Joint Orientation of Epipoles

Ondřej Chum, Tomáš Werner, and Tomáš Pajdla

Center for Machine Perception, `http://cmp.felk.cvut.cz`
Czech Technical University, Karlovo nám. 13, 121 35 Prague
Czech Republic

Abstract

It is known that epipolar constraint can be augmented with orientation by formulating it in the oriented projective geometry. This oriented epipolar constraint requires knowing the orientations (signs of overall scales) of epipoles and fundamental matrix. The current belief is that these orientations cannot be obtained from the fundamental matrix only and that additional information is needed, typically, a single correct point correspondence. In contrary to this, we show that fundamental matrix alone encodes orientation of epipoles up to their common scale sign. We present two formulations of this fact. The algebraic formulation gives a closed formula to compute the second epipole from fundamental matrix and the first epipole. The geometric formulation is in terms of the conic formed by intersections of corresponding epipolar lines in the common image plane; we show that the epipoles always lie on different antipodal components of the spherical interpretation of this conic. Further, we show that, under mild assumptions, fundamental matrix can discriminate between two classes of mutual position of a pair of directional cameras.

1 Introduction

Real[1] cameras project points from space into images such that points on a single *half-ray* emanating from a projection center into infinity project into a single image point. Most of real cameras used in photography and computer vision, however, posses a view angle smaller than $180°$, and therefore they may never see points on both half-rays of one camera ray. Thus, the images of points in space may be represented by complete lines through the projection center instead of half-lines since there is maximally one visible half-ray on each ray through the projection center. This representation allows to model the image plane by the *projective plane* at the cost of loosing the ability to represent the 'front' and the 'back' of a camera and consequently of loosing some of the constraints that are useful in solving the image correspondence problem [9, 13, 6, 5].

To express that points are in front of a camera, Stolfi introduced *oriented projective geometry* [11] and used it in computer graphics. Paper [4] introduced the concept of oriented projective geometry to computer vision and pointed out that it can be used in solving the image correspondence problem. The oriented projective geometry is implicitly expressed by the tensor algebra used in [12] to describe geometry of multiple

---

images[2]. It has been shown [2] that the existence of the 'front' of a camera can be used to upgrade the projective reconstruction [3] to the *quasi-affine reconstruction* that preserves convex hulls. Works [7] and [1] applied the above concepts to reconstruction and camera calibration respectively. Ray orientation consistency provided a constraint on five point correspondences in two images in [14].

In this work we show that the information about the 'front' of the camera, when combined with the knowledge of the standard epipolar geometry, can be used to derive a new useful constraint on the correspondence of epipolar half-lines that arise by projecting camera half-rays into images. The contributions of this paper are (i) derivation of formula for computation of jointly correctly oriented epipoles given unoriented fundamental matrix only; (ii) we show that the joint orientation of epipoles also discriminates the mutual position of directional cameras (see rows of figure 1) provided that image bases have equal handedness; (iii) the concept of Steiner conic [3] is extended into oriented epipolar geometry: it is shown that the epipoles lie on different components of the spherical interpretation of this conic; (iv) results are applied in wide baseline guided matching context, the verification of hypothesised region-to-region correspondence can be done without any search in the image transformation space.

2 Notation and Concepts

$\{\cdots\}$ denotes a set and $(\cdots)$ an ordered tuple. $\mathbb{R}^n$ is the real vector n-space. Its elements are typeset in boldface, e.g., $\mathbf{x}$. $|\mathtt{A}|$ denotes the determinant of a square matrix. Determinant of the matrix with vectors as its columns is abbreviated as $|\mathbf{x}_1,\ldots,\mathbf{x}_n|$. $[\mathbf{x}]_\times$ denotes antisymmetric 3×3 matrix such that $\mathbf{x} \times \mathbf{y} = [\mathbf{x}]_\times \mathbf{y}$. For a $(n-1) \times n$ matrix $\mathtt{A}$, the symbol $\mathtt{A}^\wedge$ denotes the column n-vector whose n-th entry is the minor of the first $n-1$ columns of $\mathtt{A}$, and so on by cyclic permutation. E.g., for $n=4$ and denoting $\mathtt{A} = [\mathbf{a}_1,\mathbf{a}_2,\mathbf{a}_3,\mathbf{a}_4]$ it is $\mathtt{A}^\wedge = [|\mathbf{a}_2,\mathbf{a}_3,\mathbf{a}_4|,|\mathbf{a}_3,\mathbf{a}_4,\mathbf{a}_1|,|\mathbf{a}_4,\mathbf{a}_1,\mathbf{a}_2|,|\mathbf{a}_1,\mathbf{a}_2,\mathbf{a}_3|]^\top$.

Symbol $\sim$ denotes equivalence of two vectors up to scale, $\mathbf{x} \sim \lambda\mathbf{x}$ for all $\lambda \neq 0$. Similarly, $\overset{+}{\sim}$ denotes equivalence of two vectors up to a positive scale, $\mathbf{x} \overset{+}{\sim} \lambda\mathbf{x}$ for all $\lambda > 0$. For scalars, $a \overset{+}{\sim} b$ means either $ab > 0$ or $a = b = 0$.

We extend the definition of these equivalences as follows. The symbol $\mathbf{x}_\sim$ denotes the set of all vectors equivalent with $\mathbf{x}$ under equivalence $\sim$, i.e., $\mathbf{x}_\sim = \{\lambda\mathbf{x} \,|\, \lambda \neq 0\}$. Similarly, $\mathbf{x}_{\overset{+}{\sim}} = \{\lambda\mathbf{x} \,|\, \lambda > 0\}$. For an ordered n-tuple $(\mathbf{x}_1,\ldots,\mathbf{x}_n)$ of elements of (possibly different) vector spaces, we define $(\mathbf{x}_1,\ldots,\mathbf{x}_n)_{\overset{+}{\sim}} = \{(\lambda_1\mathbf{x}_1,\ldots,\lambda_n\mathbf{x}_n) \,|\, \delta\lambda_1 > 0,\ldots,\delta\lambda_n > 0, \delta \neq 0\}$.

The spherical (or oriented projective) d-space is the set $\mathbb{S}^d = \{\mathbf{x} \in \mathbb{R}^{d+1} \setminus \mathbf{0}\} / \overset{+}{\sim}$, where $/$ denotes factorization by equivalence. The *oriented projective geometry* is a structure on $\mathbb{S}^d$ [11], similar to the projective geometry being a structure on the projective space.

Scene, Image and Camera[3]. The *scene* is modelled by the affine 3-space $\mathbb{A}^3$ which in turn is represented by a single open half of $\mathbb{S}^3$. This half is denoted by the pair $(\mathbb{S}^3, \pi_\infty)$ and consists of scene points that are in front of the *oriented plane at infinity* π_∞, i.e., of points $\mathbf{X}$ satisfying $\pi_\infty\mathbf{X} > 0$.

---

[2]In particular, the second formula in table 1 in the full version of [12] gives the relation between fundamental matrix and the epipoles. This observation is very close to our paper; however, the orientation consequences of the formula are not discussed in [12].

[3]Note that the vectors and matrices in this subsection represent objects of oriented projective geometry, therefore their signs matter.

The *image* is represented by $\mathbb{S}^2$. The whole of $\mathbb{S}^2$ represents the *omnidirectional* image[4]. Conventional (from, e.g., TV or photographic cameras) images are *directional*. The directional image is modelled by $\mathbb{A}^2$ which is represented by a single open half of $\mathbb{S}^2$. This half is denoted by $(\mathbb{S}^2, \mathbf{l}_\infty)$ and consists of image points that are on the positive side of the *oriented image line at infinity* $\mathbf{l}_\infty$, i.e., of points $\mathbf{x}$ satisfying $\mathbf{l}_\infty \mathbf{x} > 0$.

Camera is a linear mapping, $\mathbf{x} = \mathrm{P}\mathbf{X}$, from the scene to the image. The *directional camera* is, besides that, characterized by its oriented line at infinity, therefore denoted by $(\mathrm{P}, \mathbf{l}_\infty)$. The line $\mathbf{l}_\infty$ is a projection of the scene plane $\pi_\mathrm{P} = \mathbf{l}_\infty \mathrm{P}$ (without the camera center) passing through the camera center. Since $\pi_\mathrm{P}\mathbf{X} = (\mathbf{l}_\infty\mathrm{P})\mathbf{X} = \mathbf{l}_\infty(\mathrm{P}\mathbf{X}) = \mathbf{l}_\infty\mathbf{x}$, the condition $\mathbf{l}_\infty\mathbf{x} > 0$ is equivalent to $\pi_\mathrm{P}\mathbf{X} > 0$, saying that visible scene points $\mathbf{X}$ lie in front of π_P.

3 Oriented Epipolar Constraints

Consider a scene point represented by $\mathbf{X}$ and a pair of different cameras represented by matrices $(\mathrm{P}, \mathrm{P}')$. It is well-known [3] that, in the (unoriented) projective geometry, a point pair $(\mathbf{x}, \mathbf{x}')$ such that $\mathbf{x} \sim \mathrm{P}\mathbf{X}$ and $\mathbf{x}' \sim \mathrm{P}'\mathbf{X}$ satisfies $\mathbf{x}'^\top \mathrm{F}\mathbf{x} = 0$. The fundamental matrix F is obtained from the cameras as $\mathrm{F} = [\mathbf{e}']_\times \mathrm{P}'\mathrm{P}^+$, where $\mathrm{P}^+ = \mathrm{P}(\mathrm{P}\mathrm{P}^\top)^{-1}$, $\mathbf{e}' = \mathrm{P}'\mathbf{C}$, and $\mathbf{C}$ is an arbitrary non-zero vector satisfying $\mathrm{P}\mathbf{C} = \mathbf{0}$.

Taking orientation into account provides a more restrictive constraint on point correspondences. If $(\mathbf{x}, \mathbf{x}')$ are such that $\mathbf{x} \overset{+}{\sim} \mathrm{P}\mathbf{X}$ and $\mathbf{x}' \overset{+}{\sim} \mathrm{P}'\mathbf{X}$, it follows that [15]

$$\mathbf{e}' \times \mathbf{x}' \overset{+}{\sim} \mathrm{F}\mathbf{x}, \tag{1}$$

where again $\mathrm{F} = [\mathbf{e}']_\times \mathrm{P}'\mathrm{P}^+$ and $\mathbf{e}' = \mathrm{P}'\mathbf{C}$. Unlike in $\mathbf{x}'^\top \mathrm{F}\mathbf{x} = 0$, the scale signs of $\mathbf{x}, \mathbf{x}', \mathbf{e}'$ and F matter in (1). Therefore, $\mathbf{C}$ cannot be defined by $\mathrm{P}\mathbf{C} = \mathbf{0}$ because it leaves ambiguity in its scale sign. Instead, the 'oriented camera center' [2, 3, 15] is used, $\mathbf{C} = \mathrm{P}^\wedge$.

In the oriented framework, it is also possible to formulate a constraint on a *pair* of point correspondences that does not involve fundamental matrix. If $(\mathbf{x}_1, \mathbf{x}'_1)$ and $(\mathbf{x}_2, \mathbf{x}'_2)$ satisfy $\mathbf{x}_i \overset{+}{\sim} \mathrm{P}\mathbf{X}_i$ and $\mathbf{x}'_i \overset{+}{\sim} \mathrm{P}'\mathbf{X}_i$ for two scene points $\mathbf{X}_1$ and $\mathbf{X}_2$, it follows that

$$|\mathbf{x}'_1, \mathbf{x}'_2, \mathbf{e}'| \overset{+}{\sim} -|\mathbf{x}_1, \mathbf{x}_2, \mathbf{e}| \tag{2}$$

where $\mathbf{e} = \mathrm{P}\mathrm{P}'^\wedge$, $\mathbf{e}' = \mathrm{P}'\mathrm{P}^\wedge$. This can be proved using equality $|\mathrm{P}\mathbf{X}, \mathrm{P}\mathbf{Y}, \mathrm{P}\mathbf{Z}| = |\mathbf{X}, \mathbf{Y}, \mathbf{Z}, \mathrm{P}^\wedge|$ relating any 4-vectors $\mathbf{X}, \mathbf{Y}, \mathbf{Z}$ and 3×4 matrix P.

Equation (2) can be interpreted as a constraint on *oriented image lines* in two images (there is no constraint on unoriented lines). Let $(\mathbf{X}_1, \mathbf{X}_2)$ be a 3D line specified by its two points. Then line pair $(\mathbf{l}, \mathbf{l}')$ such that $\mathbf{l}^\top \overset{+}{\sim} \mathrm{P}\mathbf{X}_1 \times \mathrm{P}\mathbf{X}_2$ and $\mathbf{l}'^\top \overset{+}{\sim} \mathrm{P}'\mathbf{X}_1 \times \mathrm{P}'\mathbf{X}_2$ satisfy [15]

$$\mathbf{l}'\mathbf{e}' \overset{+}{\sim} -\mathbf{l}\mathbf{e}. \tag{3}$$

The constraints (1) and (3) are useful for feature matching. However, unlike when using the unoriented version $\mathbf{x}'^\top \mathrm{F}\mathbf{x} = 0$, we have to choose scale signs of the representing vectors and matrices. Scale signs of image points and lines are measured in images; however, $\mathbf{e}, \mathbf{e}'$ and F are typically known only up to *non-zero* scales because they are computed as null-vectors when solving homogeneous linear systems. There are three following relevant situations how to determine consistent scale signs of (some of) $\mathbf{e}, \mathbf{e}'$ and F:

---

[4]Cameras able to capture omnidirectional images can be built, e.g., by combining a conventional camera and a curved mirror [8].

- If $F_\sim$ and a single point correspondence $(\mathbf{x},\mathbf{x}')_{\overset{\pm}{\approx}}$ are known, $(\mathbf{e}',F)_{\overset{\pm}{\approx}}$, where $\mathbf{e}'$, F and all correspondences satisfy (1), can be obtained.
- If $F_\sim$ and a single oriented line correspondence $(\mathbf{l},\mathbf{l}')_{\overset{\pm}{\approx}}$ is known, $(\mathbf{e},\mathbf{e}')_{\overset{\pm}{\approx}}$, where $\mathbf{e}$, $\mathbf{e}'$ and all line correspondences satisfy (3), can be obtained. However, $(\mathbf{e}',F)_{\overset{\pm}{\approx}}$ satisfying (1) remains unknown.
- If $F_\sim$ is known, it was believed that neither $(\mathbf{e}',F)_{\overset{\pm}{\approx}}$ satisfying (1) nor $(\mathbf{e},\mathbf{e}')_{\overset{\pm}{\approx}}$ satisfying (2) can be obtained. However, this is not true: *this paper shows that $(\mathbf{e},\mathbf{e}')_{\overset{\pm}{\approx}}$, where $\mathbf{e}$, $\mathbf{e}'$ and all correspondences satisfy (1), can be obtained from arbitrarily scaled fundamental matrix only*.

4   Invariance of Joint Orientation of Epipoles

It is natural to define epipoles as tensors associated with the camera pair (P,P') [12],

$$\mathbf{e} = PP'^\wedge , \quad \mathbf{e}' = P'P^\wedge . \tag{4}$$

In the (unoriented) projective geometry, the image points represented by $\mathbf{e}$ and $\mathbf{e}'$ are invariant to scene homographies (i.e., changing the scene coordinate basis) and overall scales of the camera matrices. To see this, consider two camera pairs, $(\hat{P},\hat{P}')$ and (P,P'), related by $\hat{P} = \delta PT^{-1}$ and $\hat{P}' = \delta'P'T^{-1}$, where T is a scene homography and δ,δ' are scales of camera matrices. Using the fact that

$$(HPT^{-1})^\wedge = |H||T|^{-1}TP^\wedge , \tag{5}$$

we obtain

$$\begin{aligned}
\hat{\mathbf{e}} &= \hat{P}\hat{P}'^\wedge = (\delta\delta'^3|T|^{-1})\mathbf{e} = \sigma\mathbf{e} , \\
\hat{\mathbf{e}}' &= \hat{P}'\hat{P}^\wedge = (\delta^3\delta'|T|^{-1})\mathbf{e}' = \sigma'\mathbf{e}' .
\end{aligned} \tag{6}$$

Since $\sigma\sigma' \neq 0$, it is $\hat{\mathbf{e}} \sim \mathbf{e}$, hence $\hat{\mathbf{e}}$ and $\mathbf{e}$ represent the same image point. Similarly, $\hat{\mathbf{e}}' \sim \mathbf{e}'$.

In the oriented projective geometry, this is not true in general because σ and/or σ' might be negative. E.g., if $\delta = \delta' = 1$ and $|T| < 0$, it is $\hat{\mathbf{e}} \overset{\pm}{\sim} -\mathbf{e}$ and $\hat{\mathbf{e}}' \overset{\pm}{\sim} -\mathbf{e}'$. Since $\mathbf{e}$ and $-\mathbf{e}$ are different (mutually antipodal points) in the oriented projective geometry, the epipoles are not invariant to a mirroring scene homography. However, the following theorem says that the epipoles can swap sign only *simultaneously*. In other words, it is $\sigma\sigma' > 0$, which can be seen from (6).

Theorem 1 *Let $\mathbf{e} = PP'^\wedge$ and $\mathbf{e}' = P'P^\wedge$. The set $(\mathbf{e},\mathbf{e}')_{\overset{\pm}{\approx}}$ is invariant to scene homographies and scaling P and P'.*

5   Recovering Epipoles from Fundamental Matrix

The set $(\mathbf{e},\mathbf{e}')_{\overset{\pm}{\approx}}$ mentioned in theorem 1 can be obtained from the fundamental matrix associated with the camera pair by computing arbitrary camera matrices consistent with F and then letting $\mathbf{e} = PP'^\wedge$ and $\mathbf{e}' = P'P^\wedge$. The following theorem gives another way to do this: choose any vector $\mathbf{e}'$ satisfying $\mathbf{e}'^\top F = \mathbf{0}$ and compute $\mathbf{e}$ using one of the equivalent formulas in (7).
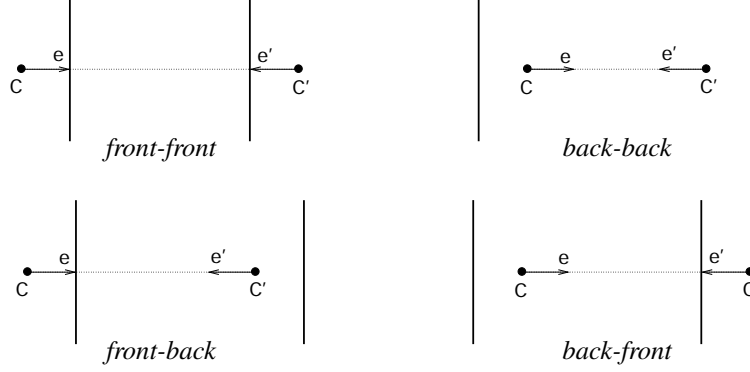
Figure 1: Four configuration classes of a pair of directional cameras in the affine scene.

Theorem 2 *Let* $(\mathtt{P}, \mathtt{P}')$ *be a pair of camera matrices and* $\mathtt{F}$ *be an arbitrarily scaled fundamental matrix consistent with it. Then vectors* $\mathbf{e} = \mathtt{P}\mathtt{P}'^{\wedge}$ *and* $\mathbf{e}' = \mathtt{P}'\mathtt{P}^{\wedge}$ *satisfy*

$$\begin{bmatrix} \mathbf{e} \\ 0 \end{bmatrix} \doteq [\mathtt{F} \,|\, \mathbf{e}']^{\wedge}, \quad [\mathbf{e}]_{\times} \doteq \mathtt{F}^{\top}[\mathbf{e}']_{\times}\mathtt{F} . \tag{7}$$

Proof. The family of pairs of camera matrices consistent with $\mathtt{F}$ is, up to their non-zero scales and a scene homography, given [3] by $\mathtt{P} = [\mathtt{I}\,|\,\mathbf{0}]$ and $\mathtt{P}' = [\mathtt{H}\,|\,\mathbf{e}']$ where $\mathbf{e}'$ is any vector satisfying $\mathbf{e}'^{\top}\mathtt{F} = \mathbf{0}$ and $\mathtt{H} = [\mathbf{e}']_{\times}\mathtt{F}$.

Clearly, it is $\mathtt{P}'\mathtt{P}^{\wedge} = \mathbf{e}'$. The first epipole is given by $\mathbf{e} = \mathtt{P}\mathtt{P}'^{\wedge} = [\mathtt{I}\,|\,\mathbf{0}][\mathtt{H}\,|\,\mathbf{e}']^{\wedge}$. Definition of $\mathtt{A}^{\wedge}$ and using the identity $|\mathbf{x}, \mathbf{x} \times \mathbf{y}, \mathbf{x} \times \mathbf{z}| = (\mathbf{x}^{\top}\mathbf{x})\,|\mathbf{x}, \mathbf{y}, \mathbf{z}|$ yields

$$\mathbf{e} = - \begin{bmatrix} |\mathbf{e}', \mathbf{h}_2, \mathbf{h}_3| \\ |\mathbf{h}_1, \mathbf{e}', \mathbf{h}_3| \\ |\mathbf{h}_1, \mathbf{h}_2, \mathbf{e}'| \end{bmatrix} = -(\mathbf{e}'^{\top}\mathbf{e}') \begin{bmatrix} |\mathbf{e}', \mathbf{f}_2, \mathbf{f}_3| \\ |\mathbf{f}_1, \mathbf{e}', \mathbf{f}_3| \\ |\mathbf{f}_1, \mathbf{f}_2, \mathbf{e}'| \end{bmatrix} \tag{8}$$

where $\mathbf{h}_i$ denote columns of $\mathtt{H}$ and $\mathbf{f}_i$ denote columns of $\mathtt{F}$. Using the fact $\mathbf{e}'^{\top}\mathbf{e}' > 0$ and definitions of $\mathtt{A}^{\wedge}$ and $[\mathbf{x}]_{\times}$, (8) can be re-arranged to the left or right formula in (7). □

6 Mutual Position of Two Directional Cameras

Non-invariance of the epipoles to scene homographies is caused by non-covariance of the camera center defined by $\mathtt{P}^{\wedge}$. By applying a homography $\mathtt{T}$ to the scene, scene points transform as $\mathbf{X} \mapsto \mathtt{T}\mathbf{X}$, while $\mathtt{P}^{\wedge}$ transforms as $\mathtt{P}^{\wedge} \mapsto |\mathtt{T}|^{-1}\mathtt{T}\mathtt{P}^{\wedge}$. If $|\mathtt{T}| < 0$, $\mathtt{P}^{\wedge}$ is replaced by its antipode. Also, $\mathtt{P}^{\wedge}$ becomes its antipode when transforming the image by a mirroring homography.

This causes confusion when trying to define the concept of 'camera center' and, for directional cameras, relations 'the second camera is in front/behind the first camera'. These have clear meaning only in the affine scene, i.e., with respect to the plane at infinity. $\mathtt{P}^{\wedge}$ represents the camera center *only if* $\pi_{\infty}\mathtt{P}^{\wedge} > 0$, otherwise, $\mathtt{P}^{\wedge}$ does not belong to the (affine) scene $(\mathbb{S}^3, \pi_{\infty})$. If $\pi_{\infty}\mathtt{P}'^{\wedge} > 0$, then the sign of $\mathbf{l}_{\infty}\mathbf{e} = \pi_{\mathtt{P}}\mathtt{P}'^{\wedge}$ correctly determines whether the second camera lies in front of ($\mathbf{l}_{\infty}\mathbf{e} > 0$) or behind ($\mathbf{l}_{\infty}\mathbf{e} < 0$) the first camera.

Let $(\mathtt{P}, \mathbf{l}_{\infty})$ and $(\mathtt{P}', \mathbf{l}'_{\infty})$ be two directional cameras in the affine scene $(\mathbb{S}^3, \pi_{\infty})$. Assume that[5] $\pi_{\infty}\mathtt{P}^{\wedge} > 0$ and $\pi_{\infty}\mathtt{P}'^{\wedge} > 0$. Then the signs of $\mathbf{l}_{\infty}\mathbf{e}$ and $\mathbf{l}'_{\infty}\mathbf{e}'$ correctly discriminate

---

[5]E.g., we can choose $\pi_{\infty} = [0,0,0,1]^{\top}$, $\mathtt{P} = \mathtt{K}\mathtt{R}[\mathtt{I}\,|\,-\mathbf{t}]$, $\mathtt{P}' = \mathtt{K}'\mathtt{R}'[\mathtt{I}\,|\,-\mathbf{t}']$ where $|\mathtt{K}\mathtt{R}| > 0, |\mathtt{K}'\mathtt{R}'| > 0$.

between four classes of camera pair configurations shown in figure 1. Assume that the camera pair undergoes an unknown scene homography and scaling of the camera matrices (e.g., this new camera pair is obtained by projective reconstruction from image correspondences). We have seen that the signs of $\mathbf{l}_\infty\mathbf{e}$ and $\mathbf{l}'_\infty\mathbf{e}'$ are not invariant to these transformations, therefore we cannot distinguish between the four classes any more. However, the sign of the *product* $s = (\mathbf{l}_\infty\mathbf{e})(\mathbf{l}'_\infty\mathbf{e}')$ is invariant to these transformations, which follows from theorem 1. The sign of s discriminates between two classes of camera pair configurations, corresponding to the first ($s > 0$) and the second ($s < 0$) row of figure 1.

The following theorem, central to the paper, is a consequence of this fact.

Theorem 3 *Let* $(\mathtt{P},\mathbf{l}_\infty)$ *and* $(\mathtt{P}',\mathbf{l}'_\infty)$ *be two directional cameras in the affine scene* $(\mathbb{S}^3, \pi_\infty)$ *such that* $\pi_\infty\mathtt{P}^\wedge > 0$ *and* $\pi_\infty\mathtt{P}'^\wedge > 0$. *Let* $\mathtt{F}$ *be a fundamental matrix associated with them. Let* $\mathbf{e}$ *and* $\mathbf{e}'$ *satisfy* $\mathbf{e}'^\top\mathtt{F} = \mathbf{0}$ *and (7). Then the camera configuration corresponds to the first row of figure 1 if* $(\mathbf{l}_\infty\mathbf{e})(\mathbf{l}'_\infty\mathbf{e}') > 0$ *and to the second row if* $(\mathbf{l}_\infty\mathbf{e})(\mathbf{l}'_\infty\mathbf{e}') < 0$.

7 Geometric Interpretation

If the two image planes are identified by a projective isomorphism, epipolar geometry can be given *geometric* interpretation [3, p. 231]. It involves the epipoles, the conic represented by the symmetric part of the fundamental matrix, $\mathtt{F}_s = (\mathtt{F} + \mathtt{F}^\top)/2$, and the point $\mathbf{x}_a$ represented by its antisymmetric part[6], $[\mathbf{x}_a]_\times = \mathtt{F}_a = (\mathtt{F} - \mathtt{F}^\top)/2$. In this section, we re-formulate this interpretation in the oriented projective geometry.

The unoriented situation is shown in figure 8.10a in [3, p. 233]. The conic given by $\mathbf{x}^\top\mathtt{F}_s\mathbf{x} = 0$ passes through both epipoles, $\mathbf{e}^\top\mathtt{F}_s\mathbf{e} = \mathbf{e}'^\top\mathtt{F}_s\mathbf{e}' = 0$, which follows from $\mathtt{F}\mathbf{e} = \mathtt{F}^\top\mathbf{e}' = \mathbf{0}$. Recall Steiner's theorem [10], saying that intersections of corresponding lines of two projectively related line pencils form a conic. Epipolar lines are such pencils, therefore this conic is $\mathtt{F}_s$. Point $\mathbf{x}_a$ has pole-polar relationship with line $\mathbf{e} \times \mathbf{e}'$ joining the epipoles, $\mathtt{F}_s\mathbf{x}_a \sim \mathbf{e} \times \mathbf{e}'$. In other words, lines $\mathbf{x}_a \times \mathbf{e}$ and $\mathbf{x}_a \times \mathbf{e}'$ are tangent to the conic.

The following theorem augments this construction with orientation.

Theorem 4 *Let* $(\mathtt{P},\mathtt{P}')$ *be camera matrices and* $\mathtt{F}$ *fundamental matrix consistent with them such that* $|\mathtt{F} + \mathtt{F}^\top| \neq 0$. *Let* $\mathbf{e} = \mathtt{P}\mathtt{P}'^\wedge$, $\mathbf{e}' = \mathtt{P}'\mathtt{P}^\wedge$, $\mathtt{F}_s = (\mathtt{F} + \mathtt{F}^\top)/2$, *and* $[\mathbf{x}_a]_\times = (\mathtt{F} - \mathtt{F}^\top)/2$. *Then*

1. *The points represented by* $\mathbf{e}$ *and* $\mathbf{e}'$ *lie on different antipodal components of the spherical conic[7] represented by* $\mathtt{F}_s$ *(see figure 2a).*
2. *The pole-polar relation of* $\mathbf{x}_a$ *and* $\mathbf{e} \times \mathbf{e}'$ *is oriented,* $\mathtt{F}_s\mathbf{x}_a \overset{+}{\sim} \mathbf{e} \times \mathbf{e}'$.

Proof. Item 1 will be proved by finding a line separating the antipodal components of $\mathtt{F}_s$ and at the same time the epipoles. Clearly, any tangent to $\mathtt{F}_s$ not incident with any epipole separates the antipodal components of $\mathtt{F}_s$. It remains to show that this tangent also separates the epipoles.

---

[6]Note, $\mathtt{F}_s$ can be rank deficient or zero. It occurs, e.g., if $\mathbf{e} \sim \mathbf{e}'$. The case $|\mathtt{F}_s| = 0$ can always be transformed to $|\mathtt{F}_s| \neq 0$ by proper image homographies. Thus, we consider only case $|\mathtt{F}_s| \neq 0$ without loss of generality.

[7]In oriented projective geometry, a non-degenerate conic consists of *two disconnected antipodal components*, as shown in figure 2a. The sign of $\mathtt{F}_s$ does not distinguish between these antipodes, it rather says whether $\mathbf{x}^\top\mathtt{F}_s\mathbf{x} > 0$ defines interior or exterior points. The sign of $\mathtt{F}_s$ can be thought of as an arrow on the conic aiming either inside or outside.
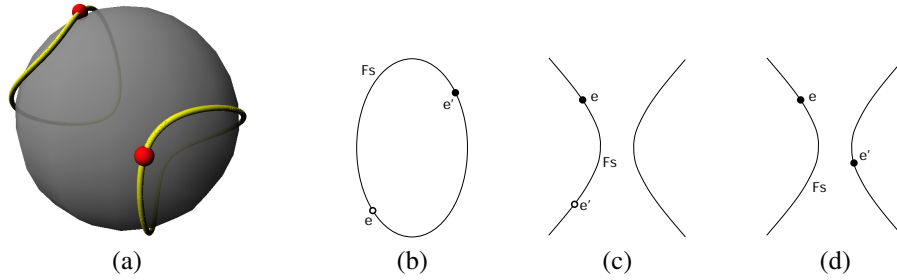
Figure 2: Epiboles are on the antipodal parts of spherical conic F_s *(a). 3 configurations of epiboles on affine conic* F_s*; signs of* $\mathbf{l}_\infty \mathbf{e}$ *and* $\mathbf{l}'_\infty \mathbf{e}'$ *denoted by filled or empty circles (b,c,d).*

A line $\mathbf{l}$ separates points $\mathbf{e}$ and $\mathbf{e}'$ iff $(\mathbf{le})(\mathbf{le}') < 0$. We need to prove this inequality for the tangent $\mathbf{l}^\top = F_s \mathbf{x}$ at any point $\mathbf{x}$ on the conic not identical to any epipole, i.e., $(\mathbf{e}^\top F_s \mathbf{x})(\mathbf{e}'^\top F_s \mathbf{x}) < 0$. Since $2F_s = F + F^\top$ and $F\mathbf{e} = F^\top \mathbf{e}' = \mathbf{0}$, it is $\mathbf{e}^\top F_s = \frac{1}{2}\mathbf{e}^\top F$ and $\mathbf{e}'^\top F_s = \frac{1}{2}\mathbf{e}'^\top F^\top$. Substitution gives $(\mathbf{e}^\top F \mathbf{x})(\mathbf{e}'^\top F^\top \mathbf{x}) < 0$. We know that $F\mathbf{x} = \delta \mathbf{e}' \times \mathbf{x}$ and $F^\top \mathbf{x} = \delta \mathbf{e} \times \mathbf{x}$ for $\delta \neq 0$. Substitution gives $\delta^2 |\mathbf{e}, \mathbf{e}', \mathbf{x}||\mathbf{e}', \mathbf{e}, \mathbf{x}| < 0$ which is true.

Item 2 can be proved analytically. By algebraic manipulations, it can be shown that if $F = [\mathbf{e}']_\times P'P^+$, then $F_s \mathbf{x}_a = \frac{1}{2}\mathbf{e} \times \mathbf{e}'$. $\qquad\qquad\square$

Note that the theorem holds for *any* cameras, omnidirectional or directional. For directional cameras in particular, item 1 of the theorem, along with affine classification of F_s and theorem 3, allows to discriminate between the two classes of camera configurations corresponding to the rows in figure 1. It is $(\mathbf{l}_\infty \mathbf{e})(\mathbf{l}'_\infty \mathbf{e}') > 0$ if and only if F_s is a hyperbola and the epipoles lie on its different branches, because branches of a hyperbola correspond to the antipodal components of the spherical interpretation of the conic (figure 2bcd).

Disambiguating re-composition. It has been noticed [3, p. 233] that the objects represented by the symmetric and the antisymmetric parts of the fundamental matrix define the epipolar geometry only up to swapping the images. This refers to the situation in projective geometry, i.e., we are given F_s and $\mathbf{x}_a$ up to *non-zero* scales and want to get back F. Indeed, the two tangents to a conic from a point have undefined order. Algebraically, the unknown scales of F_a and $\mathbf{x}_a$ are given by condition $|\alpha F_s + \beta F_a| = 0$. It can be easily shown that the three roots of this cubic equation correspond to combinations $F_s + F_a$, $-F_s + F_a$ and F_a. The second one equals $-F^\top$, hence the reversed order of images.

In the oriented projective geometry, this re-composition ambiguity disappears, because we are given F_s and F_a up to *positive* scales. Thus we are left with the first root. This is given, quite independently, also by item 2 of theorem 4.

Horopter. Conic F_s is the locus of corresponding points $(\mathbf{x}, \mathbf{x}')$ with the same coordinates, $\mathbf{x} \sim \mathbf{x}'$. This is because $\mathbf{x}^\top F\mathbf{x} = \mathbf{x}^\top F_s \mathbf{x} = 0$. The conic is the image of the *horopter* curve in the scene, formed by 3D points $\mathbf{X}$ projecting to the same point in both images, $P\mathbf{X} \sim P'\mathbf{X}$. The horopter is a twisted cubic, passing through the two camera centers.

In the oriented projective geometry, two identical points satisfy $\mathbf{x} \overset{+}{\sim} \mathbf{x}'$ rather than just $\mathbf{x} \sim \mathbf{x}'$. The consequence is that only a *segment* of the horopter projects to the identical points. This segment is delimited by the camera centers and projects to the segment of

the conic F_s delimited by the epipoles, as given by the following theorem.

Theorem 5 *Let* (P, P') *be camera matrices and* F *fundamental matrix consistent with them such that* $|F + F^\top| \neq 0$. *Then points* $\mathbf{e} = PP'^{\wedge}$, $\mathbf{e}' = P'P^{\wedge}$ *satisfying* $F\mathbf{e} = F^\top \mathbf{e}' = \mathbf{0}$ *divide conic* $F + F^\top$ *into two segments such that one segment is the image of points* $\mathbf{X}$ *satisfying* $P\mathbf{X} \stackrel{+}{\sim} P'\mathbf{X}$ *and the other segment is the image of points satisfying* $P\mathbf{X} \stackrel{+}{\sim} -P'\mathbf{X}$.

Proof. Points $\mathbf{x}$ on conic F_s satisfy $\delta \mathbf{e}' \times \mathbf{x} = F\mathbf{x}$ for some $\delta \neq 0$. It follows from (1) that $P\mathbf{X} \stackrel{+}{\sim} P'\mathbf{X}$ for $\sigma > 0$ and $P\mathbf{X} \stackrel{+}{\sim} -P'\mathbf{X}$ for $\sigma < 0$. We have to show that δ changes sign when and only when $\mathbf{x}$ crosses the line joining $\mathbf{e}$ and $\mathbf{e}'$. Multiplying the equation by $\mathbf{e}^\top$ from the left gives $\delta |\mathbf{e}, \mathbf{e}', \mathbf{x}| = \mathbf{e}^\top F\mathbf{x}$. Since $2F_s = F + F^\top$ it is $\mathbf{e}^\top F = 2\mathbf{e}^\top F_s$, therefore $\delta |\mathbf{e}, \mathbf{e}', \mathbf{x}| = 2\mathbf{e}^\top F_s \mathbf{x}$. $|\mathbf{e}, \mathbf{e}', \mathbf{x}|$ changes sign when and only when $\mathbf{x}$ crosses the line joining $\mathbf{e}$ and $\mathbf{e}'$. But $\mathbf{e}^\top F_s \mathbf{x}$ has the same sign for all points $\mathbf{x} \not\sim \mathbf{e}$ on the conic because $F_s \mathbf{x}$ is the tangent at $\mathbf{x}$ and $\mathbf{e}$ lies on the conic. $\square$

8 Application to Wide Baseline Guided Matching

Guided matching is a process used to find additional matches based on estimated epipolar geometry. The current estimate of F restricts the search from entire images to bands around corresponding epipolar lines. Similarity measure is used to select between matches that satisfy the epipolar constraint. In case of short baseline stereo, the correlation of intensities in square neighbourhoods of the interest points is typically used as a similarity measure [3]. In wide baseline stereo, such similarity measure cannot be used since the local image to image mapping is not restricted to a translation as in the short baseline case, but may be a general affine transformation.

Many wide baseline matching algorithms use region-to-region correspondences [13, 5, 6]. Consider now two regions Ω and Ω', each in different image, that are linked by affine transformation A, $|A| > 0$. Let the centroids of the regions be $\bar{\mathbf{x}} = [\bar{x}, \bar{y}, 1]^\top$ and $\bar{\mathbf{x}}' = [\bar{x}', \bar{y}', 1]^\top$ respectively. Let the covariance two-by-two matrices of the regions be Σ and Σ' respectively:

$$\Sigma = \frac{1}{|\Omega|} \int_\Omega \begin{bmatrix} x - \bar{x} \\ y - \bar{y} \end{bmatrix} [x - \bar{x}, \ y - \bar{y}] \, \mathrm{d}x \mathrm{d}y. \tag{9}$$

The representation of regions by their centroids and covariance matrices is useful, since it can be derived from regions of any shape. Both centroids and covariance matrices are transformed covariantly under any affine transformation. The affine transformation A that links regions Ω and Ω' can be written in form $A = T'^{-1}RT$. The matrix R represents image rotation,

$$T = \begin{bmatrix} \sqrt{\Sigma} & \bar{x} \\ & \bar{y} \\ 0 \ 0 & 1 \end{bmatrix}^{-1} \quad \text{and} \quad T' = \begin{bmatrix} \sqrt{\Sigma'} & \bar{x}' \\ & \bar{y}' \\ 0 \ 0 & 1 \end{bmatrix}^{-1}, \tag{10}$$

where $\sqrt{\Sigma}^\top \sqrt{\Sigma} = \Sigma$ is obtained by the Cholesky decomposition. Transformations T and T' transform the first (respectively the second) region so that the transformed regions have their centroid in the origin and have the identity covariance matrix. Such transformation is called normalisation, see figure 3, and can be obtained directly from the centroid and covariance matrix of single region. Normalisation is defined up to unknown rotation, i.e. transformations T and T' are ones of possible normalisations. The rotation R is not

Figure 3: Picture of the Little Mole (left) and his normalised picture (middle). The covariance matrices are depicted as ellipses. The decomposition of affine transformation into two normalisations T, T′, *and image rotation* R *(right).*

defined by the covariance matrices. The decomposition of A is depicted on figure 3. The covariance matrices are depicted as ellipses.

The affine transformation linking two regions in wide baseline stereo matching problem is not known beforehand. Hence, to establish a tentative correspondence of regions an affine invariant descriptors [13, 5, 6] and/or correlation over all possible rotations R after the normalisation [13, 5] may be used. The first approach is invariant to image rotation R and the second finds R so that similarity (correlation) is maximised.

Once an estimate of the epipolar geometry is known (e.g. after robust estimation using RANSAC), neither the affine transformation A nor the rotation R can be arbitrary. In fact, the fundamental matrix F defines, under two reasonable assumptions, the rotation R, and hence A, uniquely. The first assumption is that both image coordinate bases have the same handedness and was already discussed in section 4. The second assumption is that the observed surface is opaque and cannot be seen from behind. Both assumptions guarantee the absence of 'mirroring' effect. The rotation R is uniquely defined by the following equation

$$[\mathbf{Te}]_\times [0,0,1]^\top \overset{+}{\sim} -\mathtt{R}^\top [\mathtt{T}'\mathbf{e}']_\times [0,0,1]^\top. \tag{11}$$

Equation (11) chooses the image rotation R so that the affine transformation A maps the epipolar line $[\mathbf{e}']_\times \mathbf{x}$ to $[\mathbf{e}']_\times \mathbf{x}'$ with proper orientation, i.e. $\mathtt{A}^{-\top}[\mathbf{e}]_\times \mathbf{x} \overset{+}{\sim} -[\mathbf{e}']_\times \mathbf{x}'$. The essential condition is, that the epipoles $\mathbf{e}$ and $\mathbf{e}'$ have correct joint orientation, i.e., were obtained using (7).

The described method not only allows for the search for further matches, it also enables the rejection of false inliers, i.e. region matches that both lie on the corresponding epipolar lines and can be locally mapped to each other, but the mapping is not consistent with the global epipolar geometry, as depicted on figure 4.

9 Conclusions

In this paper we showed that *fundamental matrix alone encodes joint orientation of epipoles*, i.e., an unoriented fundamental matrix and one oriented epipole uniquely determines the orientation of the other epipole. The joint orientation of epipoles also determines the mutual position of cameras (figure 1). The unoriented concept of the Steiner conic was extended to oriented epipolar geometry. The application of wide baseline stereo guided matching was introduced, exploiting the correct joint orientation of epipoles.
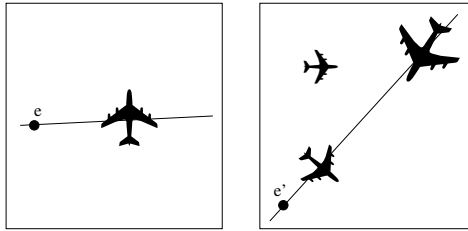
Figure 4: Which plane in the second image can correspond to the one in the first image? The smallest plane is eliminated, as it does not lie on corresponding epipolar line. And so is the biggest one because of its incorrect rotation w.r.t. the epipolar line.

References

[1] R. Hartley, L. de Agapito, I. Reid, and E. Hayman. Camera calibration and the search for infinity. volume 1, pages 510–515, Los Alamitos, California, September 1999. IEEE Computer Society.

[2] R. I. Hartley. Chirality. *Int. Jour. Computer Vision IJCV*, 26(1):41–61, 1998.

[3] R. I. Hartley and A. Zisserman. *Multiple View Geometry in Computer Vision*. Cambridge University Press, ISBN: 0521623049, 2000.

[4] S. Laveau and O. Faugeras. Oriented projective geometry for computer vision. In *ECCV96*, pages I:147–156, 1996.

[5] J. Matas, O. Chum, M. Urban, and T. Pajdla. Robust wide baseline stereo from maximally stable extremal regions. In *Proc. of BMVC*, volume 1, pages 384–393, London, UK, September 2002. BMVA.

[6] K. Mikolajczyk and C. Schmid. An affine invariant interest point detector. In *Proc. ECCV*, volume 1, pages 128–142, 2002.

[7] D. Nister. Calibration with robust use of cheirality by quasi-affine reconstruction of the set of camera projection centres. In *Proc. Intl. Conf. Computer Vision ICCV01*, pages II: 116–123, 2001.

[8] T. Pajdla, T. Svoboda, and V. Hlaváč. Epipolar geometry of central panoramic cameras. In R. Benosman and S. B. Kang, editors, *Panoramic Vision : Sensors, Theory, and Applications*. Springer Verlag, Berlin, Germany, 2000.

[9] F. Schaffalitzky and A. Zisserman. Viewpoint invariant texture matching and wide baseline stereo. In *Proc. 8th ICCV, Vancouver, Canada*, July 2001.

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

[11] J. Stolfi. *Oriented Projective Geometry: A Framework for Geometric Computations*. Academic Press, Inc., 1250 Sixth Avenue, San Diego, CA 92101, 1991.

[12] B. Triggs. Matching constraints and the joint image. In *ICCV95*, pages 338–343, 1995. See also the full version on www, submitted to IJCV.

[13] T. Tuytelaars and L. Van Gool. Wide baseline stereo matching based on local, affinely invariant regions. In *Proc. 11th British Machine Vision Conference*, 2000.

[14] T. Werner. Constraint on five points in two images. In *Proc. Comp. Vision and Pattern Recognition Conf. CVPR03*, 2003.

[15] T. Werner and T. Pajdla. Oriented matching constraints. In T. Cootes and C. Taylor, editors, *British Machine Vision Conference 2001*, pages 441–450, London, UK, September 2001. British Machine Vision Association.