HAL
open science

Epipolar Constraints for Multiscale Matching

William Triggs, Pashmina Bendale

To cite this version:

William Triggs, Pashmina Bendale. Epipolar Constraints for Multiscale Matching. BMVC 2010 - British Machine Vision Conference, Aug 2010, Aberystwyth, United Kingdom. pp.48.1-48.10, 10.5244/C.24.48 . hal-00565018

HAL Id: hal-00565018

https://hal.science/hal-00565018

Submitted on 10 Feb 2011

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Epipolar Constraints for Multiscale Matching

Bill Triggs
Bill.Triggs@imag.fr

Laboratoire Jean Kuntzmann
BP 53, 38041 Grenoble Cedex 9
France

Pashmina Bendale
pb397@cam.ac.uk

Signal Processing Laboratory
Department of Engineering
Cambridge CB2 1PZ, UK

Abstract

Many recent keypoint detectors associate a local scale (for multiscale detectors) or even a full affine frame (for affine-invariant detectors) to each detected keypoint. Although conventional epipolar constraints are a powerful tool for matching point-like features between pairs of images, they provide no constraint on their relative scales. We present an enhanced epipolar constraint that exploits these scales, thus providing more accurate correspondence search. The method encodes multiscale keypoints as image ellipses, invokes the classical Kruppa constraints that link corresponding ellipses, reduces these to constraints on 1-D quadratic forms on the pencil (1-D family) of epipolar lines, and enforces a scale-sensitive error model by a well-chosen algebraic transformation of the resulting homogeneous representation. The required projections onto the epipolar pencil are extracted from the Singular Value Decomposition of the Fundamental matrix. The final method is very simple to use. Illustrative tests yielded 2–4 fold reductions in false matches for both synthetic and real images. Matlab code is available.

1 Introduction

The conventional epipolar constraint is a powerful tool for matching keypoints (salient local features) between pairs of images, but in its standard form it treats the detected keypoints as point-like entities without intrinsic scales. In contrast, recent keypoint detectors (*c.f. e.g.* [8, 9] and their references) typically associate a scale (for multiscale detectors) or a full affine frame (for affine-invariant detectors) to each detection. In this paper we reformulate the epipolar constraint in a way that allows this additional scale information to be brought into play, thus providing a more selective overall matching process.

Specifically, we suppose that each detected keypoint can be described in terms of a circular or elliptical image region that characterizes the point's position, scale and shape, and that between pairs of images the ellipses of corresponding pairs of keypoints are in approximate correspondence, *i.e.* consistent with one another under the inter-image epipolar geometry. This will allow us to apply the classical Kruppa (conic correspondence) constraints[1]. We

[1] Note that these constraints do not control the feature scales themselves, but rather the way that they must change (for true correspondences) as we move away from the epipole along the epipolar line.

Work supported by a PhD grant from the Gates Cambridge Trust and by travel grants from the European network PASCAL2. MATLAB code is available on http://ljk.imag.fr/membres/Bill.Triggs/src.

© 2010. The copyright of this document resides with its authors.
It may be distributed unchanged freely in print or electronic forms.

will convert these to a simple angular 'epipolar pencil' representation that is convenient for matching, and construct an appropriate correspondence error metric in terms of this.

Before starting, we make several caveats. Firstly, note that our method is purely geometric. We treat ellipses as featureless geometric primitives, taking no account of the finer issue of establishing detailed pointwise correspondence between their underlying image content. Our approach is thus complementary to appearance-based correspondence methods such as multiscale or affine visual descriptors (*c.f. e.g.* [8, 10, 13] and their references). Both approaches provide useful constraints on correspondence and one should use whichever is most convenient for the problem in hand (and, when possible, combine them). In particular, if two keypoints correspond to the same salient region of a locally smooth surface in 3D, one would expect their ellipses to correspond under both our geometric constraints and appearance-based ones.

Secondly, throughout this paper we assume that the epipolar geometry (Fundamental matrix) is known. Any standard method can be used to estimate it [5]. We build an error model for uncertain multiscale points, not one for uncertain epipolar geometries.

Finally, note that even for fixed-scale keypoints, one can often model the uncertainties in their estimated positions with associated uncertainty ellipses. This is entirely different from the above use of ellipses to model associated image regions: although the resulting keypoint positions should still correspond modulo the uncertainties, there is no particular reason to expect their uncertainty ellipses to correspond because a keypoint may be localized well in one image and poorly in the other. Although it is not our focus here, our framework can handle this situation simply by omitting (13), the error term penalizing scale mismatch.

Prior Work. There has been a large amount of work on multiscale keypoint detectors and descriptors – *e.g.* [8, 9, 10, 12, 13], to give only a few examples. The use of epipolar geometry for inter-image keypoint matching and dense stereo is very well established, and the closely related 'Kruppa constraints' for correspondence between conics are also well known [3, 5, 6]. However there seems to be surprisingly little work dealing explicitly with integrating scales into epipolar geometry, perhaps because the authors interested in multiscale detection have tended to take descriptor-centric approaches not geometry-centric ones. The closest work that we are aware of is Forssén & Moe [4]. This uses elliptical blob features to estimate and apply epipolar geometry, but in a less unified way than the method given here.

2 Derivation of the Model

We now derive our model for multiscale keypoint correspondence via ellipses. The derivation takes some time but the final model is straightforward to use and is summarized at the end. We assume familiarity with the standard projective formulation of vision geometry (image projection, epipolar geometry, conics, *etc.*). For details see *e.g.* [5]. '$\sim$' denotes equality up to scale, and for a 3-vector $\mathbf{v}$, '$[\mathbf{v}]_\times$' denotes the corresponding 3×3 skew "cross product" matrix.

Epipolar Pencil Coordinates. Before starting on ellipses, we reexpress some basic facts about epipolar geometry in a useful form (*c.f. e.g.* [5] §18.4 'Kruppa Equations'). Let $\mathbf{P}, \mathbf{P}'$ be the 3×4 perspective projection matrices of two images, called respectively 'left' and 'right' below. Let $\mathbf{F}$ be the corresponding fundamental matrix – $\mathbf{x}^\top \mathbf{F} \mathbf{x}' = 0$ is the epipolar constraint between left image point $\mathbf{x}$ and right image point $\mathbf{x}'$ – and let $\mathbf{e}$ and $\mathbf{e}'$ be the corresponding epipoles. From the Singular Value Decomposition (SVD) of $\mathbf{F}$, we can extract

the following representation

$$\mathbf{F} = \mathbf{U}\mathbf{S}\mathbf{V}^\top = \begin{pmatrix} -\mathbf{v} & \mathbf{u} & \mathbf{e} \end{pmatrix} \begin{pmatrix} \mu & & \\ & \nu & \\ & & 0 \end{pmatrix} \begin{pmatrix} \mathbf{u}' & \mathbf{v}' & \mathbf{e}' \end{pmatrix}^\top = \mathbf{u}\,\nu\,\mathbf{v}'^\top - \mathbf{v}\,\mu\,\mathbf{u}'^\top \tag{1}$$

Here, the 'twisted' naming in the U matrix of the SVD ensures that when viewed as points, the pairs $(\mathbf{u},\mathbf{u}')$ and $(\mathbf{v},\mathbf{v}')$ are in epipolar correspondence: $\mathbf{u}^\top\mathbf{F}\mathbf{u}' = 0 = \mathbf{v}^\top\mathbf{F}\mathbf{v}'$. Moreover, viewed as line-vectors, $\{\mathbf{u},\mathbf{v}\}$ and $\{\mathbf{u}',\mathbf{v}'\}$ respectively form bases for the pencils (1-D linear families) of epipolar lines in the left and right images, and $(\mathbf{u},\mathbf{u}')$ and $(\mathbf{v},\mathbf{v}')$ are corresponding pairs of epipolar lines: right point-vector $\mathbf{v}'$ lies on the right epipolar line $\mathbf{u}'$ (as $\mathbf{u}'^\top\mathbf{v}'$=0) and has left epipolar line $\mathbf{F}\mathbf{v}' \sim \mathbf{u}$, and vice versa. Finally, the left epipole $\mathbf{e}$ lies on the lines $\mathbf{u}$ and $\mathbf{v}$ and we can choose its sign so that $[\mathbf{e}]_\times = \mathbf{u}\mathbf{v}^\top - \mathbf{v}\mathbf{u}^\top$.

To make this more concrete, define the 2×3 "epipolar pencil projection" matrices[2]

$$\mathbf{B} \equiv \begin{pmatrix} \mathbf{u} & \mathbf{v} \end{pmatrix}^\top, \qquad \mathbf{B}' \equiv \begin{pmatrix} \mu\,\mathbf{u}' & \nu\,\mathbf{v}' \end{pmatrix}^\top \tag{2}$$

and note that

$$\mathbf{F} = \mathbf{B}^\top \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \mathbf{B}', \qquad [\mathbf{e}]_\times = \mathbf{u}\mathbf{v}^\top - \mathbf{v}\mathbf{u}^\top = \mathbf{B}^\top \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \mathbf{B}, \qquad \mathbf{B}\mathbf{B}^\top = \mathbf{I}_{2\times 2} \tag{3}$$

It follows that points $\mathbf{x},\mathbf{x}'$ are in epipolar correspondence, $\mathbf{x}^\top\mathbf{F}\mathbf{x}' = \mathbf{x}^\top\mathbf{B}^\top \left(\begin{smallmatrix} 0 & 1 \\ -1 & 0 \end{smallmatrix}\right) \mathbf{B}'\mathbf{x}' = 0$, if and only if the 2-vectors $\left(\begin{smallmatrix} \alpha \\ \beta \end{smallmatrix}\right) = \mathbf{B}\mathbf{x}$ and $\left(\begin{smallmatrix} \alpha' \\ \beta' \end{smallmatrix}\right) = \mathbf{B}'\mathbf{x}'$ are equal up to scale. We can view $(\alpha,\beta)^\top, (\alpha',\beta')^\top$ as homogeneous 2D coordinates for the 'projections' of $\mathbf{x},\mathbf{x}'$ onto the left epipolar pencil under $\mathbf{B},\mathbf{B}'$, with epipolar correspondence if and only if the projections agree up to scale. Moreover, $\mathbf{x}'$ has left epipolar line $\beta'\mathbf{u} - \alpha'\mathbf{v}$, and $\mathbf{x} = \alpha\mathbf{u} + \beta\mathbf{v} + \gamma\mathbf{e}$ lies on left epipolar line $\beta\mathbf{u} - \alpha\mathbf{v}$, so the corresponding epipolar lines have the dual coordinates $(\beta,-\alpha)$ and $(\beta',-\alpha')$. Thus, the projections $\mathbf{B},\mathbf{B}'$ allow us to reduce image-based epipolar geometry calculations to 1D epipolar pencil ones. We will apply this to conics below. The appendix shows how to extend this to spherical images and signed epipolar geometry.

Conic Correspondence on the Epipolar Pencil. The geometry of projections and epipolar constraints for quadrics is well known [5] – see fig. 1. Let Q be a 3D quadric – here typically a (possibly planar) ellipsoid. Its images in the two cameras are 2D conics q,q' representing the envelopes of the image rays tangent to Q. Exactly two epipolar planes are tangent to Q in 3D, and in each image these generate a pair of corresponding epipolar lines that are tangent to the conics q,q'. Conversely, given any two image conics q,q' with corresponding epipolar lines, it turns out that there is a 3D quadric Q (in fact a 1 parameter family of them) whose images are q and q'. The epipolar constraint between conics is thus the requirement that the two epipolar lines tangent to q are in epipolar correspondence with the two epipolar lines tangent to q'.

In formulae, if we represent Q in dual (hyperplane) form by 4×4 symmetric matrix $\mathbf{Q}$ and q,q' in dual (line) form by 3×3 symmetric matrices $\mathbf{q},\mathbf{q}'$, the image projections are simply

[2] Well-normalized image coordinates should be used to evaluate the $\mathbf{B}$'s: $\mathbf{F} \to \mathbf{K}_1^{-\top}\mathbf{F}\mathbf{K}_2^{-1}$ on input to the SVD and $\mathbf{B}_1 \to \mathbf{B}_1\mathbf{K}_1$, $\mathbf{B}_2 \to \mathbf{B}_2\mathbf{K}_2$ on output, where $\mathbf{K}_i \sim \begin{pmatrix} 1/f_x & & -p_x/f_x \\ & 1/f_y & -p_y/f_y \\ & & 1 \end{pmatrix}$ are nominal calibration matrices with "focal lengths" / normalization scales f_x, f_y (in pixels) and "principal point" / image centre $(p_x,p_y)^\top$. The exact values used are not critical, but if unnormalized pixel coordinates are used the resulting epipolar pencil coordinates tend to be very distorted, leading to unintuitive behaviour in the below angular error estimates.

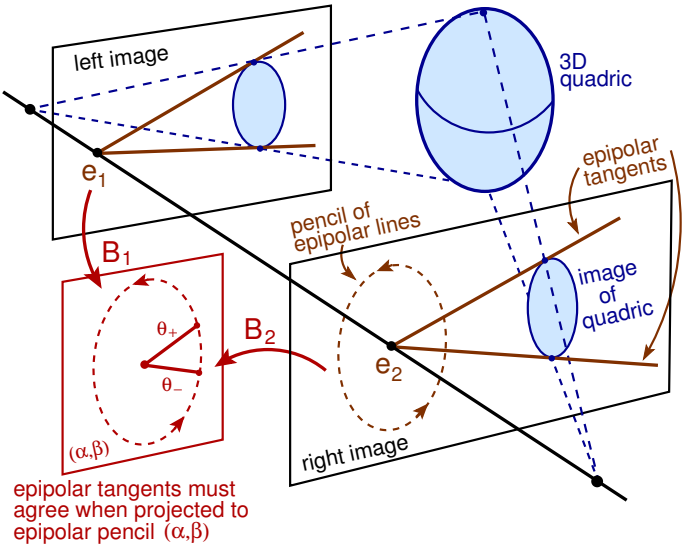

Figure 1: Projection of a 3D quadric to two image conics, and then projection via $\mathbf{B},\mathbf{B}'$ of the pair of epipolar lines tangent to each conic (or equivalently, of the quadratic form defined by the conic) to the epipolar pencil. The reduced Kruppa constraints state that the two projections agree.

$\mathbf{q} \sim \mathbf{P}\mathbf{Q}\mathbf{P}^\top$ and $\mathbf{q}' \sim \mathbf{P}'\mathbf{Q}\mathbf{P}'^\top$ [11]. More explicitly, the dual form of an image ellipse with 2D centre $\boldsymbol{c}$ and 2D covariance matrix $\boldsymbol{V}$ is

$$\mathbf{q} = \begin{pmatrix} \boldsymbol{c}\boldsymbol{c}^\top - \boldsymbol{V} & \boldsymbol{c} \\ \boldsymbol{c}^\top & 1 \end{pmatrix} \tag{4}$$

so an image line $\boldsymbol{n}^\top \binom{x}{y} - d = 0$ is tangent to q if and only if $\begin{pmatrix} \boldsymbol{n}^\top & -d \end{pmatrix} \mathbf{q} \binom{\boldsymbol{n}}{-d} = (\boldsymbol{n}^\top \boldsymbol{c} - d)^2 - \boldsymbol{n}'^\top \boldsymbol{V}\boldsymbol{n} = 0$. For a circle, $\boldsymbol{V} = r^2 \boldsymbol{I}$, the tangency equation becomes $(\boldsymbol{n}^\top \boldsymbol{c} - d)^2 = r^2$, *i.e.* any line that passes at distance r from the centre $\boldsymbol{c}$ is tangent to the circle. Analogous forms hold for the 3D quadric $\mathbf{Q}$ and its tangent planes, with 3D $\boldsymbol{c},\boldsymbol{V},\boldsymbol{n}$.

In this framework, conic correspondence turns out to be encapsulated by the 'Kruppa constraints'[3] [3, 5]

$$[\mathbf{e}]_\times \, \mathbf{q} \, [\mathbf{e}]_\times^\top \;\sim\; \mathbf{F}\mathbf{q}'\mathbf{F}^\top \tag{5}$$

Each side of this equation is a 3×3 symmetric rank 2 matrix of the form $\mathbf{l}_1\mathbf{l}_2^\top + \mathbf{l}_2\mathbf{l}_1^\top$ where $\mathbf{l}_1,\mathbf{l}_2$ are the left epipolar lines tangent q or q', as appropriate. Applying (3) and cancelling extraneous factors of $\mathbf{B}^\top \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$ gives the 2×2 symmetric 'Reduced Kruppa Constraints'[4] [5]

$$\mathbf{B}\mathbf{q}\mathbf{B}^\top \sim \mathbf{B}'\mathbf{q}'\mathbf{B}'^\top, \qquad \text{or in coordinate form} \qquad \begin{pmatrix} q_{vv} \\ q_{uv} \\ q_{uu} \end{pmatrix} \sim \begin{pmatrix} \nu^2 q'_{vv} \\ \mu\nu q'_{uv} \\ \mu^2 q'_{uu} \end{pmatrix} \tag{6}$$

[3] An alternative form is that conics correspond iff the symmetric part of $[\mathbf{e}]_\times \mathbf{q}\mathbf{F}\mathbf{q}'\mathbf{F}^\top$ vanishes. Under correspondence this matrix is $\sim [\mathbf{e}]_\times$. *C.f.* a homography $\mathbf{H}$ is consistent with $\mathbf{F}$ iff $\mathbf{F}\mathbf{H}$ is skew (and $\sim [e]_\times$).

[4] Kruppa constraints were originally developed for camera autocalibration [3, 5]. In this context, $\mathbf{q},\mathbf{q}'$ are chosen to be the "images of the dual absolute quadric" $\mathbf{q} = \mathbf{K}\mathbf{K}^\top$, $\mathbf{q}' = \mathbf{K}'\mathbf{K}'^\top$, where $\mathbf{K},\mathbf{K}'$ are the camera internal parameter matrices. Their reduced quadratic (6) determines the metric circle structure ("circular points") of (α,β) – *i.e.* which pairs of epipolar lines correspond to orthogonal pairs of 3D epipolar planes.

where the coordinates are $q_{uv} \equiv \mathbf{u}^\top \mathbf{q}\, \mathbf{v}$, $q'_{uu} \equiv \mathbf{u}'^\top \mathbf{q}'\, \mathbf{u}'$, *etc*. These equations provide two independent algebraic constraints representing the two epipolar tangencies. Below we will use $(a,b,c)^\top$ to denote either the left or right hand side of (6), as needed.

Evaluating the dual conic $\mathbf{q}$ on the epipolar line vector corresponding to 'point' $(\alpha,\beta)^\top$ induces a corresponding reduced quadratic form on the pencil

$$\begin{pmatrix}\beta & -\alpha\end{pmatrix} \mathbf{B}\,\mathbf{q}\,\mathbf{B}^\top \begin{pmatrix}\beta\\ -\alpha\end{pmatrix} \;=\; a\,\alpha^2 - 2\,b\,\alpha\beta + c\beta^2 \tag{7}$$

and similarly for $\mathbf{q}'$ with $\mathbf{B}'$. For example, inserting (4) into (7) gives $(\alpha_c\,\beta - \beta_c\,\alpha)^2 - \begin{pmatrix}\beta & -\alpha\end{pmatrix} \tilde{\boldsymbol{V}} \begin{pmatrix}\beta\\ -\alpha\end{pmatrix}$ where $\begin{pmatrix}\alpha_c\\ \beta_c\end{pmatrix} \equiv \mathbf{B}\begin{pmatrix}\boldsymbol{c}\\ 1\end{pmatrix}$ and $\tilde{\boldsymbol{V}} \equiv \mathbf{B}\begin{pmatrix}\boldsymbol{V} & \boldsymbol{0}\\ \boldsymbol{0} & 0\end{pmatrix}\mathbf{B}^\top$. If $\boldsymbol{V} \to \boldsymbol{0}$ this has a double root at $(\alpha_c,\beta_c)^\top$ as expected.

Algebraic Weakness of Kruppa Form. To gain further intuition, we briefly switch to angular coordinates $(\alpha,\beta)^\top \sim (\cos\theta,\sin\theta)^\top$, letting the roots of the reduced quadratic be[5] $\theta_\pm = \bar\theta \pm \delta\theta$. Then up to scale, (7) becomes

$$\sin(\theta-\theta_+)\,\sin(\theta-\theta_-) \;=\; (\sin\theta\cos\theta_+ - \cos\theta\sin\theta_+)(\sin\theta\cos\theta_- - \cos\theta\sin\theta_-) \tag{8}$$

$$\;=\; \big(\cos 2\delta\theta - \cos 2(\theta-\bar\theta)\big)/2 \tag{9}$$

where the various forms follow from standard trigonometric identities, and

$$\begin{pmatrix}a\\ b\\ c\end{pmatrix} \sim \begin{pmatrix}\sin\theta_+\sin\theta_-\\ (\cos\theta_+\sin\theta_- + \sin\theta_+\cos\theta_-)/2\\ \cos\theta_+\cos\theta_-\end{pmatrix} = \frac{1}{2}\begin{pmatrix}\cos 2\delta\theta - \cos 2\bar\theta\\ \sin 2\bar\theta\\ \cos 2\delta\theta + \cos 2\bar\theta\end{pmatrix} \tag{10}$$

$$\begin{pmatrix}p\\ q\\ r\end{pmatrix} \equiv \begin{pmatrix}c-a\\ 2b\\ c+a\end{pmatrix} \sim \begin{pmatrix}\cos 2\bar\theta\\ \sin 2\bar\theta\\ \cos 2\delta\theta\end{pmatrix} \tag{11}$$

We would like to use reduced Kruppa vectors $(a,b,c)^\top$ or $(p,q,r)^\top$ to quantify the extent to which two ellipses are in epipolar correspondence. The $(\cos 2\bar\theta, \sin 2\bar\theta)$ components of (11) should provide good control over errors in the mean epipolar line direction $\bar\theta$. Unfortunately, the $\cos 2\delta\theta$ term provides much less control over $\delta\theta$. In particular, for ellipses that are small relative to their distance to the epipole, $\delta\theta$ is small and $\cos^2 2\delta\theta \approx 1 - 4\delta\theta^2$ provides only a weak second order constraint on it: the Kruppa constraints are algebraically complete but unsuitable as error models because they provide too little control over the sizes of the small ellipses that make up most of the keypoint population. One way around this would be to introduce a fourth coordinate $s \equiv \sqrt{p^2+q^2-r^2} \sim \sin 2\delta\theta$ into (11), to give better control of $\delta\theta$. We will embed an analogous transformation in our error metric. Henceforth we suppose that $(p,q,r)^\top$ has been rescaled to make $p^2+q^2=1$, so that (11) becomes an equality.

Epipolar Pencil Error Model. We need to design an error weighting that reflects our expectations regarding uncertainties in keypoint positions and scales. A great many models are possible. Here we develop just one as an example, based on the assumption that both the position uncertainty and the scale uncertainty of a typical keypoint are proportional to

[5] To ensure consistent signs we take $\theta_- \le \theta_+ < \theta_- + \pi$, *i.e.* $0 \le \delta\theta < \frac{\pi}{2}$. Also, we assume that the quadratics (7) have real roots ($b^2 \ge 4ac$, or $r^2 \le p^2+q^2 = 1$ below), *i.e.* that neither image ellipse contains (surrounds) its epipole. This is true for most keypoints. If not, we can either discard the point, or note that (5), (6) still provide valid constraints on the match – the geometric interpretation is now that the circular points implied on the line $(\alpha,\beta)^\top$ by the two quadratics must agree – and, *e.g.*, use $r-1$ in place of $1-r$ in (12), (13) below, noting that for a match of this kind, *both* quadratics must have $r > 1$.

its scale and that deviations are approximately Gaussian. To concretize this algebraically on the epipolar pencil we will use $\sigma \equiv \sqrt{(1-r)/2} = \sin\delta\theta$ as a surrogate for the ellipse scale, as this scales linearly with $\delta\theta$ at small angles and increases smoothly to 1 at very large ones $\delta\theta \approx \frac{\pi}{2}$ or $\theta_+ - \theta_- \approx \pi$. (Moreover, $-\sigma^2 = (r-1)/2$ is the minimum value of (8)). Assuming independent Gaussian errors of (small) standard deviation $\delta\theta$ in $\bar\theta$, the standard (two-sample z-test) statistic for deviations of the means is $(\bar\theta - \bar\theta')^2/(\delta\theta^2 + \delta\theta'^2)$. We will algebraize this as

$$d_{\bar\theta} \equiv \frac{\sin^2 2(\bar\theta - \bar\theta')}{\sin^2\delta\theta + \sin^2\delta\theta'} = \frac{(pq'-qp')^2}{\sigma^2+\sigma'^2} = \frac{(pq'-qp')^2}{1-(r+r')/2} \tag{12}$$

Various penalties (log-normal, *etc.*) are possible for the errors in $\delta\theta$, but one viable statistic is $\big(\delta\theta/\delta\theta'\big)^k + \big(\delta\theta'/\delta\theta\big)^k - 2$ for some k. This is symmetric, scale invariant, zero at equality and grows as $O(\delta\theta^{-k})$ when either $\delta\theta$ shrinks to zero with the other held constant. We will algebraize this as

$$d_{\delta\theta} \equiv \left(\frac{\sin\delta\theta}{\sin\delta\theta'}\right)^k + \left(\frac{\sin\delta\theta'}{\sin\delta\theta}\right)^k - 2 = \left(\frac{\sigma}{\sigma'}\right)^k + \left(\frac{\sigma'}{\sigma}\right)^k - 2 = \left(\frac{1-r}{1-r'}\right)^{k/2} + \left(\frac{1-r'}{1-r}\right)^{k/2} - 2 \tag{13}$$

Our final error model will be a weighted[6] sum of (12) and (13), and for simplicity we will use $k=2$ below. Although we will not consider them here, other more heuristic error models could also be used, for example weighting squared distances between $(p,q,r,s)^\top$ vectors with a suitable function such as $\frac{1}{\sigma^2+\sigma'^2}$, $\frac{1}{\sigma\sigma'}$ or $\frac{1}{\sigma^2} + \frac{1}{\sigma'^2}$.

Summary of Method. The final method is very straightforward. For each ellipse, the matrix $\mathbf{q}$ or $\mathbf{q}'$ (4) is projected using $\mathbf{B}$ or $\mathbf{B}'$ (from the SVD of $\mathbf{F}$) to obtain $(a,b,c)^\top$ (6), normalized to give $(p,q,r)^\top$ (11), and – see appendix (14) – optionally unwrapped to signed form. 'Distances' between these vectors are then computed using a weighted sum of (12) and (13) and used to decide whether pairs of ellipses might correspond.

Further Points. Here we opted for algebraic representations for simplicity, but it would also have been possible to explicitly recover and use $\bar\theta, \delta\theta$. In particular, if an efficient data structure for inlier search is required, one could represent search intervals as rectangles in $(\bar\theta, \delta\theta)$ coordinates and use some data structure such as a box tree that allows efficient search for all of the left image rectangles containing an observed pair $(\bar\theta', \delta\theta')$ from the right image. This would require cutting the $\bar\theta$ circle to a flat interval $[0, 2\pi]$ (so some points would generate two rectangles), and using the given $(\bar\theta, \delta\theta)$ value, the acceptance thresholds on (12)-(13) and, for (12), a bound on r': $r' \in [-1, 1]$, to derive search bounds on $(\bar\theta', \delta\theta')$.

The main limitation of the epipolar pencil representation is that it suppresses all information about positions along epipolar lines. If such information is useful[7], it must be applied separately from the epipolar pencil method. The above derivations also assume that $\mathbf{F}, \mathbf{B}, \mathbf{B}'$ are known exactly and we currently make no effort to incorporate uncertainties in these into the computations. Although the resulting effective geometric search regions (wedge-shaped ones starting at the epipoles) are actually statistically more correct and simpler to use than the common parallel strips along epipolar lines, in practice it is wise to allow a few additional pixels of slack to account for uncertainties in $\mathbf{F}$, especially for keypoints near the epipoles.

[6] *E.g.*, with the above weightings and for small $\delta\theta$, a factor-of-2 error in $\delta\theta$ has about the same penalty as a position error of $2\delta\theta$.

[7] For example, if the scene is bounded by known 3D half spaces or by the plane at infinity, their homographies can be used to limit the search for a given point's correspondent along its epipolar line.

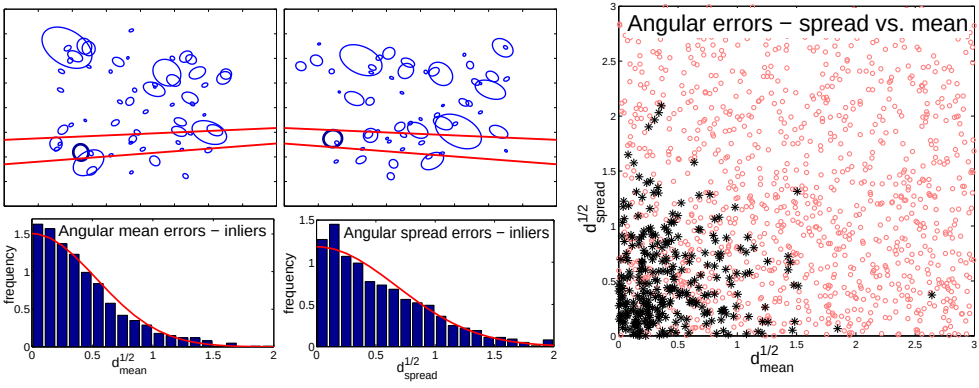

Figure 2: Experiments on synthetic data. Top left: left and right images of a scene containing random 3D ellipsoids, showing the epipolar lines tangent to a selected ellipse in the right image, and the corresponding lines in the left image almost tangent to the corresponding (noise perturbed) left ellipse. Bottom left: distributions of matching penalty values for correct but noise perturbed correspondences, for (left) the $\bar{\theta}$ (mean angle) penalty (12), (right) the $\delta\theta$ (angular spread) penalty (13). For each penalty d we histogram $\sqrt{d}$ (|linear error| rather than squared error) to show that the penalties behave roughly like χ^2_1 variables, *i.e.* the linear errors resemble half-Gaussians (red curves). Right: scatter plot of $\delta\theta$ penalty versus $\bar{\theta}$ penalty values over a large dataset. The black '*'s are correct matches and the red 'o's are incorrect ones. Again we plot linear errors $\sqrt{d}$. Clearly both the mean and the spread terms are useful for distinguishing inliers from outliers.

3 Experiments

We now describe some illustrative experiments with the method[8] on both synthetic data and a real image dataset. A more detailed study will be published later.

Synthetic Data. We generate artificial scenes consisting of N 3D ellipsoids with random centres in the cube $[-1,1]^3$, random scales distributed as s^{-2} in the interval $[0.005, 0.1]$, and random ellipticities with log-normal density of standard deviation 30%. These are viewed by two inwards-facing perspective cameras 4 units from the cube centre and 60° apart. The resulting image ellipses are perturbed in position and scale by Gaussian noise with standard deviation 33% of the ellipse radius. The ground truth epipolar geometry is used. Fig. 2 (top left) shows an example of the image pairs generated and their epipolar geometry.

With these settings we find that for true correspondences, the errors underlying the $\bar{\theta}$ (12) and $\delta\theta$ (13) penalty terms are approximately jointly Gaussian (see fig. 2, bottom left and right), so that the penalties $d_{\bar{\theta}}, d_{\delta\theta}$ themselves have scaled 1 d.o.f. χ^2 distributions. In contrast, the distribution of errors for incorrect matches is much broader and is approximately uniform near the origin. This implies that a near-optimal inlier-outlier decision rule is to threshold the χ^2_2 variable $d_{\bar{\theta}}/\mu_{\bar{\theta}} + d_{\delta\theta}/\mu_{\delta\theta}$, where $\mu_{\bar{\theta}}, \mu_{\delta\theta}$ are the empirical means of the penalty functions (*i.e.* the variances of the underlying errors) for true matches. At a fixed percentage of false rejections, we find that using this rule reduces the number of false positives by a factor of around 2 to 2.5 relative to classical epipolar thresholding based on $d_{\bar{\theta}}$ alone. This gain holds across a wide range of feature densities N, rejection percentages,

[8] A Matlab implementation is available from http://ljk.imag.fr/membres/Bill.Triggs/src.

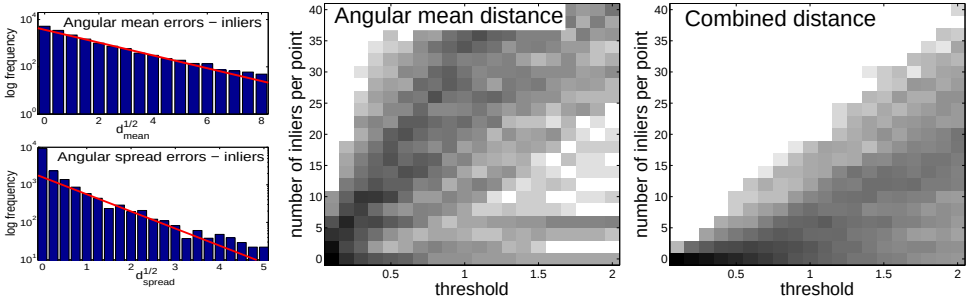

Figure 3: Experiments on real data. (Left) For SIFT interest points, the distributions of $\sqrt{d_{\bar\theta}}$ and $\sqrt{d_{\delta\theta}}$ for corresponding features appear to be exponential with medians $m_{\bar\theta} \approx 1.0$ and $m_{\delta\theta} \approx 0.6$ (*i.e.* scale lengths 1.5 and 0.9). (Middle) The histogram over an image pair of numbers of candidate matches satisfing the epipolar correspondence rule $\sqrt{d_{\bar\theta}}/m_{\bar\theta} < t$, for varying thresholds t. Darkness is proportional to log frequency. (Right) The corresponding histogram for the combined decision rule $\sqrt{d_{\bar\theta}}/m_{\bar\theta} + \sqrt{d_{\delta\theta}}/m_{\delta\theta} < t$, which is approximately optimal for independent exponential variables against a uniform background of outliers. The combined rule is about 4 times more selective, producing many fewer incorrect correspondences.

scene parameters, *etc*. It is increased by reduced uncertainty in, or broader distributions of, the ellipse scales, but we believe that our settings for these are representative of real detectors. For frontal camera motion (epipole at the image centre) the distribution of $\delta\theta$ values becomes somewhat broader and the gain is increased to around 4. Note that points with particularly large scales and ones that lie near the epipole are associated with broad sectors of epipolar lines, and therefore tend to match many other points under $d_{\bar\theta}$ alone. Adding $d_{\delta\theta}$ is particularly useful for eliminating these. On the other hand, there are typically many points with similar scales so $d_{\delta\theta}$ alone is not useful – it is only useful in combination with $d_{\bar\theta}$.

Real data. We also tested the method using SIFT interest points on a real dataset consisting of calibrated images of toy cars on a turntable[9] [1]. As a surrogate for ground-truth correspondences, we used conventional epipolar constraints over a triangle of images – 'Reference', 'Auxilliary' and 'Test' – selecting point pairs that corresponded both geometrically and by least squares patch matching in Reference and Auxilliary, and accepting any point within a generous region around the intersection of their two epipolar lines in Test as an inlier for the purposes of the evaluation. The resulting correspondences are far from perfect, but they suffice for an initial proof-of-concept test of the method on real data. Fig. 3 (left) shows that the distributions of $\sqrt{d_{\bar\theta}}$ and $\sqrt{d_{\delta\theta}}$ for such 'inliers' are approximately exponential (Cauchy-like), not Gaussian. A corresponding scatter plot (not shown) demonstrates that the two error metrics again provide very complementary information for correspondence search. Fig. 3 (middle) and (right) show that selecting possible correspondences by thresholding a weighted combination of the two metrics produces far fewer false matches than using epipolar line distances $d_{\bar\theta}$ alone. Similar conclusions are reached if the putative inliers for the tests are found using SIFT descriptor matching instead of 3-image epipolar constraints.

[9] http://www-sigproc.eng.cam.ac.uk/imu

4 Summary and Conclusions

We have introduced a framework for epipolar correspondence search that provides tighter constraints for matching multiscale keypoints by constraining the features' relative scales as well as their positions. The method works by representing uncertain keypoints as image ellipses, using 2×3 projection matrices $\mathbf{B},\mathbf{B}'$ extracted from the fundamental matrix $\mathbf{F}$ to project these onto the epipolar pencil – *i.e.* w.r.t. $(\alpha,\beta)^\top$ coordinates on the 1-D family of left epipolar lines – and formulating the quality of the match in terms of an algebraic error model in these coordinates. The method is elegant, simple to use (probably simpler than the traditional image search along epipolar strips) and gives a substantial reduction in false correspondences – typically a factor of 2–4 in our synthetic and real experiments.

5 Appendix: Spherical Images and Unwrapping

This appendix extends the above epipolar pencil constructions to the 'spherical' (also called 'oriented' or 'signed') approach to projective vision in which image points are identified with points on the camera's viewing sphere (the sphere of incoming 3D visual rays at its centre) [2, 7]. In this framework, epipolar lines correspond to *half* great circles joining the epipole to its antipode (*c.f.* the lines of longitude between the north and south poles of the earth): given the 3D line joining the two camera centres, take the circle of 3D half planes with this line as edge and project each half plane (on edge) into the two images making its two epipolar half-circles; the image of any 3D point on the half plane lies on both half-circles. In terms of a flat image, there is thus a full circle of distinct epipolar half-lines, and for any given point $\mathbf{x}'$, only the correct half of its epipolar line (to the left of the epipole or to the right of it) need be searched for its correspondent $\mathbf{x}$. Hence, this framework provides slightly sharper epipolar constraints whenever the epipole is in the image (notably for omnidirectional cameras).

Algebraically, with appropriate choices of signs, we can write this as $\mathbf{F}\mathbf{x}' \sim [e]_\times \mathbf{x}$ with equality up to positive rescalings. Using (3) and cancelling a common factor of $\mathbf{B}^\top \left(\begin{smallmatrix} 0 & 1 \\ -1 & 0 \end{smallmatrix}\right)$, we find that $\mathbf{B}'\mathbf{x}' \sim \mathbf{B}\mathbf{x}$, again with equality up to positive rescalings. So the epipolar pencil coordinates $(\alpha,\beta)^\top$ fully support the circular structure: the coordinates of corresponding points coincide up to *positive* rescalings and we can view them as angular coordinates $(\cos\theta,\sin\theta)^\top$ with $0 \le \theta < 2\pi$ and correspondence possible if and only if $\theta = \theta'$. The necessary signs for $\mathbf{B},\mathbf{B}'$ can not be recovered from $\mathbf{F}$ alone[10] so the easiest approach is to take a pair of corresponding spherical points, check that their pencil coordinates correspond, and if not flip the sign of $\mathbf{B}$. This will suffice for us here, although further refinements to ensure right-handedness, put $\mathbf{u},\mathbf{v}$ in canonical positions, *etc.*, are possible.

We can extend this to epipolar pencil conic matching. The basic formulae for conics and Kruppa constraints are intrinsically unsigned because converting $\left(\begin{smallmatrix} c \\ 1 \end{smallmatrix}\right)$ to its antipode $\left(\begin{smallmatrix} -c \\ -1 \end{smallmatrix}\right)$ leaves $\mathbf{q}$ unchanged in (4). This is reflected in the systematic appearance of $2\bar\theta$ in (10), *etc.* To correct for this we can algebraically 'unwrap' the first two rows of (11). Using multiple angle formulae and assuming normalization to $p^2+q^2=1$, we have

$$\begin{pmatrix} \cos\bar\theta \\ \sin\bar\theta \end{pmatrix} = \pm\begin{pmatrix} \sqrt{(1+p)/2} \\ q/\sqrt{2(1+p)} \end{pmatrix} = \pm\begin{pmatrix} q/\sqrt{2(1-p)} \\ \sqrt{(1-p)/2} \end{pmatrix} \tag{14}$$

[10] $\mathbf{F}$ itself can usually only be recovered up to sign, and even then the SVD of $\mathbf{F}$ is invariant under $(\mathbf{u},\mathbf{v}') \leftrightarrow (-\mathbf{u},-\mathbf{v}')$, $(\mathbf{v},\mathbf{u}') \leftrightarrow (-\mathbf{v},-\mathbf{u}')$, $\mathbf{e} \leftrightarrow -\mathbf{e}$, and $\mathbf{e}' \leftrightarrow -\mathbf{e}'$, some of which affect $\mathbf{B},\mathbf{B}'$.

where for numerical stability the first form is preferred if $p > 0$ and the second if $p < 0$. To choose the sign we take any signed point within (the desired branch of) the ellipse – *e.g.* its centre $\binom{c}{1}$ – find its pencil projection $\mathbf{Bx}$ or $\mathbf{B}'\mathbf{x}' \sim \binom{\cos\theta}{\sin\theta}$, choose the sign that aligns (14) the best with this direction, and replace the first two coordinates of (11) with the result (14) so that the feature vector encodes the desired half of the epipolar line. The right hand side of (12) is unchanged but it now represents only $\sin^2(\bar{\theta} - \bar{\theta}')$ so it should be scaled up by 4 to preserve the relative weightings of position and scale errors.

References

[1] P. Bendale, B. Triggs, and N. Kingsbury. Multiscale keypoint analysis based on complex wavelets. In *British Machine Vision Conference*, August 2010.

[2] O. Chum, T. Werner, and T. Pajdla. Joint orientation of epipoles. In *British Machine Vision Conference*, 2003.

[3] O. Faugeras, Q.-T. Luong, and S.J. Maybank. Camera self calibration: Theory and experiments. In *European Conf. Computer Vision*, Santa Margherita Ligure, Italy, May 1992. Springer-Verlag.

[4] P-E. Forssén and A. Moe. Blobs in epipolar geometry. In *SSBA Symposium on Image Analysis*, pages 82–85, Uppsala, 2004. See also Linköping Univ. Tech. Report LiTH-ISY-R-2620.

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

[6] F. Kahl and A. Heyden. Using conic correspondence in two images to estimate the epipolar geometry. In *Int. Conf. Computer Vision*, 1998.

[7] S. Laveau and O. Faugeras. Oriented projective geometry for computer vision. In *European Conf. Computer Vision*, pages 147–156. Springer-Verlag, 1996.

[8] D.G. Lowe. Distinctive image features from scale-invariant keypoints. *Int. J. Computer Vision*, 60(2):91–110, 2004.

[9] K. Mikolajczyk and C. Schmid. Scale and affine invariant interest point detectors. *Int. J. Computer Vision*, 60(1):63–86, 2004.

[10] K. Mikolajczyk, T. Tuytelaars, C. Schmid, A. Zisserman, J. Matas, F. Schaffalitzky, T. Kadir, and L. Van Gool. A comparison of affine region detectors. *Int. J. Computer Vision*, 65(1/2):43–72, 2005.

[11] B. Triggs. Autocalibration and the absolute quadric. In *IEEE Conf. Computer Vision & Pattern Recognition*, Puerto Rico, 1997.

[12] B. Triggs. Detecting keypoints with stable position, orientation and scale under illumination changes. In *European Conf. Computer Vision*, pages IV 100–113, May 2004.

[13] S. Winder, G. Hua, and M. Brown. Picking the best DAISY. In *IEEE Conf. Computer Vision & Pattern Recognition*, 2009.