Many recent keypoint detectors [3, 4] associate a local scale (for multiscale detectors) or even a full affine frame (for affine-invariant detectors) to each detected keypoint. Conventional epipolar geometry [2] only constrains the relative positions of corresponding keypoints, not their relative scales. We present an enhanced epipolar constraint that exploits both positions and scales, thus making correspondence search 2–4 times more accurate in practice.

The method works as follows:
- encode multiscale keypoints as image ellipses;
- invoke the ‘Kruppa constraints’ that link corresponding ellipses [2];
- project to the “epipolar pencil” (the 1-D family of epipolar lines) to get reduced constraints linking 1-D quadratic forms on the pencil;
- enforce a scale-sensitive (angular position, angular width) error model by a well-chosen algebraic transformation of this representation.

Figure 1 illustrates the projection process. The $2 \times 3$ matrices $B, B'$ generating the projections onto the epipolar pencil are extracted from the singular value decomposition of the fundamental matrix [2] via $F = USV^\top = F = B' \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} B'$, whence $B \equiv (u_2, -u_1)$ and $B' \equiv (S_{11} v_1, S_{22} v_2)^\top$. Possibly-corresponding pairs of multiscale keypoints are represented as image ellipses – $3 \times 3$ symmetric dual-form conic matrices $q, q'$ – and projected onto the epipolar pencil by $BqB'$ and $B'q'B'$. The reduced Kruppa constraints say that if the ellipses correspond, these two symmetric $2 \times 2$ matrices agree up to scale. Algebraically, this turns out to provide a strong constraint on the correspondence of keypoint centres, but only a weak second order one on the correspondence of their scales. To work around this, we algebraically transform the constraint to the following (keypoint position, keypoint scale) “normalized distance” model:

$$d_{\theta} = \frac{\sin 2(\theta - \theta')}{\sin \theta + \sin \theta'} = \frac{(p - q'p')^2}{1 - (r + r')/2} \quad (1)$$

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

Here, $(p, q, r)$ and $(p', q', r')$ (with scaling to $p^2 + q^2 = 1$) linearly encode the 3 independent entries of the matrices $BqB'$ and $B'q'B'$. and $(\theta, \delta\theta)$ and $(\theta', \delta\theta')$ are corresponding mean angles. angular width representations (with angles measured in projective coordinates). The above forms have appropriate small and large angle limits and close relationships to standard statistical error models. The formulae above assume standard full-epipolar-line matching but the method also handles signed (epipolar half-line) correspondance and omnidirectional images via a twofold unwrapping process based on oriented projective geometry.

The final method is very simple to use – perhaps even simpler than standard epipolar line search – and it gives good results on both synthetic images (see fig. 2) and real images (see fig. 3). In both cases, incorporating the additional scale constraint into the epipolar matching process cuts the number of false positive matches by a factor of 2–4 over a wide range of camera geometries and imaging conditions.


