noRANSAC for fundamental matrix estimation

Fabio Bellavia
fabio.bellavia@unipa.it

Domenico Tegolo
domenico.tegolo@unipa.it

Department of Mathematics
and Computer Science
University of Palermo
Palermo, Italy

The estimation of the fundamental matrix from a set of corresponding points is a relevant topic in epipolar stereo geometry [2]. Due to the high amount of outliers between the matches, RANSAC-based approaches [1] have been used to obtain the fundamental matrix.

We introduce a new normalized epipolar error measure which takes into account the shape of the features used as matches [3] and does not introduce any relevant computational cost.

Moreover, a new evaluation strategy is described as a valid tool to compare the estimated fundamental matrices. It does not rely on the inlier ratio, which could not correspond to the best allowable fundamental matrix estimated model, but it makes use of a reference ground truth fundamental matrix obtained by a set of corresponding points given by the user.

Let $\mathscr{R}_1$, $\mathscr{R}_2$ be two elliptical feature patches belonging respectively to the images I_1, I_2, centred in $\mathbf{x}_1$, $\mathbf{x}_2$ as commonly extracted by feature detectors [3], with minor and major axes respectively α_{min_i}, α_{max_i}, $i \in \{1,2\}$. The error measure κ_i in the image I_i, for the feature pair $(\mathscr{R}_1, \mathscr{R}_2)$ is defined as

$$\kappa_i = \min\left(\frac{d(\mathbf{x}_i, \mathbf{l}_i)}{\alpha_{min_i}}, 1\right) \tag{1}$$

that is, the epipolar distance $d(\mathbf{x}_i, \mathbf{l}_i)$ between the feature centre $\mathbf{x}_i$ and its epipolar line $\mathbf{l}_i$, computed by using the corresponding point in the other image, is normalized by the minor axis of the feature ellipse $\mathscr{R}_i$.

The error κ_i achieves the maximal value of 1 roughly when the supposed reprojected feature ellipse would not touch the actual ellipse, as shown in Figure 1. Clearly, when accidentally a wrong feature $\mathscr{R}_i$ lies close to the correct epipolar line $\mathbf{l}_i$, the error κ_i is misleading, as it also happens for both the Sampson error and the epipolar distance $d(\mathbf{x}_i, \mathbf{l}_i)$.

The proposed error κ_i does not depend on the image scale and provides a soft threshold t to be used by RANSAC approaches, thus a possible overfitting on matches derived by a non-optimal choice of t can be alleviated.

Finally, the error on both the image is combined into a vector $[\kappa_1\ \kappa_2]$ and leads to different error measures. In particular the L_1, L_2 and L_∞ norms have been used, denoted as the symmetric, geometric and max errors respectively.

In order to compare fundamental matrices estimated by different algorithms on non-synthetic data, the inlier ratio is commonly adopted [4]. Although the maximization of the number of inliers coincides with the formulation of the optimization problem used by RANSAC-based approaches, i.e. find the best F compatible with the largest input dataset, this does not always correspond to the desired real solution.

For istance, by increasing the threshold t, a large consensus set of points is usually found which could wrongly lead to include outliers. Moreover, when threshold errors cannot be comparable due to the different error measures adopted, it could be misleading to compare methods for the fundamental matrix estimation according to the inlier ratio. Furthermore, the theoretical best model obtained by a robust estimator algorithm, could not meet the correct solution in some frequent degenerate cases, for instance when a dominant plane is present and the initial set of matches contains an high fraction of outliers.

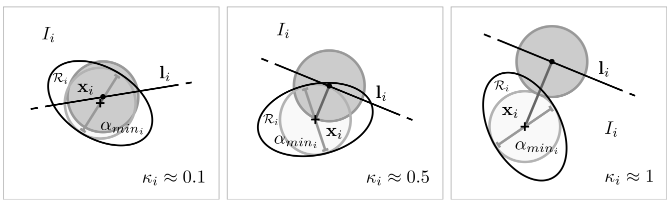

Figure 1: Examples of different values of the normalized epipolar distance κ_i. The dark grey circle represents the reprojected feature supposed by the error measure κ_i, the light grey circle is the approximation of the feature ellipse $\mathscr{R}_i$. The epipolar distance $d(\mathbf{x}_i, \mathbf{l}_i)$ is given by the dark grey segment joining the centres of the two circles and the minor axis α_{min_l} is shown as the light grey segment.

In order to deal with these issues a new relation between lines on the image is defined. Given the true epipolar line $\mathbf{l}_{\bar{i}}$ and its estimation $\widetilde{\mathbf{l}}_{\bar{i}}$ on the image $I_{\bar{i}}$ corresponding to the point $\mathbf{x}_i$ on the other image I_i, a cone is obtained by intersecting the respective two half-planes so that the minimum intersection angle is considered (in the case of parallel lines the non-empty intersection is taken). The resulting surface $\varphi_i(\mathbf{x}_i)$ on $I_{\bar{i}}$, normalized to the image area, can be seen as the minimum amount of work needed to move the estimated epipolar line to the correct one.

Thus an indirect measure between the fundamental matrices F and its estimation $\widetilde{\mathrm{F}}$ can be draw out for each point in the stereo pair. The corresponding error surface is almost continuous on the image, as shown in Figure 2, and defines a fingerprint of the difference between the matrices. This error measure φ_i can be low for an high image portion, because for a finite epipole $\mathbf{e}_i$ and its estimation $\widetilde{\mathbf{e}}_i$ the corresponding epipolar line pencils share a common line for which $\varphi_i = 0$ and due to the continuity near this map area low values can occur (see Figure 2 (d-f)). However, the maximum ς_i of φ_i

$$\varsigma_i = \max_{\mathbf{x}_i \in I_i} \varphi_i(\mathbf{x}_i) \tag{2}$$

can give a good indication about the precision of the matrix estimation $\widetilde{\mathrm{F}}$ with respect to the true fundamental matrix F when points from the image I_i are projected to image $I_{\bar{i}}$. The maximum ς on both images $\varsigma = \max(\varsigma_1, \varsigma_2)$ is finally used as error value.

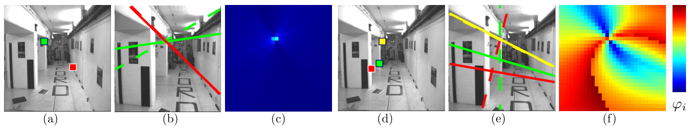

Figure 2: Two different estimations of the same fundamental matrix. The points on the image I_i (a,d) correspond to the estimated (solid) and the true (dashed) epipolar lines on the other image $I_{\bar{i}}$ (b,e). Clearly the model (a-c) is better, as it is confirmed by inspecting the sampled maps φ_i (c,f). In the model (d-f) there is a discontinuity between the red and the green points, where the minimum angle made by the epipolar lines switches. When both the true and the estimated yellow epipolar lines pass through both the estimated and the true epipoles $\varphi_i(\mathbf{x}_i) = 0\%$.

According to the new proposed evaluations strategy, the new *no*rmalized epipolar distance provides better results when applied to RANSAC or MLESAC, defined as noRANSAC and noMLESAC respectively, especially with the geometric and the max errors. Moreover, it does not depend on the input image scale, which makes it more robust and allows a stable threshold selection for RANSAC-based approaches.

Details of the proposed methods, of the experimental evaluation and results are described in the paper.

[1] M.A. Fischler and R.C. Bolles. Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography. *Communications of the ACM*, 24(6):381–395, 1981. ISSN 0001-0782.

[2] R. Hartley and A. Zisserman. *Multiple View Geometry in Computer Vision*. Cambridge University Press, 2000. ISBN 0-521-62304-9.

[3] K. Mikolajczyk, T. Tuytelaars, C. Schmid, A. Zisserman, J. Matas, F. Schaffalitzky, T. Kadir, and L. Van Gool. A comparison of affine region detectors. *International Journal of Computer Vision*, 65(1-2): 43–72, 2005. ISSN 0920-5691.

[4] P.H.S. Torr and A. Zisserman. MLESAC: a new robust estimator with application to estimating image geometry. *Computer Vision and Image Understanding*, 78(1):138–156, 2000. ISSN 1077-3142.