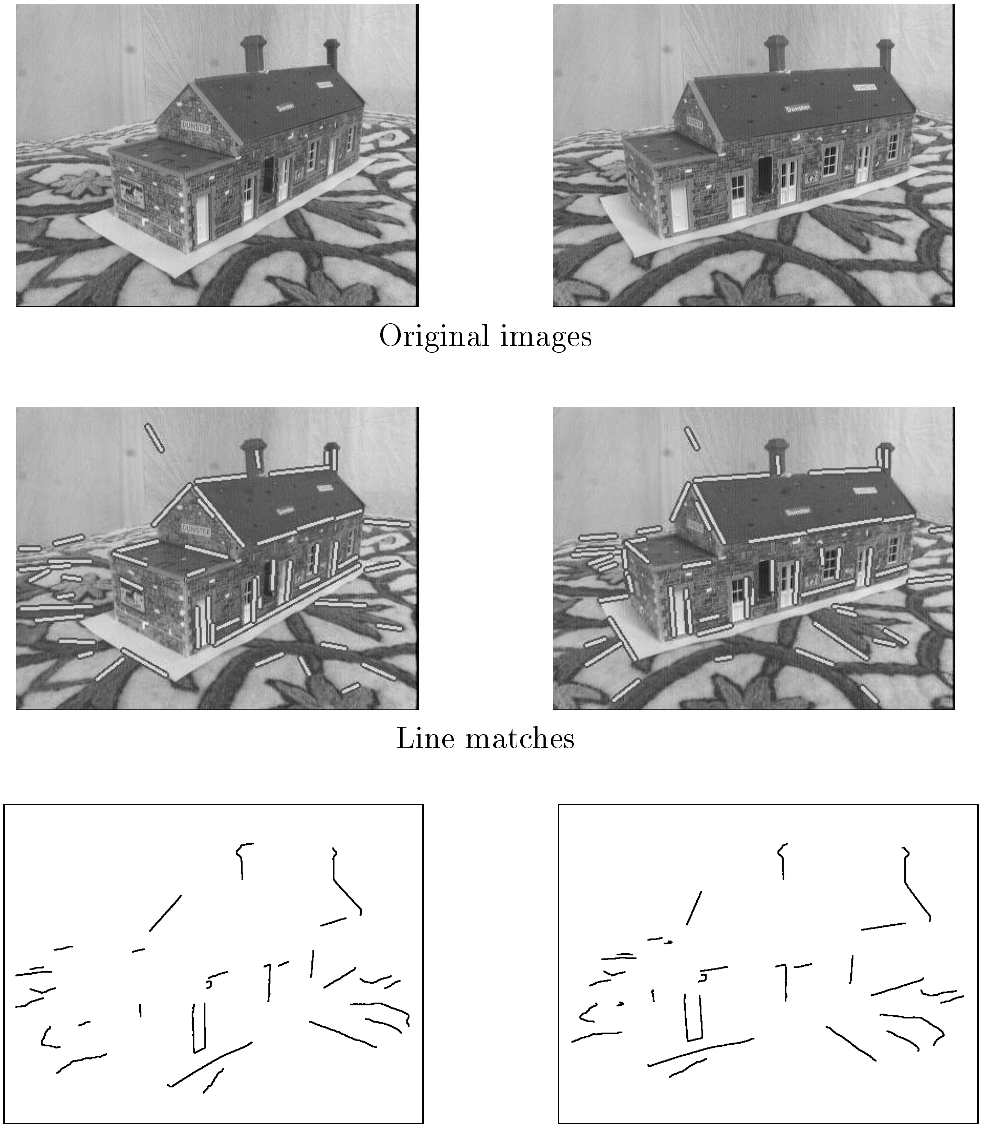
The three view matching algorithm uses the trifocal constraint as well as the two view algorithm for an intensity neighbourhood verification. There are four stages:

1. **Two view matching**: Putative curve matches between image 1 and 2 are determined using steps 1 and 2 of the algorithm described in section 4.2, that is curve pairs with no overlap or a low similarity score are rejected.

2. **Trifocal geometry verification**:

   Line segments. Given a putative correspondence between line segments in the first and second views an (infinite) line is determined in the third view (via the trifocal tensor). Line segments in the third view within a distance threshold of this transferred (infinite) line are a possible match. In the second step the trifocal tensor (via the three pairwise fundamental matrices) determines the common parts of the three segments. If there is no common part then the match is rejected. Otherwise, the line segments in the three views satisfy the geometric constraints and is a possible line triplet match.

   Curves. Given a putative correspondence between curves, a curve is generated in the third view. For each point in the first view, a putative correspondence is determined in the second (via the epipolar geometry) and this two view match determines a point in the third view (via the trifocal tensor). Curves which have a sufficient overlap within a



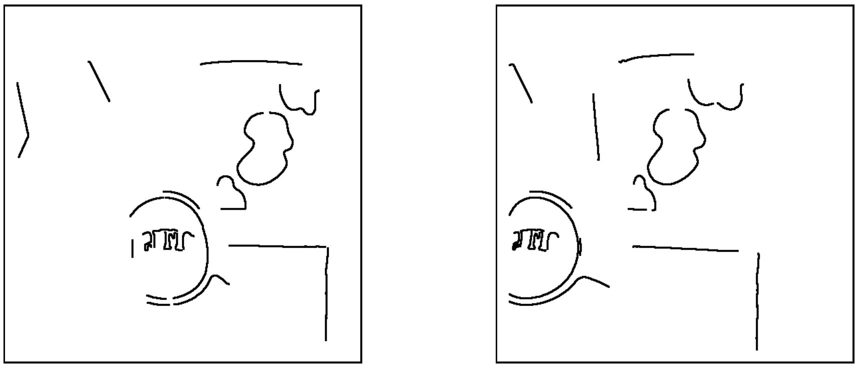
Original images

Line matches

Curve matches

Figure 19: **Wide baseline two view matching II**: There is significantly different foreshortening between the planes of the house in the two images. If only lines are matched, for the 120/125 detected lines, there are 53 matches and of these 41 are correctly matched. If instead edgel chains are matched directly, i.e. without fitting lines, then 93 and 89 curve segments are detected in the left and right images, respectively, and 28 of these are matched. There is one mis-match.

Original images



Curve matches

Figure 20: **Wide baseline two view matching III**: There are 37 and 48 curve segments extracted in the left and right images, respectively. Of these 16 are matched, with 2 of these matches incorrect.

distance threshold of the "generated" curve are kept as potential matches. Notice that no direct use is made of the tangent line and curvature transfer in the verification, though these certainly could also be employed.

3. Photometric verification:

A similarity measure is computed for each triplet match satisfying the three view geometric constraints. The similarity score is computed between views 2/3, as described in section 4.1.2. The similarity measure is the average of this score and the score between views 1/2 computed in step 1.

4. Winner takes all matching:

The matches are determined by a winner takes all scheme over the three views based on the similarity measures. This rules out multiple matches for an individual line or curve.

The complexity for lines and curves are discussed separately. If n is the average number of line segment in each image, the complexity of finding triplets that satisfy the trifocal constraint is $\mathcal{O}(n^2)$. If n lines satisfy the trifocal constraint (in general less), the complexity of the photometric verification is $\mathcal{O}(2nl)$ where l is the average length of a line segment. The factor 2 is due to the fact that we compare views 1/2 as well as views 2/3. The overall complexity is then $\mathcal{O}(2nl + n^2) \approx \mathcal{O}(n^2)$ as l is in general of the same order of magnitude as n.

In the case of curves, the complexity of finding point triplets that satisfy the trifocal constraint is $\mathcal{O}(ln^2)$. As only part of the potential triplets satisfy the trifocal constraint, the complexity of photometric verification is at most of the same order of magnitude. The overall complexity is then $\mathcal{O}(n^3)$ as l is in general of the same order of magnitude as n.

4.4.1 Matching results

Results are given for the three view matching algorithm applied to the same scenes as the two view matching.

Line matching. Figures 21 and 23 display matches for 3 views of an aerial scene. Figure 23 is a detail of figure 21. The line matching performance is excellent: about 35% of the lines are matched, and of these matches 100% are correct.

These three view line matches may be used to determine 3D lines. The 3D line is estimated by minimizing the error in each view between the measured (segmented) line and the 3D line projected into that view. The error is the sum of the perpendicular distances between the ends of the measured line and the projected line. Examples are given in figures 22 and 24. Figure 25 shows a 3D texture mapped model built from the 3D lines. For this example line grouping into planar facets was done manually. However, using the same line matching algorithm, comparable results have been obtained automatically [4] without any manual intervention.

Figures 26 and 27 show two further examples of 3 view line matching. Again 100% of the line matches are correct.

Curve matching. Figure 28 shows the results of three view curve matching for the "bottle" images. Note, that the outline of the bottle is correctly *not* matched over the three views since it is an apparent contour and not rigidly attached to the scene. All the matches are correct. The 3D reconstruction of these matched curves is shown in figure 29. A further example of three view curve matching is given in figure 30. For this example the matches are 100% correct.

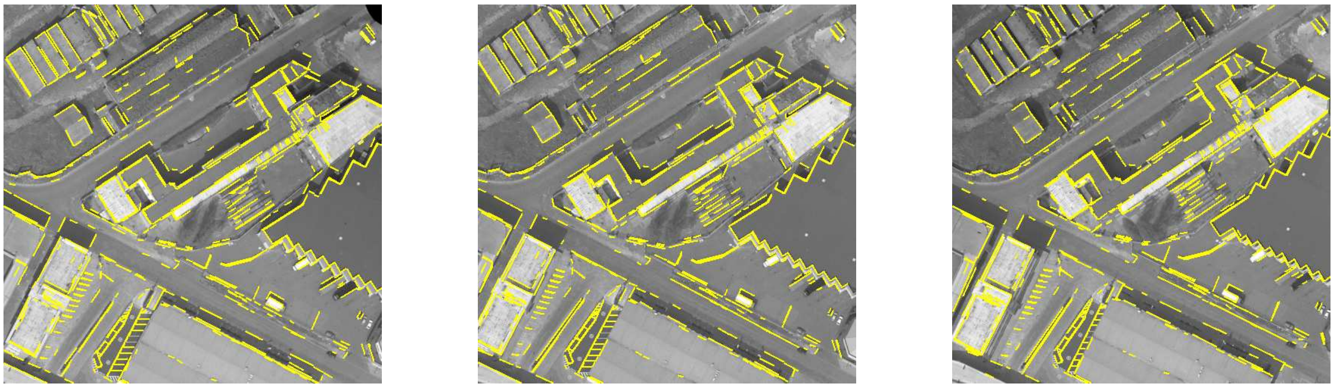


Figure 21: **Three view line matching I**: The aerial images are 2000 × 2000 pixels. There are 2012/2151/1966 line segments detected in the three views. Of these 679 are matched. The correctness of these matches can be verified from figure 22.

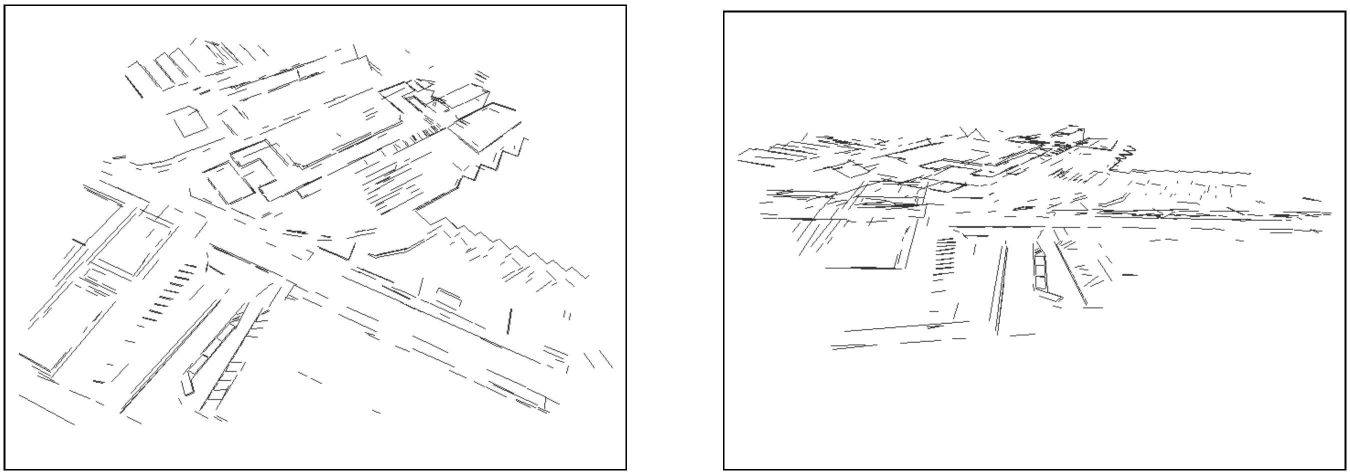


Figure 22: Two views of 3D lines computed from the matches shown in figure 21. The 3D lines are shown after merging and growing, see section 4.5, and are computed by minimizing reprojection error over the three views.

Figure 23: **Three view line matching II**: These images are a sub-part (600 × 600 pixels) of those in figure 21. There are 248/236/212 detected line segments and 89 lines are matched, all of which are correct.

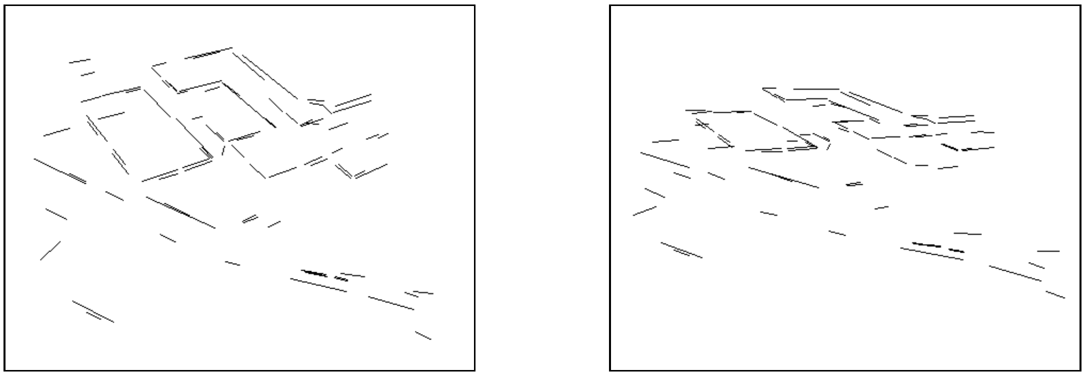


Figure 24: Two views of the 3D lines computed from the line matches of figure 23.

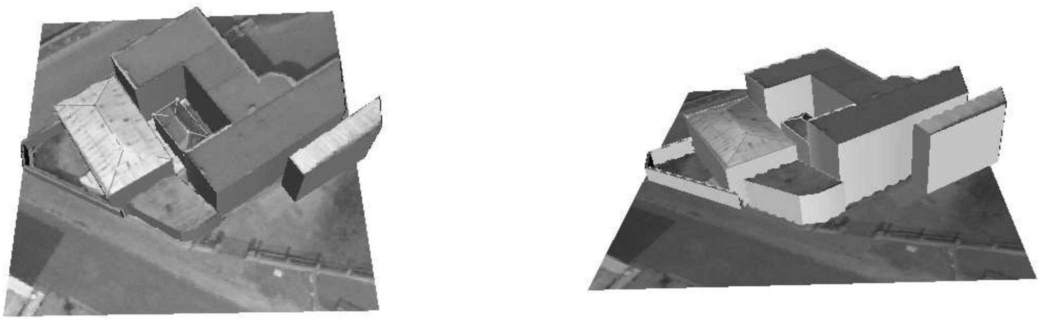


Figure 25: Rendered 3D model built from the 3D lines of figure 24.

Figure 26: **Three view line matching III**: The images are 512×512 pixels. There are 282/285/285 line segments in the first/second/third views. Of these 130 are matched, all the line matches are correct.

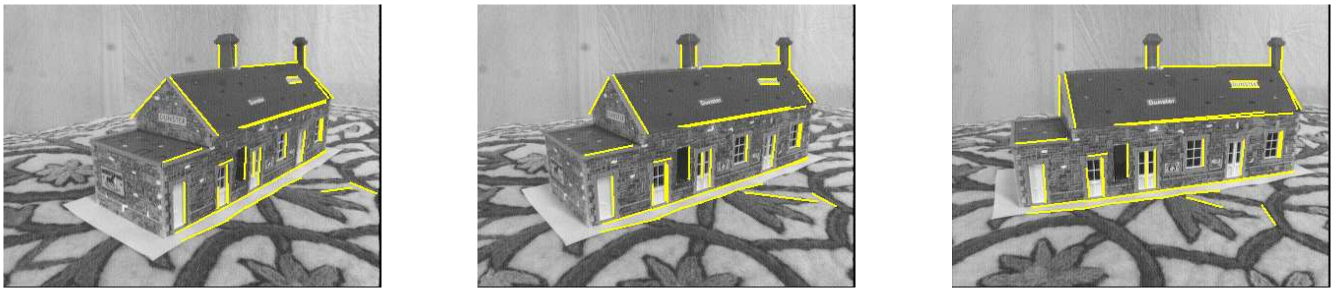
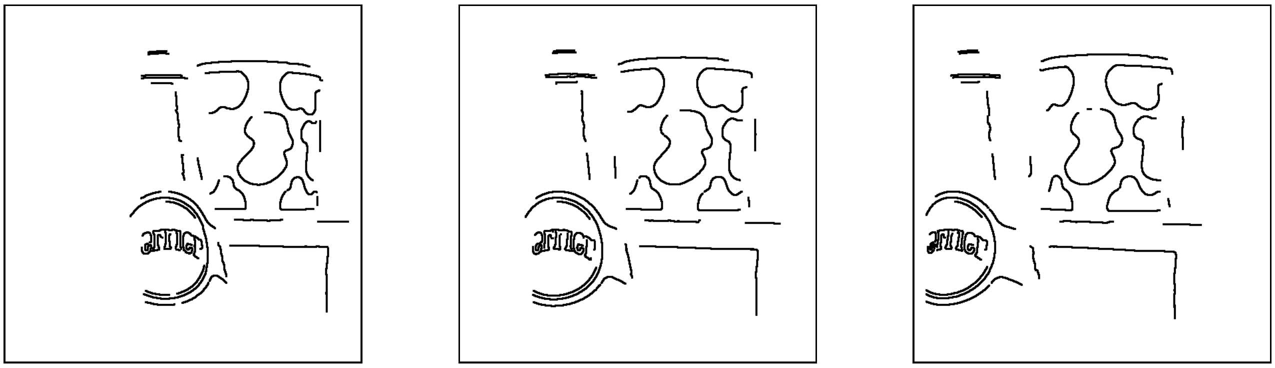


Figure 27: **Three view line matching IV**: There are 120/109/97 line segments in the first/second/third views. Of these 37 are matched, all the line matches are correct.

Original images



Curve matches

Figure 28: **Three view curve matching I**: There are 85/85/90 curve segments detected on the first/second/third views respectively. Of these 25 are matched, and all the matches are correct. Note, that the outline of a specularity is matched because over these three views it behaves as a rigid object. The reflection (of a light source) does not lie on the bottle surface however.

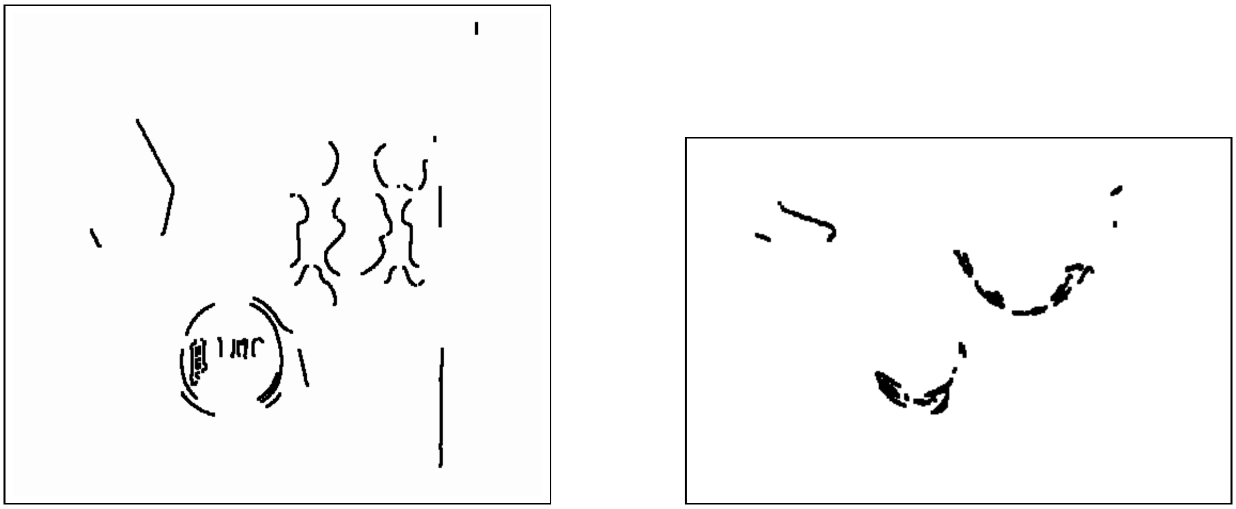
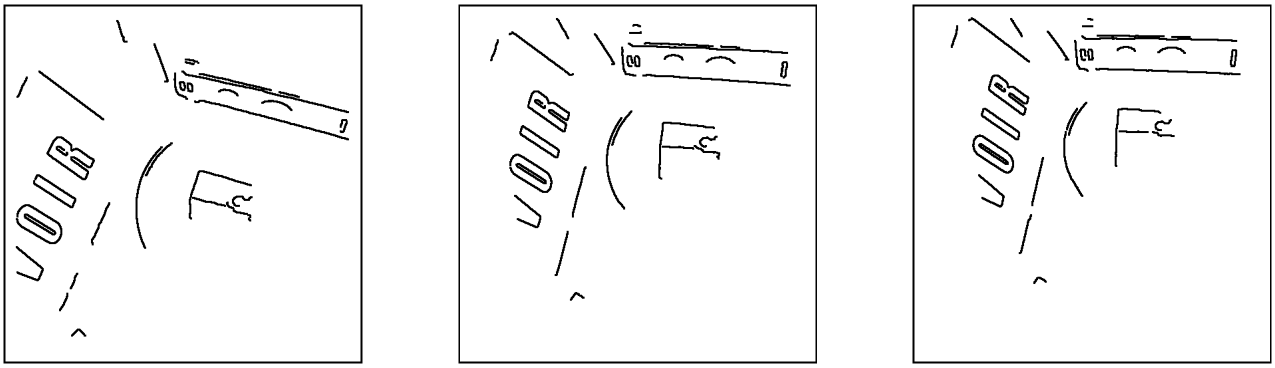


Figure 29: Two views of the 3D curves computed from the matched curves of figure 28. The right figure is a plan view which clearly shows the circular cross-section of the bottle and mug.

Original images



Curve matches

Figure 30: **Three view curve matching II**: There are 80/89/84 curve segments detected on the first/second/third views respectively. All of the 30 matched curves are correct.

4.5 Merging and Growing

In this section we describe two improvements to the three view matching of the previous section. The first improvement – merging – overcomes a limitation of the winner takes all strategy employed in the two and three view algorithms. The second improvement – growing – overcomes the deficiency in line segmentation.

Merging. During the line extraction, lines are often erroneously broken into several segments. Under "winner takes all" only one of these segments may be matched over three views. Suppose, for example, that in view one both lines $L1a$ and $L1b$ match line $L2$ in view two and line $L3$ in view 3. Even if lines $L1a$ and $L1b$ are collinear, only one of the two triplets $(L1a, L2, L3)$ and $(L1b, L2, L3)$ will be selected. If no further action is taken the effect of line fragmentation in detection is that line $L1b$ will be unmatched.

If $L1a$ and $L1b$ are collinear and the two triplets $(L1a, L2, L3)$ and $(L1b, L2, L3)$ are among the triplets before applying the winner takes all strategy, then it is likely that $L1a$ and $L1b$ arose from a single (but fragmented) line. The lines are merged if it is possible to fill in the gap between $L1a$ and $L1b$, that is if there is correlation support for the gap over 3 views. In this manner a line match is generated which is the *union* of the lines detected in each of the views. Furthermore 3D lines estimated from this correspondence are improved because longer line segments are available in each view.

Growing. Another deficiency of line detection, other than fragmentation, is that a fitted line may be shorter than the true line. However, the missing edgels are in general not the same in the different images. Thus if a short line is matched to a longer line, the end points may be extended provided the "grown" edgels satisfy the photometric correlation tests over the three views.

Results. Figure 31 shows the results obtained for merging and growing in the image. Almost 10% additional line matches are generated by merging, and more than 70% of the all matched lines are grown. The significant benefits of this improvement are demonstrated in figure 32, which compares the 3D lines computed from these correspondences before and after the line clean-up.

Extensions. Both merging and growing can also be applied to the two view matching algorithm. However, there is not the independent verification of any change as there is with 3-view geometry. Merging and growing have only be implemented here for lines, but there is no reason why they should not produce equivalent improvements for curves.

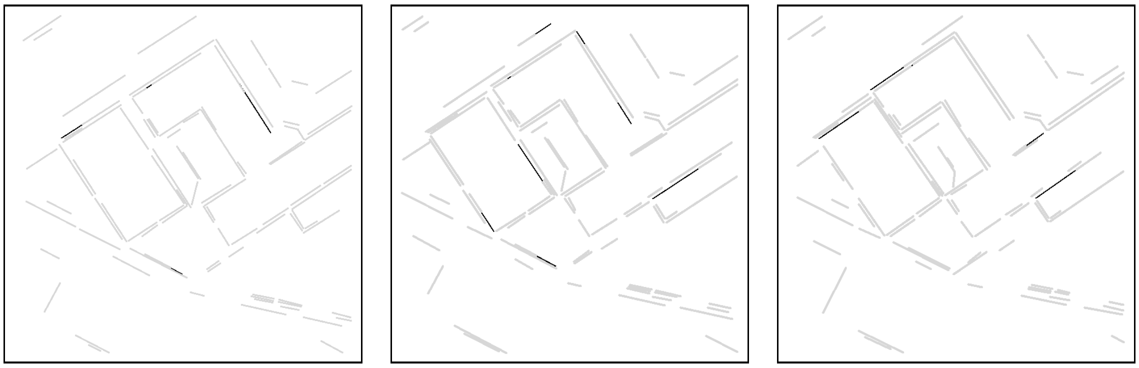
5 Discussion and extensions

We have derived and implemented new geometric results for lines, conics and general curves over multiple views. These results have been applied to automatic line/curve matching over two and three views using a combination of geometric and photometric constraints.

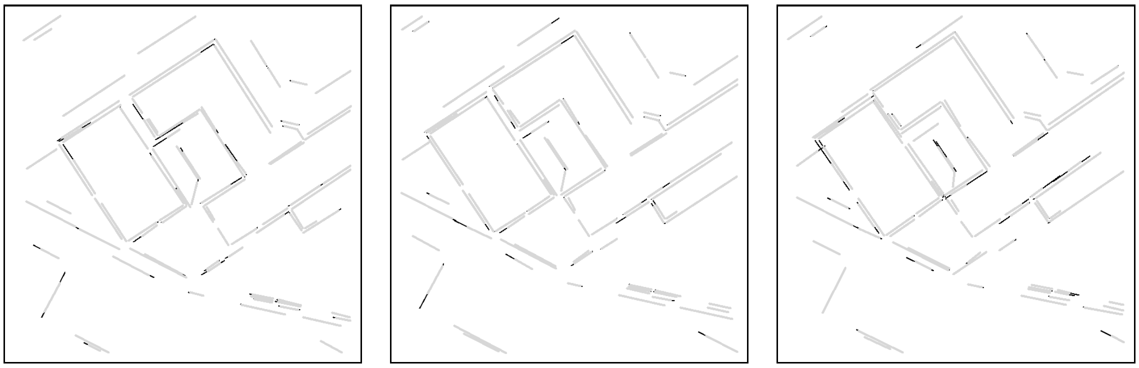
As with any algorithm there remain variations that could be tested. For example: the benefits of using higher order spline fitting so that the curvature is C^1 continuous (rather than C^0 at present); computing the homography for the curve using a nearby point on the curve (rather than the curvature); replacing the winner takes all matching scheme by something less exclusive; using transferred lines and curvature into the third view to verify curve matches.

Original images

Matched line segments

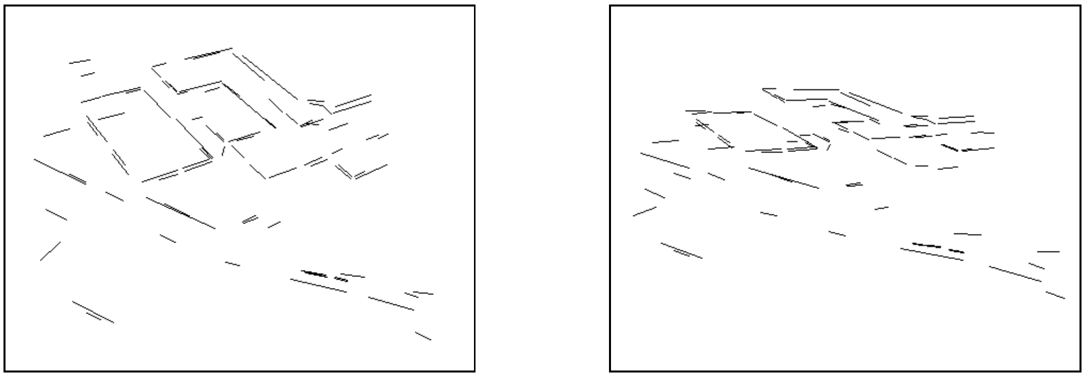


Additional line segments after merging shown in black

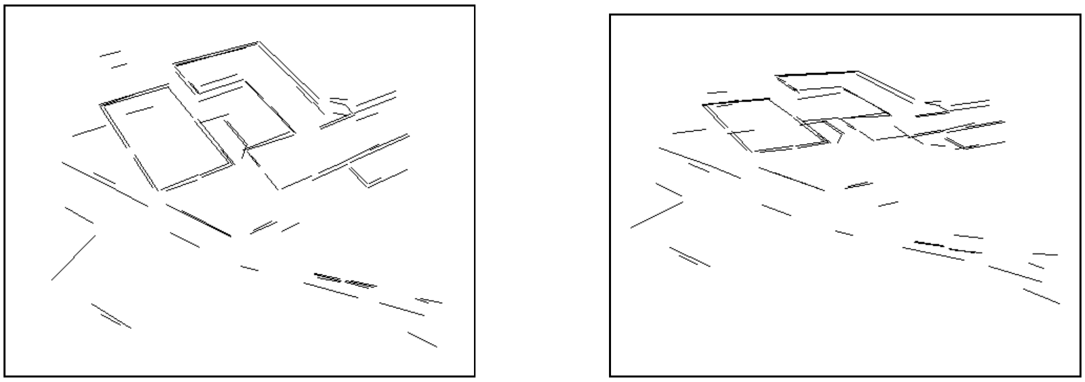


Grown edgels shown in black

Figure 31: Results for merging and growing for the 89 matches of figure 23. 4/8/5 line segments are merged and 626/478/717 edgels are grown in the first/second/third images respectively.



3D lines before merging and growing



3D lines after merging and growing

Figure 32: Two views of the 3D lines computed from the three view correspondences, before and after merging and growing. The improvements are clear — longer segments, shorter gaps, and better parallelism.

Nevertheless, it is worth drawing attention to four points: first, lines and curves can be matched between views without requiring prior grouping if photometric constraints are employed. In three views virtually perfect matches are obtained. In two views matching performance varies according to the quality and number of the extracted curves. It would be worth investigating if the addition of grouping constraints, and the ordering constraint if applicable, could raise the two view matching performance to that of the three view matching. For example, lines are often segmented from edgel chains and the grouping provided by the original chain could be employed.

Second, it has been demonstrated that local planar homographies can ameliorate the effects of wide base line viewing. In the case of lines, and tangent lines at low curvature points, computing the homography involved a one dimensional search. In the case of curves, the homography provided by the osculating plane was used. However, this plane will often *not* coincide with the local tangent plane of the surface (think of the circular curve at the mouth of a mug). It is worth questioning why the algorithm succeeds in this case. A partial answer is that any of the homographies in the family defined by the tangent line correspondence corrects for the effects of image rotation and foreshortening in the line direction, so that only the foreshortening perpendicular to the line direction may be in error. Also, in order to disambiguate putative curve matches, the homography need only provide sufficient correction to correctly rank order the matches.

Third, one of the possibly unexpected benefits of multiple view matching is that deficiencies of segmentation in one view can be corrected if the segmentation fails at a different point in another view. This is well demonstrated by the improvements in 3D lines after merging and growing.

Finally, the wide baseline algorithm may be applied in the short baseline case, but is more expensive. The choice of algorithm should also be automatic, using the cheaper short baseline algorithm where possible. Although we have not investigated the choice, it is likely that the process that generates the fundamental and trifocal tensors will have sufficient information to choose which of the short or long range algorithms is appropriate.

Acknowledgements

We are very grateful to Andrew Fitzgibbon for both discussions and software, and for discussions with Caroline Baillard, Amnon Shashua, Bill Triggs, and Thierry Viéville. Financial support for this work was provided by the EU Esprit Project IMPACT and the UK EPSRC IUE Implementation Project GR/L05969.

A Proofs

A.1 The plane $\boldsymbol{\pi}$ given a homography $\mathtt{H}$ and camera matrices

Given camera matrices $\mathtt{P} = [\mathtt{I} \mid \mathbf{0}]$, $\mathtt{P}' = [\mathtt{A} \mid \mathbf{e}']$, and a homography $\mathtt{H}$, the problem is to determine the coordinates of the plane $\boldsymbol{\pi}$ which induced $\mathtt{H}$. Note $\mathtt{A}$ and $\mathbf{e}'$ are known up to a common scale, and $\mathtt{H}$ only up to scale. From (1) the homography has the form

$$\mathtt{H} = \lambda(\mathtt{A} + \mathbf{e}'\mathbf{v}^\top) \tag{8}$$

for the plane $\boldsymbol{\pi} = (-\mathbf{v}^\top, 1)^\top$, where the unknown scale parameter λ has been included explicitly.

The solution is in two stages: first the scale parameter is determined by pre-multiplying (8) by $[\mathbf{e}']_\times$, so that

$$[\mathbf{e}']_\times \mathtt{H} = \lambda[\mathbf{e}']_\times \mathtt{A}$$

and λ is simply the ratio of the matrix elements on either side of the equation. If the matrix $[\mathbf{e}']_\times \mathtt{H}$ is written as the 9-vector $\tilde{\mathbf{h}}$, and $[\mathbf{e}']_\times \mathtt{A}$ as the 9-vector $\tilde{\mathbf{r}}$ then $\tilde{\mathbf{h}} = \lambda\tilde{\mathbf{r}}$, and $\lambda = \tilde{\mathbf{r}}.\tilde{\mathbf{h}}/(\tilde{\mathbf{r}}.\tilde{\mathbf{r}})$.

The second stage is then to determine $\mathbf{v}$ given λ. Pre-multiplying (8) by $\mathbf{e}'$, gives

$$\mathbf{v} = \left(\mathtt{H}/\lambda - \mathtt{A}\right)^\top \mathbf{e}'/||\mathbf{e}'||^2$$

and thus the plane $\boldsymbol{\pi}$ has coordinates

$$\boldsymbol{\pi} = \begin{pmatrix} -\mathbf{v} \\ 1 \end{pmatrix} = \frac{1}{||\mathbf{e}'||^2} \begin{pmatrix} -\left(\mathtt{H}/\lambda - \mathtt{A}\right)^\top \mathbf{e}' \\ ||\mathbf{e}'||^2 \end{pmatrix}$$

A.2 Lines

In this section the proof is given of Proposition 1. Luong and Viéville [34] show that the homography between two images induced by a world plane $\boldsymbol{\pi}$ is given by

$$\mathtt{H}(\mathbf{v}) = -[\mathbf{e}']_\times \mathtt{F}/||\mathbf{e}'||^2 + \mathbf{e}'\mathbf{v}^\top$$

and this may be generalized to

$$\mathtt{H}(\mathbf{v}) = -[\mathbf{b}]_\times \mathtt{F}/(\mathbf{b}.\mathbf{e}') + \mathbf{e}'\mathbf{v}^\top \tag{9}$$

for any 3-vector $\mathbf{b}$, provided that $\mathbf{b}.\mathbf{e}' \neq 0$. The 3-vector $\mathbf{v}$ parametrizes the 3-parameter family of planes in 3-space. Under the homography the images of points on $\boldsymbol{\pi}$ are mapped as $\mathbf{x}' = \mathtt{H}\mathbf{x}$.

Since the case that $\mathbf{l}'.\mathbf{e}' = 0$ is excluded in the proposition, we may choose $\mathbf{b} = \mathbf{l}'$. Then

$$\mathtt{H}(\mathbf{v}) = -[\mathbf{l}']_\times \mathtt{F}/(\mathbf{l}'.\mathbf{e}') + \mathbf{e}'\mathbf{v}^\top \tag{10}$$

and when $\mathbf{v} = \mathbf{0}$ the homography is degenerate and corresponds to a plane through the second camera centre which intersects the second image plane in the line $\mathbf{l}'$.

Given the correspondence of image lines, $\mathbf{l}$, $\mathbf{l}'$, the homographies induced by planes containing the line in 3-space are reduced from a 3-parameter to a one-parameter family. From (3) a line transforms under a homography as $\mathbf{l} = \mathtt{H}^\top\mathbf{l}'$, and imposing this relation on (10) we obtain

$$\begin{aligned} \mathbf{l} &= \mathtt{H}^\top\mathbf{l}' = \left(\mathtt{F}^\top[\mathbf{l}']_\times/(\mathbf{l}'.\mathbf{e}') + \mathbf{v}\mathbf{e}'^\top\right)\mathbf{l}' \\ &= \mathbf{v}(\mathbf{e}'.\mathbf{l}') \end{aligned}$$

Including the scale factor explicitly as $-\mu\mathbf{l} = \mathbf{v}(\mathbf{e}'.\mathbf{l}')$, (10) becomes

$$\tilde{\mathtt{H}}(\mu) = -\left([\mathbf{l}']_\times \mathtt{F} + \mu\mathbf{e}'\mathbf{l}^\top\right)/(\mathbf{l}'.\mathbf{e}')$$

The factor of $-1/(\mathbf{l}'.\mathbf{e}')$ may be dropped for the homogeneous matrix to give $\mathtt{H}(\mu) = [\mathbf{l}']_\times \mathtt{F} + \mu\mathbf{e}'\mathbf{l}^\top$.

The plane of the homography in the projective reconstruction with cameras $\mathtt{P} = [\mathtt{I}|\mathbf{0}]$, $\mathtt{P}' = [\mathtt{A}|\mathbf{e}']$ may be determined from (2) as

$$\mathbf{v} = \left(\mathtt{H}/\lambda - \mathtt{A}\right)^\top \mathbf{e}'/||\mathbf{e}'||^2$$

A short calculation shows that if $\mathtt{F} = [\mathbf{e}']_\times \mathtt{A}$ (where the equality is exact to fix the scales) then the scale factor λ which satisfies $\lambda[\mathbf{e}']_\times \mathtt{A} = [\mathbf{e}']_\times \mathtt{H}$ has the value $-(\mathbf{l}'.\mathbf{e}')$, since $[\mathbf{e}']_\times [\mathbf{l}']_\times \mathtt{F} = -(\mathbf{l}'.\mathbf{e}')\mathtt{F}$. Thus

$$\mathbf{v} = -\left(\left([\mathbf{l}']_\times \mathtt{F} + \mu\mathbf{e}'\mathbf{l}^\top\right)/(\mathbf{l}'.\mathbf{e}') + \mathtt{A}\right)^\top \mathbf{e}'/||\mathbf{e}'||^2$$

This completes the proof of Proposition 1.

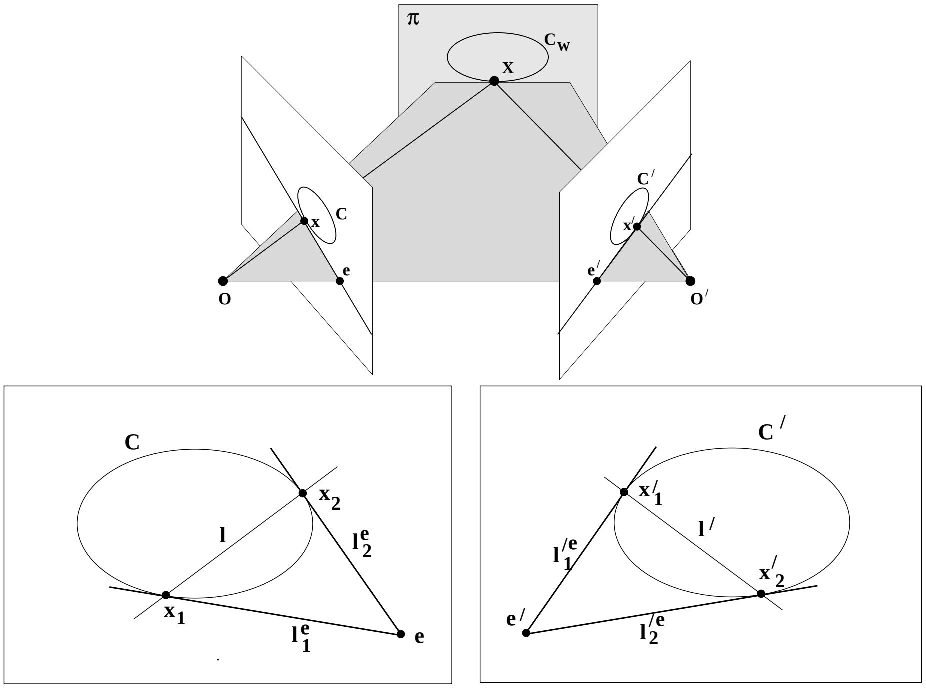


Figure 33: Epipolar tangency for a conic. Upper: The epipolar plane is tangent to the world conic $\mathtt{C}_W$ on the plane $\boldsymbol{\pi}$. In the image the corresponding epipolar lines are tangent to the imaged conics, $\mathtt{C}$ and $\mathtt{C}'$, and the points of tangency $\mathbf{x}$ and $\mathbf{x}'$ correspond. Lower: The two epipolar planes tangent to the world conic $\mathtt{C}_W$ define corresponding epipolar lines, $\mathbf{l}_1^e \leftrightarrow \mathbf{l}_1'^e$ and $\mathbf{l}_2^e \leftrightarrow \mathbf{l}_2'^e$, which are tangent to the imaged conics, and corresponding epipolar tangent points, $\mathbf{x}_1 \leftrightarrow \mathbf{x}_1'$, $\mathbf{x}_2 \leftrightarrow \mathbf{x}_2'$. The lines $\mathbf{l}$ and $\mathbf{l}'$ (which are the polars of the epipole in each image) also correspond. These lines are the image of a line on the plane $\boldsymbol{\pi}$ which is the polar with respect to the world conic of the point at which the baseline pierces $\boldsymbol{\pi}$.

A.3 Conics

In this section the proof is given of Proposition 2. The proof is in two parts. First, we build on the result of the previous section and determine a pencil of planes from corresponding lines obtained from the image conics. Then the pencil is reduced to a two fold ambiguity.

Corresponding image conics must satisfy an epipolar tangency condition in order to be consistent with the epipolar geometry. Suppose an epipolar line $\mathbf{l}^e$ is tangent to the conic $\mathtt{C}$ in the first image, then the corresponding epipolar line $\mathbf{l}'^e$ is tangent to the conic $\mathtt{C}'$ in the second image. The geometry is illustrated in figure 33. This is a necessary condition for conics to correspond, but is not sufficient to ensure that the conics do correspond. To see that epipolar tangency is not sufficient, consider the constraints on a conic in the second image: the conic in the first image has two epipolar tangents and this generates two epipolar tangency constraints in the second image; however a conic is defined by tangency to five lines, so there is a 3-parameter family of conics which satisfy the two epipolar tangency constraints in the second image.

The line $\mathbf{l}$ of figure 33 intersects the conic $\mathtt{C}$ at the points $\mathbf{x}_1, \mathbf{x}_2$ of tangency of lines from $\mathbf{e}$. A line constructed in this way from a point and a conic is known as the *polar* line, and is given by $\mathbf{l} = \mathtt{C}\mathbf{e}$ [52]. Similarly, $\mathbf{l}' = \mathtt{C}'\mathbf{e}'$. It is evident from figure 33 that the lines $\mathbf{l}$ and $\mathbf{l}'$ correspond. Consequently, the plane $\boldsymbol{\pi}$ and associated homography $\mathtt{H}$ of the world conic are determined up to a one-parameter family by Proposition 1 as:

$$\begin{aligned} \mathtt{H}(\mu) &= [\mathbf{l}']_\times \mathtt{F} + \mu \mathbf{e}' \mathbf{l}^\top \\ &= [\mathtt{C}'\mathbf{e}']_\times \mathtt{F} + \mu \mathbf{e}' (\mathtt{C}\mathbf{e})^\top \end{aligned} \tag{11}$$

From (4) a conic transforms under a homography as $\mathtt{C} = \mathtt{H}^\top \mathtt{C}' \mathtt{H}$, and imposing this relation on (11) we obtain

$$\begin{aligned} \mathtt{C} &= \mathtt{H}^\top \mathtt{C}' \mathtt{H} = \left(-\mathtt{F}^\top [\mathtt{C}'\mathbf{e}']_\times + \mu (\mathtt{C}\mathbf{e}) \mathbf{e}'^\top\right) \mathtt{C}' \left([\mathtt{C}'\mathbf{e}']_\times \mathtt{F} + \mu \mathbf{e}' (\mathtt{C}\mathbf{e})^\top\right) \\ &= -\mathtt{F}^\top [\mathtt{C}'\mathbf{e}']_\times \mathtt{C}' [\mathtt{C}'\mathbf{e}']_\times \mathtt{F} + \mu (\mathtt{C}\mathbf{e}) \mathbf{e}'^\top \mathtt{C}' [\mathtt{C}'\mathbf{e}']_\times \mathtt{F} - \mu \mathtt{F}^\top [\mathtt{C}'\mathbf{e}']_\times \mathtt{C}' \mathbf{e}' (\mathtt{C}\mathbf{e})^\top + \mu^2 (\mathtt{C}\mathbf{e}) \mathbf{e}'^\top \mathtt{C}' \mathbf{e}' (\mathtt{C}\mathbf{e})^\top \\ &= -\mathtt{F}^\top [\mathtt{C}'\mathbf{e}']_\times \mathtt{C}' [\mathtt{C}'\mathbf{e}']_\times \mathtt{F} + \mu^2 (\mathbf{e}'^\top \mathtt{C}' \mathbf{e}') (\mathtt{C}\mathbf{e}) (\mathtt{C}\mathbf{e})^\top \end{aligned} \tag{12}$$

The terms linear in μ are both zero because $[\mathtt{C}'\mathbf{e}']_\times (\mathtt{C}'\mathbf{e}') = \mathbf{0}$.

Equation (12) is homogeneous, and thus contains a hidden scale factor γ which is now made explicit

$$\gamma \mathtt{C} = -\mathtt{F}^\top [\mathtt{C}'\mathbf{e}']_\times \mathtt{C}' [\mathtt{C}'\mathbf{e}']_\times \mathtt{F} + \mu^2 (\mathbf{e}'^\top \mathtt{C}' \mathbf{e}') (\mathtt{C}\mathbf{e}) (\mathtt{C}\mathbf{e})^\top$$

The scale factor may be determined by pre- and post- multiplying by $\mathbf{e}$. Since $\mathtt{F}\mathbf{e} = \mathbf{0}$ we obtain

$$\gamma (\mathbf{e}^\top \mathtt{C}\mathbf{e}) = \mu^2 (\mathbf{e}'^\top \mathtt{C}' \mathbf{e}') (\mathbf{e}^\top (\mathtt{C}\mathbf{e}))((\mathtt{C}\mathbf{e})^\top \mathbf{e}) = \mu^2 (\mathbf{e}'^\top \mathtt{C}' \mathbf{e}') (\mathbf{e}^\top \mathtt{C}\mathbf{e})^2$$

and hence $\gamma = \mu^2 (\mathbf{e}'^\top \mathtt{C}' \mathbf{e}')(\mathbf{e}^\top \mathtt{C}\mathbf{e})$. Note, it is permissible to divide by $\mathbf{e}^\top \mathtt{C}\mathbf{e}$ since Proposition 2 excludes the case that $\mathbf{e}^\top \mathtt{C}\mathbf{e} = 0$. With γ determined (12) becomes

$$\mu^2 \left[(\mathtt{C}\mathbf{e})(\mathtt{C}\mathbf{e})^\top - (\mathbf{e}^\top \mathtt{C}\mathbf{e}) \mathtt{C}\right] (\mathbf{e}'^\top \mathtt{C}' \mathbf{e}') = \mathtt{F}^\top [\mathtt{C}'\mathbf{e}']_\times \mathtt{C}' [\mathtt{C}'\mathbf{e}']_\times \mathtt{F}$$

The coordinates of a plane in a projective reconstruction are determined as in Proposition 1 but with $\mathbf{l} = \mathtt{C}\mathbf{e}$ and $\mathbf{l}' = \mathtt{C}'\mathbf{e}'$, and this completes the proof of Proposition 2.

It is straightforward to verify that the homography of Proposition 2 does indeed satisfy the epipolar tangency constraint. The two epipolar lines tangent to $\mathtt{C}$ from $\mathbf{e}$ may be represented as a degenerate line (dual) conic as $(\mathtt{C}\mathbf{e})(\mathtt{C}\mathbf{e})^\top - (\mathbf{e}^\top \mathtt{C}\mathbf{e})\mathtt{C}$. Since this is a dual conic, it maps as $\mathtt{C}^{*'} = \mathtt{H}\mathtt{C}^* \mathtt{H}^\top$ under the point transformation $\mathbf{x}' = \mathtt{H}\mathbf{x}$. A short calculation shows that the mapped conic is $(\mathtt{C}'\mathbf{e}')(\mathtt{C}'\mathbf{e}')^\top - (\mathbf{e}'^\top \mathtt{C}'\mathbf{e}')\mathtt{C}'$, which is the degenerate line conic representing the epipolar lines tangent to $\mathtt{C}'$ from $\mathbf{e}'$ as required.

A.4 Curves

In this section the proof is given of Proposition 3 and Proposition 4.

Given corresponding points on two curves $\mathbf{c}(t)$ and $\mathbf{c}'(u)$, which are locally related by a homography $\mathtt{H}$, we will derive the relation between the curvatures κ and κ' at these points. Note that in general the parametrization of the two curves is different, and is not arc length. In order to avoid having to keep track of parametrizations, we will represent the curvature of $\mathbf{c}'(u)$ by an osculating circle at $\mathbf{x}'$, map this circle under the homography $\mathtt{H}$, and then determine the curvature of the resulting conic at $\mathbf{x}$. The disadvantage of this approach is that only the magnitude of the curvature is modelled, not the sign. The reason for going from $\mathbf{x}'$ to $\mathbf{x}$, rather than $\mathbf{x}$ to $\mathbf{x}'$ is because the conic transformation is a simpler expression in this direction.

We choose a local coordinate system such that the point $\mathbf{x}$ is at the origin, and the x-axis is the tangent line to the curve at $\mathbf{x}$, and denote this system by $\tilde{\mathbf{x}}$. The coordinates $\tilde{\mathbf{x}}$ and $\mathbf{x}$ are related by a Euclidean transformation. Similarly $\tilde{\mathbf{x}}'$ is the corresponding local coordinate system for $\mathbf{x}'$. It may be verified that the homography $\tilde{\mathbf{x}}' = \tilde{\mathtt{H}}\tilde{\mathbf{x}}$ that maps the origin of $\tilde{\mathbf{x}}$ to the origin of $\tilde{\mathbf{x}}'$, and the x-axis of $\tilde{\mathbf{x}}$ to the x-axis of $\tilde{\mathbf{x}}'$ has the form:

$$\tilde{\mathtt{H}} = \begin{bmatrix} \alpha & \beta & 0 \\ 0 & \gamma & 0 \\ a & b & c \end{bmatrix}$$

The osculating circle at $\tilde{\mathbf{x}}'$ has radius $\rho' = 1/\kappa'$ with centre $(0, \rho')$, and matrix representation

$$\tilde{\mathbf{C}}' = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & -\rho' \\ 0 & -\rho' & 0 \end{bmatrix}$$

A conic transforms as $\tilde{\mathbf{C}} = \tilde{\mathtt{H}}^\top \tilde{\mathbf{C}}' \tilde{\mathtt{H}}$, and so the circle $\tilde{\mathbf{C}}'$is mapped to the conic $\tilde{\mathbf{C}}$ where

$$\tilde{\mathbf{C}} = \tilde{\mathtt{H}}^\top \tilde{\mathbf{C}}' \tilde{\mathtt{H}} = \begin{bmatrix} \alpha & 0 & a \\ \beta & \gamma & b \\ 0 & 0 & c \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & -\rho' \\ 0 & -\rho' & 0 \end{bmatrix} \begin{bmatrix} \alpha & \beta & 0 \\ 0 & \gamma & 0 \\ a & b & c \end{bmatrix} = \begin{bmatrix} \alpha^2 & \alpha\beta - a\rho'\gamma & 0 \\ \alpha\beta - a\rho'\gamma & \beta^2 + \gamma^2 - 2b\rho\gamma & -c\rho'\gamma \\ 0 & -c\rho'\gamma & 0 \end{bmatrix}$$

We now need to determine the curvature of the conic $\tilde{\mathbf{C}}$ at $\tilde{\mathbf{x}}$. The conic $\tilde{\mathbf{x}}^\top \tilde{\mathbf{C}} \tilde{\mathbf{x}} = 0$ with $\tilde{\mathbf{x}} = (x, y, 1)^\top$ is an implicit function of x and y. Write this implicit function as $\phi(x, y) = 0$, then the curvature is given by

$$\kappa = \frac{\phi_y^2\, \phi_{xx} - 2\phi_x\, \phi_y\, \phi_{xy} + \phi_x^2\, \phi_{yy}}{(\phi_x^2 + \phi_y^2)^{\frac{3}{2}}}$$

We write this in terms of the conic matrix $\tilde{\mathbf{C}}'$ and tangent line $\tilde{\mathbf{l}}$ at $\tilde{\mathbf{x}}$, which (with exact equality, not just up to scale) is

$$\tilde{\mathbf{l}} = (\tilde{l}_1, \tilde{l}_2, \tilde{l}_3)^\top = \tilde{\mathbf{C}}(x, y, 1)^\top = (0, -c\rho'\gamma, 0)^\top$$

then $\phi_x = 2\tilde{l}_1, \phi_y = 2\tilde{l}_2$ and we obtain

$$\kappa = \frac{\tilde{l}_2^2\, \tilde{\mathbf{C}}_{11} - 2\tilde{l}_1\, \tilde{l}_2\, \tilde{\mathbf{C}}_{12} + \tilde{l}_1^2\, \tilde{\mathbf{C}}_{22}}{(\tilde{l}_1^2 + \tilde{l}_2^2)^{\frac{3}{2}}} = \frac{\alpha^2}{c\rho'\gamma}$$

and hence

$$\kappa' = \frac{c\gamma}{\alpha^2}\, \kappa$$

Note, this shows that mapped curvature only depends on the local linearization of the homography (an affine transformation), since it does not depend on the map of the line at infinity $(a, b, c)^\top$. It also does not depend on the affine shear β.

It only remains to relate the elements α, γ, c of $\tilde{\mathtt{H}}$ to those of $\mathtt{H}, \mathbf{x}, \mathbf{x}', \mathbf{l}$ and $\mathbf{l}'$. First since $\tilde{\mathtt{H}}$ and $\mathtt{H}$ are related by Euclidean transformations it follows that

$$\det(\mathtt{H}) = \det(\tilde{\mathtt{H}}) = \alpha\gamma c.$$

Next, from the map of the origin points $(0, 0, 1)^\top \rightarrow (0, 0, 1)^\top$ under $\tilde{\mathbf{x}}' = \tilde{\mathtt{H}}\tilde{\mathbf{x}}$ it follows that $\tilde{x}_3' = c\tilde{x}_3$. Now $\mathbf{x}'$ and $\tilde{\mathbf{x}}'$ are related by a Euclidean transformation so that $x_3' = \tilde{x}_3'$, and similarly $x_3 = \tilde{x}_3$. Thus we have

$$c = \frac{\tilde{x}_3'}{\tilde{x}_3} = \frac{x_3'}{x_3}$$

In a similar manner, from the map of the x-axes under $\tilde{\mathbf{l}} = \tilde{\mathtt{H}}^\top \tilde{\mathbf{l}}'$ it can be shown that

$$\gamma = \frac{(l_1^2 + l_2^2)^{\frac{1}{2}}}{(l_1'^2 + l_2'^2)^{\frac{1}{2}}}$$

Putting these together gives

$$\kappa' = \frac{c\gamma}{\alpha^2}\kappa = \frac{1}{|\mathtt{H}|^2}\left(\frac{x_3'}{x_3}\right)^3 \frac{(l_1^2 + l_2^2)^{\frac{3}{2}}}{(l_1'^2 + l_2'^2)^{\frac{3}{2}}}\kappa \tag{13}$$

which is the relation between the curvature magnitudes. Inserting sign($|\mathtt{H}|$) to take account of the curvature sign completes the derivation of Proposition 3.

Equation (13) simplifies if the relation between corresponding curves is affine, for then $c = 1$ and $x_3' = x_3$. Suppose the affine homography matrix $\mathtt{H}_\mathrm{A}$ is written as

$$\mathtt{H}_\mathrm{A} = \begin{bmatrix} \mathtt{M} & \mathbf{m} \\ \mathbf{0}^\top & 1 \end{bmatrix}$$

and that $\mathbf{t} = (-l_2, l_1)^\top / (l_1^2 + l_2^2)^{\frac{1}{2}}$ is the unit tangent vector to the curve at $\mathbf{x}$ then

$$\kappa' = \frac{|\mathtt{M}|}{||\mathtt{M}\mathbf{t}||^3}\kappa$$

We now derive Proposition 4. The starting point is $\mathtt{H}(\mu) = \mathtt{A} + \mu\mathbf{e}'\mathbf{l}^\top$, where $\mathtt{A} = [\mathbf{l}']_\times \mathtt{F}$ and $\mathbf{l}, \mathbf{l}'$ are the corresponding tangent lines to the curve. The objective is to solve for μ from the corresponding curvatures κ, κ' using (7).

We require expressions for $\mathbf{x}'$, $\mathbf{l}'$ and $|\mathtt{H}|$. Since $\mathbf{l}.\mathbf{x} = 0$, we obtain for $\mathbf{x}'$:

$$\mathbf{x}' = \mathtt{H}(\mu)\mathbf{x} = (\mathtt{A} + \mu\mathbf{e}'\mathbf{l}^\top)\mathbf{x} = \mathtt{A}\mathbf{x}$$

For $\mathbf{l}'$ we obtain :

$$\mathbf{l}' = \mathtt{H}^{-\top}\mathbf{l} = (\mathtt{A}^\top + \mu\mathbf{l}\mathbf{e}'^\top)^{-1}\mathbf{l} = \frac{1}{|\mathtt{A}^\top + \mu\mathbf{l}\mathbf{e}'^\top|}\mathrm{adj}\,(\mathtt{A}^\top + \mu\mathbf{l}\mathbf{e}'^\top)\mathbf{l} = \frac{1}{|\mathtt{A}^\top + \mu\mathbf{l}\mathbf{e}'^\top|}\mathrm{adj}\,(\mathtt{A}^\top)\mathbf{l}$$

with adj ($\mathtt{M}$) the adjoint of $\mathtt{M}$ defined as adj $(\mathtt{M})\mathtt{M} = |\mathtt{M}|\mathtt{I}$. Proposition 4 gives the explicit form for the adjoint.

The determinant $|\mathtt{H}| = |\mathtt{A} + \mu\mathbf{e}'\mathbf{l}^\top|$ can be simplified as follows. Since $\mathtt{A} = [\mathbf{l}']_\times\mathtt{F}$ is rank 2, it may be written as $\mathtt{A} = \mathbf{a}\mathbf{b}^\top + \mathbf{c}\mathbf{d}^\top$. Then

$$\mathtt{H} = \mathtt{A} + \mu\mathbf{e}'\mathbf{l}^\top = \mathbf{a}\mathbf{b}^\top + \mathbf{c}\mathbf{d}^\top + \mu\mathbf{e}'\mathbf{l}^\top = [\mathbf{a}\ \mathbf{c}\ \mu\mathbf{e}'][\mathbf{b}\ \mathbf{d}\ \mathbf{l}]^\top = [\mathbf{a}\ \mathbf{c}\ \mathbf{e}']\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \mu \end{bmatrix}[\mathbf{b}\ \mathbf{d}\ \mathbf{l}]^\top$$

Taking determinants

$$|\mathtt{H}| = \mu\ |[\mathbf{a}\ \mathbf{c}\ \mathbf{e}'][\mathbf{b}\ \mathbf{d}\ \mathbf{l}]^\top| = \mu\ |\mathbf{a}\mathbf{b}^\top + \mathbf{c}\mathbf{d}^\top + \mathbf{e}'\mathbf{l}^\top| = \mu\ |\mathtt{A} + \mathbf{e}'\mathbf{l}^\top|$$

Using these expressions, and noting that for a matrix $\mathtt{M}$, $|\mathtt{M}| = |\mathtt{M}^\top|$, gives the result of Proposition 4.

A.5 Linear relation amongst curvatures in three views

The objective here is to show that if a plane curve is imaged in three views, then the imaged curvatures at corresponding points are related as $a_1\kappa + a_2\kappa' + a_3\kappa'' = 0$.

There are two steps: first, the homography $\mathtt{H}_{12}$ (and implicitly the support plane of the curve) are computed from corresponding curvatures κ, κ' in the first two views, and it is shown that $\mathtt{H}_{12}$ is linear in the ratio κ'/κ; second, the computed plane is used to transfer the curvature κ to the curvature κ'' in the third view using $\mathtt{H}_{13}$.

Given the corresponding curvatures $\kappa \leftrightarrow \kappa'$ it follows from Proposition 4 that $\mu = \alpha\frac{\kappa'}{\kappa}$, and hence the homography is $\mathtt{H}_{12} = \mathtt{A} + \alpha\frac{\kappa'}{\kappa}\mathbf{e}'\mathbf{l}^\top$. Then from Proposition 5 $\mathtt{H}_{13} = \mathtt{B} + \alpha\frac{\kappa'}{\kappa}\mathbf{e}'\mathbf{l}^\top$.

We now use $\mathtt{H}_{13}$ and Proposition 3 to compute κ'' from κ as

$$\kappa'' = \frac{(x_3'')^3}{|\mathtt{H}_{13}|^2((l_1'')^2 + (l_2'')^2)^{3/2}}\,\kappa$$

and examine the form and degree of each term in turn. We will write $\mathbf{a}$ for $\alpha\frac{\kappa'}{\kappa}\mathbf{e}'$. First x_3'':

$$\mathbf{x}'' = \mathtt{H}_{13}\mathbf{x} = (\mathtt{B} + \mathbf{a}\mathbf{l}^\top)\mathbf{x} = \mathtt{B}\mathbf{x}$$

the last step follows because $\mathbf{x}$ lies on the line $\mathbf{l}$ and so $\mathbf{l}^\top\mathbf{x} = 0$. Thus x_3'' is of zeroth degree in κ'/κ.

$$|\mathtt{H}_{13}| = |\mathtt{B} + \mathbf{a}\mathbf{l}^\top| = |\mathtt{B}||\mathtt{I} + \mathtt{B}^{-1}\mathbf{a}\mathbf{l}^\top| = |\mathtt{B}|(1 + \mathbf{l}^\top\mathtt{B}^{-1}\mathbf{a})$$

Thus $|\mathtt{H}_{13}|$ is linear in κ'/κ. Finally,

$$\mathbf{l}'' = \mathtt{H}_{13}{}^{-\top}\mathbf{l} = (\mathtt{I} - \frac{\mathtt{B}^{-\top}\mathbf{l}\mathbf{a}^\top}{1 + \mathbf{l}^\top\mathtt{B}^{-1}\mathbf{a}})\mathtt{B}^{-\top}\mathbf{l} = \frac{\mathtt{B}^{-\top}\mathbf{l}}{1 + \mathbf{l}^\top\mathtt{B}^{-1}\mathbf{a}}$$

Collecting this together we find

$$\kappa'' = \frac{(\mathtt{B}\mathbf{x})_3^3(1 + \mathbf{l}^\top\mathtt{B}^{-1}\mathbf{a})}{|\mathtt{B}|^2((\mathtt{B}^{-\top}\mathbf{l})_1^2 + (\mathtt{B}^{-\top}\mathbf{l})_2^2)^{3/2}}\,\kappa = \beta(1 + \mathbf{l}^\top\mathtt{B}^{-1}\mathbf{a})\,\kappa = \beta(1 + \mathbf{l}^\top\mathtt{B}^{-1}\alpha\frac{\kappa'}{\kappa}\mathbf{e}')\,\kappa = \beta(1 + \gamma\frac{\kappa'}{\kappa})\,\kappa$$

and hence $\kappa'' = \beta(\kappa + \gamma\kappa')$ as required.

References

[1] N. Ayache. *Stereovision and Sensor Fusion.* MIT Press 1990.

[2] N. Ayache and B. Faverjon. Efficient registration of stereo images by matching graph descriptions of edge segments. *International Journal of Computer Vision*, 1(2):107–132, 1987.

[3] N. Ayache and F. Lustman. Fast and reliable passive stereovision using three cameras. In *International Conference on Computer Vision*, pages 422–427, 1987.

[4] C. Baillard, C. Schmid, A. Zisserman, and A. Fitzgibbon. Automatic line matching and 3D reconstruction of buildings from multiple views. In *ISPRS Conference on Automatic Extraction of GIS Objects from Digital Imagery, IAPRS vol.32, Part 3-2W5*, pages 69–80, 1999.

[5] P. Beardsley, P. Torr, and A. Zisserman. 3D model acquisition from extended image sequences. In *European Conference on Computer Vision*, pages 683–695, 1996.

[6] F. Bignone, O. Henricsson, P. Fua, and M. Stricker. Automatic extraction of generic house roofs from high resolution aerial imagery. In *European Conference on Computer Vision*, pages 85–96, 1996.

[7] A. Blake and R. Cipolla. Robust estimation of surface curvature from deformation of apparent contours. In *European Conference on Computer Vision*, pages 465–474, 1990.

[8] F.L. Bookstein. Fitting conic sections to scattered data. *Computer Graphics and Image Processing*, 9:56–71, 1979.

[9] A.T. Brint and M. Brady. Stereo matching of curves by least deformation. In *International Workshop on Intelligent Robots and Systems*, pages 163–170, 1989.

[10] S. Carlsson, R. Mohr, L. Morin, C. Rothwell, L. Van Gool, F. Veillon, and A. Zisserman. Semi-local projective invariants for the recognition of smooth plane curves. *International Journal of Computer Vision*, 19(3):211–236, 1996.

[11] R.C.K. Chung and R. Nevatia. Use of monocular groupings and occlusion analysis in a hierarchical stereo system. In *Conference on Computer Vision and Pattern Recognition*, pages 50–56, 1991.

[12] R. Cipolla, K. Astrom, and Giblin P. Motion from the frontier of curved surfaces. In *International Conference on Computer Vision*, pages 269–275, 1995.

[13] R. Cipolla and A. Zisserman. Qualitative shape from deformation of image curves. *International Journal of Computer Vision*, 8(1):53–69, 1992.

[14] G. Cross and A. Zisserman. Quadric reconstruction from dual-space geometry. In *International Conference on Computer Vision*, pages 25–31, 1998.

[15] J.L. Crowley and P. Stelmazyk. Measurement and integration of 3D structures by tracking edges lines. In *European Conference on Computer Vision*, pages 269–280, 1990.

[16] R. Deriche and O. Faugeras. Tracking line segments. In *European Conference on Computer Vision*, pages 259–267, 1990.

[17] O. Faugeras. What can be seen in three dimensions with an uncalibrated stereo rig? In *European Conference on Computer Vision*, pages 563–578, 1992.

[18] O. Faugeras. *Three-Dimensional Computer Vision - A Geometric Viewpoint*. Artificial intelligence. MIT Press, 1993.

[19] O. Faugeras and R. Keriven. Complete dense stereovision using level set methods. In *European Conference on Computer Vision*, pages 379–393, 1998.

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

[21] O. Faugeras and L. Robert. What can two images tell us about a third one? *International Journal of Computer Vision*, 18(1):5–19, 1996.

[22] Y. Gdalyahu and D. Weinshall. Flexible syntactic matching of curves and its application to automatic hierarchical classification of silhouettes. IEEE *Transactions on Pattern Analysis and Machine Intelligence*, 21(12):1312–1328, 1999.

[23] P. Gros. Matching and clustering: Two steps towards object modelling in computer vision. *International Journal of Robotics Research*, 14(6):633–642, 1995.

[24] A.W. Gruen. Adaptative least squares correlation: a powerful image matching technique. *S. Afr. Journal of Photogrammetry, Remote Sensing and Cartography*, 14(3):175–187, 1985.

[25] R.I. Hartley. A linear method for reconstruction from lines and points. In *International Conference on Computer Vision*, pages 882–887, 1995.

[26] R.I. Hartley. Lines and points in three views and the trifocal tensor. *International Journal of Computer Vision*, 22(2):125–140, 1997.

[27] R.I. Hartley, R. Gupta, and T. Chang. Stereo from uncalibrated cameras. In *Conference on Computer Vision and Pattern Recognition*, pages 761–764, 1992.

[28] R. Hartley and A. Zisserman. *Multiple View Geometry in Computer Vision.* Cambridge University Press, 2000.

[29] P. Havaldar and G. Medioni. Segmented shape descriptions from 3-view stereo. In *International Conference on Computer Vision*, pages 102–108, 1995.

[30] R. Horaud and T. Skordas. Stereo correspondence through feature grouping and maximal cliques. IEEE *Transactions on Pattern Analysis and Machine Intelligence*, 11(11):1168–1180, 1989.

[31] D.P. Huttenlocher, G.A. Klanderman, and W.J. Rucklidge. Comparing images using the Hausdorff distance. IEEE *Transactions on Pattern Analysis and Machine Intelligence*, 15(9):850–863, 1993.

[32] F. Kahl and A. Heyden. Using conic correspondences in two images to estimate the epipolar geometry. In *International Conference on Computer Vision*, pages 761–766, 1998.

[33] P.J. Laurent. Courbes ouvertes ou fermées par B-splines régularisées. Technical Report RR 652, IMAG, Grenoble, France, 1987.

[34] Q. T. Luong and T. Vieville. Canonical representations for the geometries of multiple projective views. *Computer Vision and Image Understanding*, 64(2):193–229, 1996.

[35] S. Ma. Conics-based stereo, motion estimation, and pose determination. *International Journal of Computer Vision*, 10(1):7–25, 1993.

[36] S.J. Maybank and O.D. Faugeras. A theory of self calibration of a moving camera. *International Journal of Computer Vision*, 8(2):123–151, 1992.

[37] G. Médioni and R. Nevatia. Segment-based stereo matching. *Computer Vision, Graphics and Image Processing*, 31:2–18, 1985.

[38] S.K. Nayar and R.M. Bolle. Reflectance based object recognition. *International Journal of Computer Vision*, 17(3):219–240, 1996.

[39] S. B. Pollard, T. P. Pridmore, J. Porrill, J. E. W. Mayhew, and J. P. Frisby. Geometrical modeling from multiple stereo views. *International Journal of Robotics Research*, 8(4):132–138, 1989.

[40] S.B. Pollard, J.E.W. Mayhew, and J.P. Frisby. PMF: A stereo correspondence algorithm using a disparity gradient constraint. *Perception*, 14:449–470, 1985.

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

[42] P. Pritchett and A. Zisserman. Wide baseline stereo matching. In *International Conference on Computer Vision*, pages 754–760, 1998.

[43] L. Quan. Conic reconstruction and correspondence from two views. IEEE *Transactions on Pattern Analysis and Machine Intelligence*, 18(2):151–160, 1996.

[44] P. Remagnino, P. Brand, and R. Mohr. Correlation techniques in adaptative template matching with uncalibrated cameras. In *Vision Geometry III, SPIE's international symposium on photonic sensors & control for commercial applications*, vol. 2356, pages 252–253, 1994.

[45] L. Robert and O.D. Faugeras. Curve-based stereo: Figural continuity and curvature. In *Conference on Computer Vision and Pattern Recognition*, pages 57–62, 1991.

[46] C. Rothwell, J.L. Mundy, and B. Hoffman. Representing objects using topology. In *Object Representation in Computer Vision II LNCS 1144*, pages 79–108. Springer-Verlag, 1996.

[47] C. Rothwell, A. Zisserman, D. Forsyth, and J. Mundy. Fast recognition using algebraic invariants. In *Geometric Invariance in Computer Vision*, pages 398–407. MIT Press, 1992.

[48] C.A. Rothwell. *Recognition using Projective Invariance.* PhD thesis, Department of Engineering Science, University of Oxford, 1993.

[49] R. Safaee-Rad, I. Tchoukanov, B. Benhabib, and K.C. Smith. 3D pose estimation from a quadratic curved feature in two perspective views. In *International Conference on Pattern Recognition*, pages 341–344, 1992.

[50] C. Schmid and A. Zisserman. Automatic line matching across views. In *Conference on Computer Vision and Pattern Recognition*, pages 666–671, 1997.

[51] C. Schmid and A. Zisserman. The geometry and matching of curves in multiple views. In *European Conference on Computer Vision*, vol. I, pages 394–409, 1998.

[52] J. Semple and G. Kneebone. *Algebraic Projective Geometry.* Oxford University Press, 1979.

[53] A. Shashua. Trilinearity in visual recognition by alignment. In *European Conference on Computer Vision*, pages 479–484, 1994.

[54] M.E. Spetsakis and J. Aloimonos. Structure from motion using line correspondences. *International Journal of Computer Vision*, 4(3):171–183, 1990.

[55] P. H. S. Torr and D. W. Murray. The development and comparison of robust methods for estimating the fundamental matrix. *International Journal of Computer Vision*, 24(3):271–300, 1997.

[56] P. H. S. Torr and A. Zisserman. Robust parameterization and computation of the trifocal tensor. *Image and Vision Computing*, 15:591–605, 1997.

[57] B. Triggs. Autocalibration and the absolute quadric. In *Conference on Computer Vision and Pattern Recognition*, pages 609–614, 1997.

[58] R. Vaillant. Using occluding contours for 3D object modeling. In *European Conference on Computer Vision*, pages 454–464, 1990.

[59] L. Van Gool, M. Brill, E. Barrett, T. Moons, and E. Pauwels. Semi-differential invariants for nonplanar curves. In *Geometric Invariance in Computer Vision*, pages 293–309. MIT Press, 1992.

[60] L. Van Gool, L.T. Moons, E. Pauwels, and A. Oosterlinck. Semi-differential invariants. In *Geometric Invariance in Computer Vision*, pages 157–192. MIT Press, 1992.

[61] V. Venkateswar and R.Chellappa. Hierarchical stereo and motion correspondence using feature groupings. *International Journal of Computer Vision*, 15(3):245–269, 1995.

[62] Y. Zhang and J.J. Gerbrands. Method for matching general stereo planar curves. *Image and Vision Computing*, 13(8):645–655, 1995.

[63] Z. Zhang. Token tracking in a cluttered scene. *Image and Vision Computing*, 12(2):110–120, 1994.

[64] Z. Zhang, R. Deriche, O. Faugeras, and Q.T. Luong. A robust technique for matching two uncalibrated images through the recovery of the unknown epipolar geometry. *Artificial Intelligence*, 78:87–119, 1995.

[65] A. Zisserman and S. Maybank. A case against epipolar geometry. In *Applications of Invariance in Computer Vision LNCS 825*, pages69–88. Springer-Verlag, 1994.