

Epipolar Geometry from Two Correspondences.

Michal Perďoch, Jiří Matas and Chum Ondřej

Center for Machine Perception, Czech Technical University, Karlovo nám. 13

Abstract

A novel algorithm for robust RANSAC-like estimation of epipolar geometry (of uncalibrated camera pair) from two correspondences of local affine frames (LAFs) is presented. Each LAF is constructed from three points independently detected on a maximally stable extremal region.

The algorithm assumes that a sufficiently accurate approximation of the fundamental matrix is obtained from two LAF correspondences by the 6-point algorithm of Stéwenius et al. The so-far-the-best hypotheses are further processed by so-called local optimization to estimate the epipolar geometry. Special attention is paid to planar sample degeneracy, since the probability of drawing two coplanar LAF correspondences is not negligible. Combining the 6-point solver, local optimization, and the degeneracy test enables RANSAC to draw samples of only two LAFs to generate hypotheses and thus to reduce the number of samples drawn.

We experimentally show that using the 6-point algorithm (approximating the real camera by camera with unit aspect ratio, zero skew, principal point in the center of image, and a common unknown focal length) generates hypotheses that are sufficient for EG estimation in LO-RANSAC framework.

1. Introduction

The problem of wide-baseline stereo has recently received considerable attention [6, 14, 7, 9]. As a rule, wide-baseline methods establish correspondence of entities which provide geometric constraints stronger than a single point-to-point correspondence. For instance, the popular SIFT keypoint operator [6] determines a location, scale and orientation; the latter two quantities implicitly define a second point. The LAF-MSER method of Obdržálek et al. [8] exploits local affine frames (LAFs), i.e. three ordered points detected in an affine-covariant way.

The seven point algorithm embedded in RANSAC is a standard method for epipolar geometry estimation [5]. Chum and Matas [1, 2] studied whether the required seven

point-to-point correspondences obtained from three correspondences of LAFs allows efficient estimation of epipolar geometry. Since the speed of RANSAC is inversely proportional to an exponential function of the sample size, drawing at random three LAF instead of seven point-to-point correspondences potentially reduces running times by orders of magnitude. The key issue is whether the spatial distribution of the nine points – three triplets of nearby point in a LAF – allows good estimates of epipolar geometry. Experiments in [2] show that standard RANSAC fails in this case. However, a simple modification called *local optimisation of the so-far-the-best solution* leads to an algorithm that benefits from the small sample size without losing efficiency; speed-ups of up to 10^3 are reported.

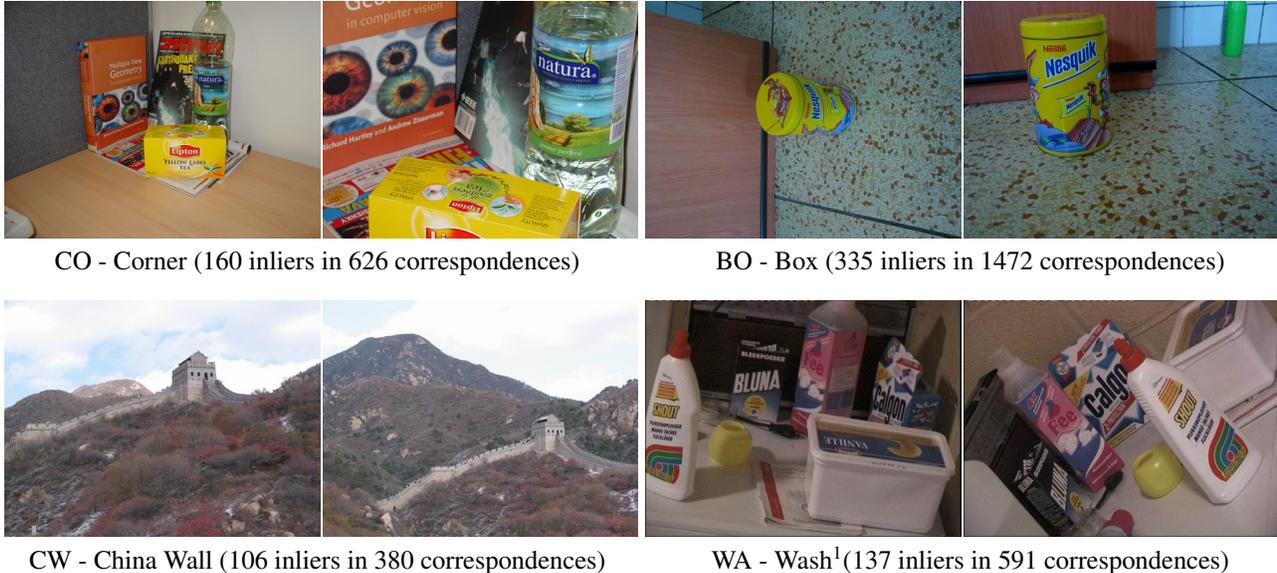
In this paper, we take the approach to an extreme and design an algorithm that computes EG from two correspondences of local affine frames using the recently proposed 6-point EG estimation algorithm of Stéwenius et al. [10, 11]. The new 2LAF-LO-RANSAC algorithm is shown experimentally to have competitive performance, requiring a very small number of RANSAC iterations.

The rest of paper is structured as follows. In Section 2 the structure of the 2LAF-LO-RANSAC algorithm and its components are described in detail. Experiments presented in Section 3 validate the design decisions and performance of the algorithm. The paper is concluded in Section 4.

2. Algorithm

We first introduce the building blocks of the 2LAF-LO-RANSAC algorithm: the RANSAC estimator [4], the local optimisation for RANSAC [1], a method detecting degenerated configurations [3] and the six-point solver for EG [10, 11].

RANSAC is a simple but powerful algorithm. Repeatedly, subsets are randomly selected from input data \mathcal{T} and model parameters θ fitting the sample are computed. The size m of the random samples is the smallest sufficient for determining model parameters. In a second step, the quality of the model parameters is evaluated on the full data set. Different cost functions may be used [13], the standard being the size of the support, i.e. the number of data points consistent with the model. The process is terminated when the likelihood



CO - Corner (160 inliers in 626 correspondences)

BO - Box (335 inliers in 1472 correspondences)

CW - China Wall (106 inliers in 380 correspondences)

WA - Wash¹(137 inliers in 591 correspondences)**Figure 1.** Images pairs and characteristics of problems used in experiments.

of finding a better model becomes low, i.e. the probability η of missing a set of inliers \mathcal{I} of size I within k samples falls under predefined threshold η_0

$$\eta = (1 - P(I))^k. \quad (1)$$

The probability $P(I)$ that an uncontaminated sample of size m is randomly selected from $N = |\mathcal{T}|$ correspondences is

$$P(I) = \frac{\binom{I}{m}}{\binom{N}{m}} = \prod_{j=0}^{m-1} \frac{I-j}{N-j} \approx \varepsilon^m, \quad (2)$$

where ε is the fraction of inliers $\varepsilon = I/N$. Let the support $\mathcal{S}(\theta, \mathcal{T}, \Delta) : \Theta \rightarrow \mathcal{T}$ be the set of data points consistent with model θ i.e.

$$\mathcal{S}(\theta, \mathcal{T}, \Delta) = \{\mathbf{x} \in \mathcal{T} | \rho(\theta, \mathbf{x}) \leq \Delta\},$$

where $\rho(\theta, \mathbf{x})$ is a distance measure (e.g. Sampson's error [5]) and Δ a user-specified threshold. The average number of samples that has to be drawn to satisfy $\eta \leq \eta_0$ is

$$k_{\eta_0}(\mathcal{S}_k) = \frac{\ln(\eta_0)}{\ln(1 - P(\mathcal{S}_k))}. \quad (3)$$

LO-RANSAC (RANSAC with local optimisation). Note that the expression for probability η controlling RANSAC termination as defined by Equations (1) and (2) is valid only under the assumption that a selection of any random sample not contaminated by outliers leads to the correct solution, i.e. to a solution supported by all inliers. However, as shown in [1], this assumption is often invalid. For standard 7-point and 8-point EG estimators, all-inlier samples fail at a rate that can be unnoticed and/or ignored², despite

the fact that the rate is not negligible — Murray and Tordoff [12] observed that about two to three times more samples were needed than predicted by theory. For algorithms that estimate EG from LAFs, we observed that the percentage of all-inlier samples leading to a correct EG estimate is negligible, less than 0.1% on the scenes shown in Fig. 1! This observation is consistent with [2], which demonstrates that EG cannot be efficiently computed from three LAF correspondences by standard RANSAC. However, it is also shown in [1] that the so called "local optimisation" overcomes the problem. We show, if an appropriate local optimisation is adopted, the result also holds for the six point solver [11] estimating EG from two LAF correspondences.

The core insight of LO-RANSAC is the following. For the success of a particular estimator within a RANSAC framework, it is not necessary that any all-inlier sample generates an EG hypothesis that is close to optimal. It is sufficient if a significant proportion of EGs estimated from all-inlier samples has support higher (more correspondences consistent with it) than almost all EGs estimated from contaminated samples, i.e. EGs with random support. If this property holds, an EG hypothesis from an all-inlier sample is very likely to be so-far-the-best, i.e. it is the EG with the largest support observed so far. It is proved in [1] that as the RANSAC progresses, only very few so-far-the-best EG hypothesis emerge. It is therefore possible to apply even fairly complex optimisation procedures on the so-far-the-best model without affecting overall time complexity.

Degeneracy. The presence of degenerate configurations is a surprisingly common mode of failure of RANSAC. A (dominant) plane is a typical degenerate configuration in EG estimation – EG cannot be recovered from a coplanar, even an all-inlier one. In [3], it was shown that even if a part of a

¹Images of Wash were gracefully provided by Tinne Tuytelaars.

²At least the problem has not been often reported in the literature.

sample (at least five points out of a seven point sample) lies in a plane, the estimated EG may be incorrect and yet having significantly higher than random support. Unsurprisingly, the problem of degenerated configurations is acute if an EG is estimated from a small number (three or two) of LAFs. The smaller the sample the higher the probability that the whole sample comes from a single plane. Potential degeneracy is therefore checked for each so-far-the-best sample, as suggested in [3]. When a planar configuration is detected, the plane-and-parallax [3] algorithm is exploited to estimate the EG.

Six point solvers in 2LAF-LO-RANSAC. Two LAFs [8] provide six point-to-point correspondences. Practical six point solvers have been recently introduced [10, 11]. We adopted the algorithm [11], since it has the least restrictive assumptions about the camera. In [11], it is assumed that the intrinsic camera parameters are known up to a common focal length f . In the paper, the (unknown) intrinsic parameters are assumed to be: a unit aspect ratio, zero skew, and the principal point in the center of the image. Such an approximation is reasonable for many currently used sensors. Experimentally, we show that the algorithm is not affected by slight deviation from the model (i.e. different focal length, rectangular pixels see Exp. 3). Note, that the EG based on the assumptions on actually unknown intrinsic parameters is only used as an approximation of the fundamental matrix F . Such an approximation allows generation of reasonable initial hypotheses while reducing the sample size to two LAF correspondences.

Input: set of data points \mathcal{T} , thresholds η_0, Δ

Output: model θ^* with largest support \mathcal{S}^*

1. **repeat**
 2. Draw a random sample of minimal size m from data points \mathcal{T} .
 3. Model parameters θ_i computed from the minimal set by estimator E .
 4. Calculate the support \mathcal{S}_i of the model θ_i .
 5. **if** new maximum has occurred ($|\mathcal{S}_i| > |\mathcal{S}_j|$ for $\forall j < i$):
 6. **if** a degenerated configuration is detected
 7. Perform plane-and-parallax search for EG and appropriate local optimisation.
 8. **otherwise**
 9. Apply local optimisation.
 10. Store the best model θ^* with support \mathcal{S}^* .
 11. **until** the probability η of finding model with support larger than \mathcal{S}^* falls under η_0 .
-

Algorithm 1. The structure of LO-RANSAC with the degenerated configuration test.

2.1. 2LAF-LO-RANSAC algorithm

The 2LAF-LO-RANSAC algorithm is an instance of LO-RANSAC with a test for a degenerated configuration. Its structure is summarized in Alg. 1; the parameters are: \mathcal{T} is a set of LAF correspondences, the minimal size $m = 2$, the estimator E is the six point solver [11]. Handling of degenerated configurations [3] proceeds as follows. If a non-random number of correspondences consistent with a homography H is found, a *plane-and-parallax* method [3] is employed. A LAF correspondence outside the plane is sought by RANSAC sampling tentative correspondences off the plane. If a correspondence is found that leads to an EG with support significantly (i.e. non-randomly) larger than the support of the original H -dominated estimate, the plane and parallax algorithm is successful.

3. Experimental results

Four scenes (Fig. 1) with different number of correspondences and inlier ratios were chosen to evaluate efficiency and performance of the 2LAF-LO-RANSAC algorithm. In all experiments, LAFs on maximally stable extremal regions [7] were detected and three point tentative region correspondences (\mathcal{T}) established as in [8]. The quality of tentative correspondences was artificially degraded by untuning the parameters of the matching process.

Abbrev.	Estim.	Degen. test	Data points
7PTS	7pt	✓	7x LAF centers
3LAF	8pt	✓	3x LAF
3LAF ND	8pt	—	3x LAF
2LAF	6pt	✓	2x LAF
2LAF ND	6pt	—	2x LAF

Table 1. LO-RANSAC variants used.

A reference model θ_r of EG was computed using seven point LO-RANSAC and a set \mathcal{I}_r , correspondences consistent with θ_r was stored (see Fig. 1). Compared LO-RANSAC variants are listed in Tab. 1.

Experiment 1 (Efficiency of 2LAF-LO-RANSAC). In theory, i.e. according to equation (2), drawing minimal samples of size $m = 2$ should increase the probability $P(I)$ and significantly reduce the expected average number of samples k (Eq. 3). In practice, a certain fraction of all-inlier samples will lead to random EG due to the influence of noise, poor conditioning and violated assumptions e.g. of equal focal lengths. We define *efficiency* of a LO-RANSAC variant as the fraction of the all-inlier samples that leads to a model with support size (approximately) equal to I . To measure and compare the efficiency, the following synthetic experiment (see Tab. 2) was conducted.

Scn.	Alg.	$ \theta_B $	$ \theta_D $	$ \theta_N $	$ \theta_G $
CO	7PTS	6	733	261	785
	3LAF	115	779	106	641
	3LAF ND	117	0	883	458
	2LAF	184	676	140	553
	2LAF ND	163	0	837	362
BO	7PTS	0	957	43	998
	3LAF	71	912	17	918
	3LAF ND	78	0	922	900
	2LAF	209	707	84	766
	2LAF ND	224	0	776	743
CW	7PTS	19	466	515	815
	3LAF	342	517	141	341
	3LAF ND	328	0	672	246
	2LAF	298	515	187	274
	2LAF ND	307	0	693	184
WA	7PTS	110	195	695	646
	3LAF	380	369	251	353
	3LAF ND	396	0	604	180
	2LAF	425	140	435	247
	2LAF ND	408	0	592	161

Table 2. Synthetic test, results for 1000 random all-inlier samples. For ND algorithms all hypotheses are assumed non-degenerated.

One thousand random all-inlier samples are generated from the inliers of the reference epipolar geometry. For each sample, an initial hypothesis θ_0 is computed from a minimal set of data points. Subsequent steps of LO-RANSAC (see Alg. 1) variants with local optimisation and degeneracy test are executed. Column θ_B shows the number of θ_0 hypotheses that are consistent with less than an accidental number of correspondences. This fraction of all-inlier hypotheses cannot be distinguished from an outlier contaminated hypothesis and hence degrades the efficiency of LO-RANSAC. The number of degenerated and non-degenerated all-inlier samples is shown in columns θ_D and θ_N . Finally, the number of “good” hypothesis is shown in column θ_G . Hypotheses are considered “good” if the number of consistent correspondences is at least 90% of the $|\mathcal{L}_r|$, i.e. the number of inliers of reference EG.

We observed (see Tab. 2) that the efficiency of the seven point estimator does not drop under 65% even on the noisiest scene – Wash. 3LAF-LO-RANSAC and 2LAF-LO-RANSAC performs better on the scenes Corner and Box with higher inlier ratio. On the noisy and hard scenes, their efficiency drops down to around 30% for 3LAF-LO-RANSAC algorithm and 25-30% for 2LAF-LO-RANSAC algorithm. Therefore, to achieve a desired probability of success η_0 , the number of samples drawn before termination needs to take efficiency in account. Note however that in the performance test (Exp. 3) theoretical number of samples is used to observe differences between algorithms.

Experiment 2 (The robustness of the 6pt solver).

Scn.	Alg.	#inls	#samples	#LOs
CO	7PTS	162.13	42307.93	8.97
	3LAF	161.23	177.47	4.28
	3LAF ND	157.69	219.83	4.41
	2LAF	160.74	44.65	3.03
	2LAF ND	132.22	86.39	3.69
BO	7PTS	334.83	99838.93	9.13
	3LAF	334.46	256.08	4.74
	3LAF ND	321.23	290.75	4.27
	2LAF	334.29	57.02	2.79
	2LAF ND	315.58	64.45	2.81
CW	7PTS	106.00	26396.00	8.73
	3LAF	104.07	151.15	3.76
	3LAF ND	102.86	203.60	3.76
	2LAF	101.50	46.84	2.84
	2LAF ND	90.38	73.64	3.34
WA	7PTS	136.43	96872.83	8.30
	3LAF	135.14	257.93	3.10
	3LAF ND	125.23	484.80	3.73
	2LAF	130.27	69.61	2.37
	2LAF ND	121.22	104.86	2.84

Table 3. Average number of inliers, samples and number of LOs taken in 100 runs of RANSAC algorithms.

Proposed 2LAF-LO-RANSAC algorithm uses the six point solver [11] to estimate the initial hypothesis that assumes calibrated cameras up to an unknown common focal length. The first three assumptions: the principal point in the middle of the image, zero skew, and unit aspect ratio often hold with modern cameras, while the assumption of common focal length is often violated. The sensitivity to different focal lengths is tested on the Corner scene with focal length ratio of about 3. No significant decrease in performance was observed. Robustness to aspect ratio was tested on the China Wall scene. One coordinate was artificially scaled with factor from 0.5 to 2.0. Similar performance (see Fig. 2) was observed for all factors.

Experiment 3 (The speed of RANSAC algorithms).

The speed of 2LAF-LO-RANSAC, 3LAF-LO-RANSAC and seven point LO-RANSAC is summarised in Tab. 3. As expected, the average number of samples is highest for the seven point algorithm. For 2LAF-LO-RANSAC algorithm it is around one fourth (i.e. proportional to ε) of the 3LAF-LO-RANSAC algorithm and about thousand times less compared to the seven point algorithm on hard scenes.

For deeper insight in the performance of algorithms, the histograms Fig. 3 of the support size of the model found in each run were computed. Results show that the degeneracy test and the special local optimisation is necessary for 2LAF-LO-RANSAC algorithm. The 2LAF-LO-RANSAC algorithm without the degeneracy test, i.e. when the so-far-the-best models are processed only with local optimisation, fail to produce a stable solution on difficult problems. On

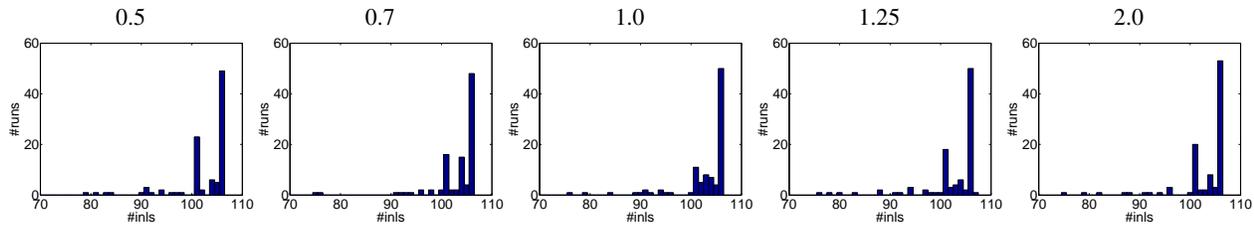


Figure 2. Number of inliers in 100 runs of 2LAF algorithm on China Wall scene with different aspect ratios.

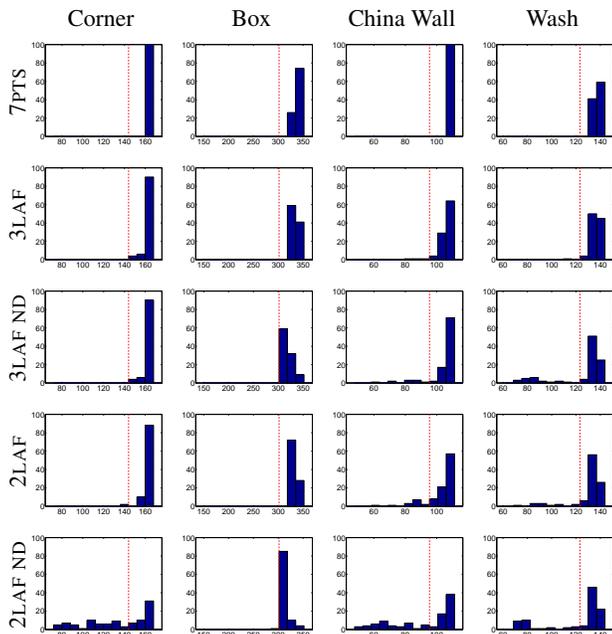


Figure 3. Number of inliers in 100 runs of RANSAC algorithms. Dashed line delimits 90% of inliers according to reference model.

the China Wall and Wash scenes, approximately 40% of executions (see Fig. 3) produced solutions with less than 90% of the reference method’s support.

4. Conclusions

The 2LAF-LO-RANSAC algorithm that estimates epipolar geometry from two LAF correspondences was proposed. Two LAF correspondences provide six point-to-point correspondences that are used to generate RANSAC hypotheses by the state-of-the-art 6pt solver [10, 11]. Experimentally, it was shown that the algorithm performs well only if the methods [2, 3] for local optimisation and degeneracy detection are used. Compared to 3LAF-LO-RANSAC algorithm, the 2LAF-LO-RANSAC algorithm requires lower number of samples, local optimisations, and significantly lower number of iterations than the seven point LO-RANSAC with degeneracy test. In experiments where the inlier ratio ranged

from 23% to 28% the number of iterations ranged from 45 to 70 for 2LAF-LO-RANSAC, 151 to 258 for 3LAF-LO-RANSAC, and 26396 to 99840 for LO-RANSAC.

References

- [1] O. Chum, J. Matas, and J. Kittler. Locally optimized ransac. In *Proc. of the DAGM*, pages 236–243, 2003.
- [2] O. Chum, J. Matas, and Š. Obdržálek. Enhancing RANSAC by generalized model optimization. In *Proc. of the ACCV*, volume 2, pages 812–817, January 2004.
- [3] O. Chum, T. Werner, and J. Matas. Two-view geometry estimation unaffected by a dominant plane. In *Proc. of the CVPR*, pages 772–780, June 2005.
- [4] M. Fischler and R. Bolles. Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography. *CACM*, 24(6):381–395.
- [5] R. Hartley and A. Zisserman. *Multiple view geometry in computer vision*. Cambridge, 2nd edition, 2003.
- [6] D. Lowe. Distinctive image features from scale-invariant keypoints. *IJCV*, 60(2):91–110, 2004.
- [7] J. Matas, O. Chum, M. Urban, and T. Pajdla. Robust wide baseline stereo from maximally stable extremal regions. In *Proceedings of the BMVC*, volume 1, pages 384–393, 2002.
- [8] J. Matas, Š. Obdržálek, and O. Chum. Local affine frames for wide-baseline stereo. In *Proc. of the ICPR*, 2002.
- [9] P. Pritchett and A. Zisserman. Wide baseline stereo matching. In *Proc. ICCV*, pages 754–760, 1998.
- [10] H. Stewénius. *Grbner Basis Methods for Minimal Problems in Computer Vision*. PhD thesis, LTH, Sweden, 2005.
- [11] H. Stewénius, D. Nister, F. Kahl, and F. Schaffalitzky. A minimal solution for relative pose with unknown focal length. In *Proc. of the CVPR*, pages 789–794. IEEE, 2005.
- [12] B. Tordoff and D. Murray. Guided sampling and consensus for motion estimation. In *Proc. 7th ECCV*, volume 1, pages 82–96. Springer-Verlag, 2002.
- [13] P. H. S. Torr and A. Zisserman. MLESAC: A new robust estimator with application to estimating image geometry. *CVIU*, 78:138–156, 2000.
- [14] T. Tuytelaars and L. Van Gool. Wide baseline stereo matching based on local, affinely invariant regions. In *Proc. of the 11th BMVC*, 2000.