A New A Contrario Approach for the Robust Determination of the Fundamental Matrix

Ferran Espuny[1], Pascal Monasse[2], and Lionel Moisan[3]

[1] School of Environmental Sciences, University of Liverpool
Ferran.Espuny-Pujol@liverpool.ac.uk
[2] Université Paris-Est, LIGM (CNRS UMR 8049),
Center for Visual Computing, ENPC, F-77455 Marne-la-Vallée
pascal.monasse@imagine.enpc.fr
[3] Université Paris Descartes, MAP5 (CNRS UMR 8145)
lionel.moisan@parisdescartes.fr

Abstract. The fundamental matrix is a two-view tensor that plays a central role in Computer Vision geometry. We address its robust estimation given correspondences between image features. We use a non-parametric estimate of the distribution of image features, and then follow a probabilistic approach to select the best possible set of inliers among the given feature correspondences. The use of this perception-based *a contrario* principle allows us to avoid the selection of a precision threshold as in RANSAC, since we provide a decision criterion that integrates all data and method parameters (total number of points, precision threshold, number of inliers given this threshold). Our proposal is analyzed in simulated and real data experiments; it yields a significant improvement of the ORSA method proposed in 2004, in terms of reprojection error and relative motion estimation, especially in situations of low inlier ratios.

Keywords: stereovision; fundamental matrix; feature matching, *a contrario* model; outlier detection

1 Introduction

Matching between features in two images is an important component of many methods for 3D reconstruction, camera motion estimation and pattern recognition [7]. An appropriate set of matches can determine the geometry between two images, also known as *epipolar geometry*, which for non-planar scenes is characterized by the *fundamental matrix*.

The fundamental matrix is a two-image tensor that associates to any point in the first image an *epipolar line* of possible corresponding points (*matches*) in the second image. In particular, the tensor evaluated at any pair of perfectly corresponding points is zero (*epipolar constraint*). All epipolar lines intersect in a point of the second image, called *epipole*, that is the projection of the optical center of the first camera; this algebraically implies that the fundamental

matrix has rank two. The so-called seven point algorithm exploits such a rank-two constraint to provide up to three possible solutions for the fundamental matrix given (the minimal number of) seven matches [7].

Feature correspondences obtained with standard detection and matching algorithms (we use SIFT features [10]) are usually unsuitable for the direct computation of the fundamental matrix due to noise and outliers. The computation of the fundamental matrix given only inlier matches has been studied elsewhere, see e.g. [9] for a review.

Given a set of tentative correspondences between two images, containing noise and outliers in an unpredictable manner, we address the problem of robustly identifying the set of inlier matches and computing the fundamental matrix. We assume generic scene(s) and camera position(s); dedicated methods have been devised to deal with nearly-degenerate configurations [18,1,4,5].

The Random Sample Consensus (RANSAC) [3] performs random sampling of at most N minimal subsets of data (7-uples of matches) to hypothesize models. Each hypothetical model is evaluated using the point matches whose residual is under a user-specified threshold τ (these matches are the estimated inliers), and the model that leads to the maximum number of estimated inliers is chosen (the final model is then re-estimated using only these inliers). The cost function that RANSAC tries to optimize assigns 0 to those matches with error under the threshold τ (inliers) and a constant penalty otherwise [17]. An inlier ratio ε is usually user-specified to decide the number of random trials N, although this decision can also be taken adaptively using the best model found so far [7].

The Optimized Random Sampling Algorithm (ORSA) [11] (detailed in Section 2) assumes a uniform spatial distribution of the features over the image. Through random sampling, it looks for a set of matches that is the least expected in terms of the precision achieved for a given number of inliers (*a contrario* criterion). It requires as parameter only the maximum number of random trials to perform (that is, the computational effort allowed by the user). The actual number of random trials can drop to as little as 10% of the specified one when a "sufficiently good" model is found early (acceleration strategy). It can be seen as a non-parametric version of RANSAC, where all possible inlier thresholds τ are implicitly used (and compared) thanks to the *a contrario* criterion. Note that this method was recently adapted to the case of homographies [12].

ORSA has been shown to outperform a fixed-threshold RANSAC approach and classical methods like M-estimators and LMedS, in particular by its ability to deal with very high outlier ratios [11,13]. However, performance lacks and even failure cases can be observed when the data points are far from being uniformly distributed, in particular when most of them lie in a small region of the image domain [11]. In the search for an adaptive improvement of ORSA, we here study the replacement of the uniform distribution of the ORSA method with an empirical distribution estimated non-parametrically using plug-in methods [16].

We only carry out a preliminary comparison with the RANSAC and ORSA methods, the latter's source code being available online[4]. Comparison with other

[4] `http://www.mi.parisdescartes.fr/~moisan/epipolar/`

modern methods was not possible due to the lack of available code. A major difference between the proposed approach and almost all existing ones is that we build a model of the distribution of 2D image features to recover an optimal threshold, independently of any matching, whereas other methods propose a model of (at best) the distribution of the 1D errors associated to the 2D matches.

The paper is organized as follows. After reviewing the ORSA method in Section 2, we adapt it in Section 3 to the case of an empirical distribution of image features. We then detail the geometric and motion estimation error measures that we use. Evaluation follows in Section 5 using simulated and real data, in both cases with groundtruth available, before we conclude in Section 6.

2 Fundamental Matrix Computation using ORSA

The ORSA method (Algorithm 1) is based on [11]: first, the use of the RANSAC paradigm, performing several random trials to generate potential true 7-point matches; second, the use of an *a contrario* criterion to select the best set of 7-point matches (lines 5–6 in Algorithm 1); third, a final sampling strategy that adaptively selects the set of potential matches (lines 9–10 in Algorithm 1).

Consider a set T of 7-point matches and one of the (up to 3) fundamental matrices estimated from it using the normalized 7-point algorithm [7]. For each of the n point matches, we can estimate a match error according to the selected fundamental matrix F. Applying a threshold τ to this error will result in a selection of estimated inliers (error smaller than τ) and outliers (error larger than τ). Depending on the value of τ, we will thus select a certain number k of potential inliers, with $7 < k \leq n$. The *a contrario* criterion associates to each of these $\binom{n-7}{k-7}$ possible sets S of k inliers (containing the initial set T) a *number of false alarms*

$$\mathrm{NFA}(S) = N(k,n) \cdot P(k,\tau) \;, \tag{1}$$

that is the product of two terms: a number of tests

$$N(k,n) = 3(n-7)\binom{n}{k}\binom{k}{7} \;, \tag{2}$$

which counts all possible couples (F,S) that can be considered, and an upper bound $P(k,\tau)$ of the probability that the maximal reprojection error of the k estimated inliers remains below τ. Assuming a uniform distribution of the image points, this upper bound can be chosen as

$$P(k,\tau) = (\alpha_0 \tau)^{k-7} \;, \tag{3}$$

where α_0 is an upper bound for the probability of a random point to have reprojection error 1 (obtained by considering the relative image area of a band of width 2 around the image diagonal), and τ is the normalized error threshold that selects the k inliers (in practice, the k-th smallest error). The form of (1) is typical of *a contrario* models (see [2,6]): it guarantees that for any $\varepsilon > 0$, the expected number of sets S that pass the test $NFA(S) < \varepsilon$ is, for random data,

less than ε. In practice, one often chooses the value $\varepsilon = 1$, so that only one false alarm is expected, on average, for random data.

Once 7 point matches have been used to compute a fundamental matrix, there remain $n-7$ point matches to test, or, more exactly, all 2^{n-7} possible subsets of these $n-7$ point matches. Hopefully, we do not need to test these 2^{n-7} possible subsets, but only $n-7$ subsets. The reason is simple: for a given cardinal k, the criterion used (NFA) is an increasing function of the maximum error τ, so the best subset with cardinal k is simply the one made of the k point matches that have the smallest error. Therefore, by sorting the matches by increasing errors, we can find the minimum NFA over the 2^{n-7} possible subsets by looking only at $n-7$ subsets (one for each possible value of k).

Algorithm 1 The ORSA Algorithm [11] to minimize the NFA (1)

Require: W the set of n potential matches, N the maximum number of trials
1: $U := W$; $NFA^* := \infty$; $S^* := \emptyset$; $opt := 0$; $iter := 1$; $MaxIter := N$
2: **while** $iter \leq MaxIter$ **do**
3: Randomly sample a set T of 7-point matches from U
4: **for all** Fundamental matrix F associated to T **do**
5: Sort the matches by increasing errors
6: $S(F) :=$ set with minimum NFA in $\{S_k(F) = \text{first } k \text{ matches}\}_{7<k\leq n}$;
7: **if** $NFA(S(F)) < NFA^*$ **then**
8: $NFA^* := NFA(S(F))$; $S^* := S(F)$
9: **if** $opt = 0$ **and** ($NFA^* < 1$ **or** $iter > 0.9N$) **then**
10: $opt := 1$; $MaxIter := (iter + 0.1N)$; $U := S^*$
11: $iter := iter + 1$
12: **return** Best inlier set S^* and associated fundamental matrix F

If we have k inliers among n matches, the number of samples to be drawn to have a probability q of selecting an inlier minimal sample of matches is [11]

$$N = \log(1-q)/\log\left(1 - \prod_{i=0}^{6} \frac{k-i}{n-i}\right). \tag{4}$$

In practice, ORSA considers that a valid fundamental matrix has been found if a NFA value below 1 is obtained (line 9 in Algorithm 1). In general, this event will occur much before the predicted number of iterations (4), either because the sampled subset contains only inliers or because it contains a majority of inliers and the outliers are close to consistent with the underlying geometry.

3 Using an Empirical Feature Distribution

We now try to improve the ORSA method by testing different empirical distributions for the image points. Since we are interested in using the empirical distribution on the image to detect matches between images, the computed matches

cannot be used for the distribution estimation. Instead, we will only use the points in the second image and model their 2D distribution using a mixture of Gaussian kernels centered at the data points. Their covariance matrix is modeled as **isotropic** (a diagonal matrix with a single parameter σ) or **anisotropic** (a symmetric positive definite 2×2 matrix H).

3.1 Method Overview

Given a set of putative matches, we first estimate the 2D density of the image points in the second image (Section 3.2) and then pre-compute the integrated density in any domain delimited by the image frame and a segment whose endpoints lie on the image frame (Section 3.3). Using a random sampling procedure, for each tested model we sort the matching probabilities (instead of the matching errors), and then apply the *a contrario* methodology (Section 3.4). The computation of such probabilities is fast thanks to the pre-computation mentioned above. We use the same sampling strategy as in ORSA: after a first significant inlier set is found or after 90% of the total iterations have been made, the procedure is restarted for 10% of the iterations, with new samples drawn only from the best inlier set found so far.

3.2 Kernel Density Estimation of the Image Distribution

The first step of our method consists in estimating the density of the points $\{x_1, \ldots, x_n\}$ in the second image. We use Kernel Density Estimation (KDE) for such purpose, meaning that we model this density as a mixture of two-dimensional *kernels* K centered at the image points:

$$f(x) = \frac{1}{n|H|} \sum K\left(H^{-1}(x - x_i)\right) \ . \tag{5}$$

The matrix H is known as *bandwidth* and it is the smoothing parameter of the density estimator. We recommend the paper [15] for a complete description of KDEs in the 1D case, from which the 2D case is just a generalization.

Among the several possible kernel functions K, we chose for simplicity reasons the Gaussian density function. In order to have a fully non-parametric density estimation, and following the final recommendations from [15], we decided to use 2D plug-in methods for the automatic selection of the bandwidth matrix H. We discarded cross-validation due to the possibility of having few data points.

Plug-in methods search for the estimator with minimum asymptotic mean integrated squared error (AMISE) by replacing (or pluging-in) an unknown high order term in its formula by an approximation made by normal reference. We chose two different parametric forms for the bandwidth matrix:

1. **Hiso** $= h \cdot$ **Id,** isotropic bandwidth. In practice, the 1D parameter h is estimated by averaging the 1D plug-in estimations (pilot estimation of derivatives) [16] on a set of radially equidistributed 2D directions, and then applying a dimension change factor equal to $n^{1/30}$.

2. **Hani** a full 2×2 symmetric positive definite matrix, anisotropic bandwidth, which we estimate in practice using a 2D plug-in method.

Instead of implementing the corresponding 1D and 2D plug-in methods, we used the following R functions[5] from the ks (kernel smoothing) package[6]: bw.SJ() and Hpi(). The R functions can be embedded into C++ using RInside[7].

3.3 Line Distribution Pre-computation

The method of Section 3.1 involves the repeated computation of the probabilities associated to the residual errors. That is, for each error $\varepsilon = \text{dist}(x', Fx)$ associated to a correspondence (x, x'), we need to integrate the density map on the line band $\{z; d(z, Fx) \leq \varepsilon\}$ to obtain the probability $\text{Prob}\{\text{dist}(Z, Fx) \leq \varepsilon\}$ (Z being a random variable following the distribution of Section 3.2). It can be computed as the difference of the integrated densities in two domains delimited by the image frame and one boundary of the band, which is the reason why we pre-compute these sums. In practice, we consider only domains delimited by segments whose endpoints have integer coordinates. Thus, any epipolar line is approximated to the closest line in this set to estimate the associated probability.

3.4 *A Contrario* Step

We follow the main *a contrario* framework of the ORSA method, based on the minimization of the number of false alarms NFA in (1). However, we need to adapt this framework to take into account that we now consider a more complex (that is, non-uniform) a priori distribution for the points of the second image.

Instead of measuring the max error τ of a set S of k potential inliers, we now decide to measure the max probability

$$\delta(S) = \max_{(x,x') \in S} G_{Fx}\left(\text{dist}(x', Fx)\right), \tag{6}$$

where for any line D of the second image, $G_D(\tau)$ represents the probability that a point Z drawn according to the distribution of Section 3.2 falls below a distance τ from D, that is,

$$G_D(\tau) = \text{Prob}\left\{\text{dist}(Z, D) \leq \tau\right\}. \tag{7}$$

Now, remarking that one has $\text{Prob}\left\{G_D(\text{dist}(Z, D)) \leq \delta\right\} \leq \delta$ (use, e.g., [6] Lemma 1 with $X = -\text{dist}(Z, D)$), it follows that the NFA property (control of the number of false alarms in case of random matches) is preserved provided that we replace, for any set S of k matches, the formula (1) of ORSA by

$$\text{NFA}(S) = N(k, n) \cdot Q(k, \delta(S)) \quad \text{with} \quad Q(k, \delta) = \delta^{k-7}. \tag{8}$$

[5] `http://www.r-project.org/`
[6] `http://cran.r-project.org/web/packages/ks/`
[7] `http://dirk.eddelbuettel.com/code/rinside.html`

4 Error Measures

4.1 Optimal Reprojection Error

In order to evaluate a given fundamental matrix F and a set of n matches $(x_i, x'_i)_{1\leq i\leq n}$ we use the implementation of Kanatani [8] [8] to compute the optimal matches for the given data; we use as measure the RMSE of the reprojection error associated to these optimal matches.

4.2 Comparing a Fundamental Matrix with Ground Truth

Given a groundtruth fundamental matrix F_0, we generate a set of 500 exact matches under F_0, their coordinates on each image being either uniformly distributed or following a local distribution around the groundtruth inliers. Then, we evaluate a new fundamental matrix F by computing the RMSE error (in pixels) of these matches under F. The generation of uniformly distributed matches was done following [19], where this procedure was used. We denote by $\mathbf{Err_1}$ the evaluation of the RMSE Gold error (in pixels) on those "global" matches.

In order to generate matches according to estimated densities, we first generated random features on the left image from a mixture of Gaussian kernels using the R function rmvnorm.mixt(). Then, for each of these features, we used the groundtruth calibration to obtain an epipolar line on the right image. We took 10 equidistributed points on each line and used the density of the second image on those points (re-weighted to sum 1) for generating a point on the line, giving in most cases an outlier correspondence. We denote by $\mathbf{Err_2}$ the evaluation of the RMSE Gold error (in pixels) on these "local" matches.

4.3 Comparing a Relative Camera Motion with a Groundtruth Reconstruction

Assume as known a groundtruth camera reconstruction (camera internal and external camera parameters),

$$P_0 = K_0 R_o^T(\mathrm{Id}, -C_0)\ , \quad P_1 = K_1 R_1^T(\mathrm{Id}, -C_1)\ , \tag{9}$$

Then, given a fundamental matrix F, we can compute the essential matrix using K_0, K_1 and so the relative camera motion R, t. We evaluate the error of camera motion with the help of the best similarity H taking the groundtruth reconstruction to the following one:

$$Q_0 = K_0(\mathrm{Id}, 0)\ , \quad Q_1 = K_1(R, t)\ . \tag{10}$$

Two measures are proposed:

1. *Rotational error (in degrees).* The distance between R_0^T and $R^T R_1^T$, i.e. the rotation angle (in degrees) of $R^T R_1^T R_0$;
2. *Translational error (in %).* The absolute error between the similarity-corrected camera centers and the groundtruth ones, divided by the absolute distance between the groundtruth camera centers, and multiplied by 100.

[8] Source code available at `http://www.iim.ics.tut.ac.jp/~sugaya/public-e.html`

5 Evaluation

5.1 Simulation

The proposed methods (Hiso, Hani) can compare favorably to ORSA in the case where the image matches are concentrated in a small part of the image and, of course, there exist no matches compatible with false epipolar geometries.

In order to exemplify this assumption, we used images with groundtruth calibration available and took 100 matches with Gold error smaller than 1 pixel; then, we estimated their density (as explained in Section 3.2) and then sampled outlier matches from that density (using the R function rmvnorm.mixt()) to create an input dataset with an approximately controlled inlier ratio. We tested the following inlier ratios: 75%, 25%, 15% and 10%; in all cases, the approximate number of inliers was 100, and the final dataset size was inversely proportional to the inlier ratio (sample input datasets are shown in Figure 1).

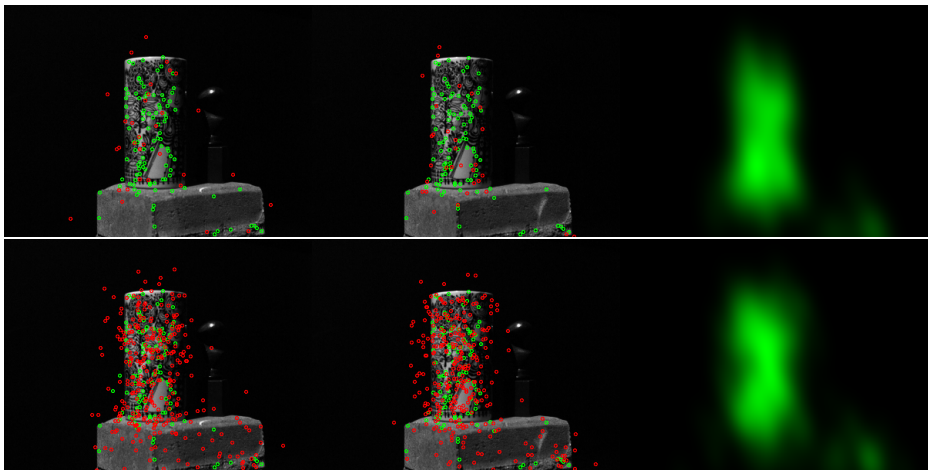

Fig. 1. Example simulation input datasets with inlier ratios 0.75 (top) and 0.25 (bottom). We show the two images with the data matches (green for inlier and red for outlier), and the estimated density (green scale). Groundtruth calibration is available at http://roboimagedata.imm.dtu.dk/data.html.

For each inlier ratio we simulated a dataset, and using it we ran each tested method 200 times and computed the average results (Table 1). Whereas the methods perform similarly for 0.75 and 0.25 inlier ratios, ORSA considers as inliers a big set of outliers for smaller inlier ratios. The other two methods improve ORSA's performance, mainly due to the localization of matches in a small image region; a slightly better performance of Hiso can be explained by the fact that the data was simulated using isotropic kernels.

For the extreme inlier ratio 0.10, our proposed methods did not always manage to find a solution: Hiso succeeded 108/200 times, and Hani 123/200; the

ORSA method always returned a solution as valid, which in most cases turned out to be wrong. In some of the successful cases, the obtained fundamental matrices were not compatible with the groundtruth internal calibration matrices and thus the camera motion could not be extracted. This happened 57/200 times for ORSA with 0.15 inlier ratio, and for the 0.10 inlier ratio it happened 42/200 times for ORSA and 2/123 times for Hani. In conclusion, the proposed methods returned a consistent solution in case of success.

Rat	**Met**	**#it**	**Thr**	**#in**	RMSE	Err_1	Err_2	R_{err}	T_{err}
0.75	ORSA	1001.03	0.55	94.24	0.11	2.35	0.83	2.48	1.31
	Hiso	1001.25	0.84	98.67	0.19	2.33	0.84	2.86	1.53
	Hani	1001.41	0.56	98.95	0.20	3.75	1.34	3.81	1.49
0.25	ORSA	1001.05	0.54	93.94	0.11	2.19	0.79	2.53	1.26
	Hiso	1101.48	0.54	103.53	0.28	3.48	1.13	4.20	1.57
	Hani	1115.51	0.59	102.99	0.21	3.56	1.59	4.44	1.76
0.15	ORSA	1001.88	40.92	533.77	11.98	36.06	14.40	36.79	16.16
	Hiso	3578.09	0.51	108.94	0.46	2.99	1.01	3.59	2.00
	Hani	4896.33	0.46	103.80	0.66	5.02	1.83	5.50	1.97
0.10	ORSA	1001.54	36.87	763.74	11.18	37.10	15.42	125.99	14.22
	Hiso	9166.88	1.57	92.98	1.08	8.13	4.12	13.95	5.70
	Hani	9308.27	1.08	101.16	1.35	9.61	4.91	15.71	5.14

Table 1. Average results over 200 runs with $S = 10000$ maximal trials each for the tested methods: ORSA, Hiso and Hani (Section 3.2). For the different inlier ratios (**Rat**), we show the (average) number of iterations (**#it**), chosen threshold (**Thr**), RMSE optimal reprojection error of the found inliers (Section 4.1), the evaluation errors Err_1, Err_2 (in pixels, Section 4.2), and the relative motion errors R_{err} (in degrees), and T_{err} (in %), explained in Section 4.3.

5.2 Real Images

We used images with groundtruth calibration available and computed their SIFT matches taking as valid a match such that the ratio of scores to the following candidate match was smaller than 0.9. This parameter is usually set to 0.6, but with our selection we are able to obtain much more matches and, interestingly, we are able to obtain smaller final errors (independently of the selected method).

We used the first images in the Strecha's castle-P19 sequence, available at `http://cvlabwww.epfl.ch/data/multiview/denseMVS.html`. The selected image pairs are very different (see Figure 2) in terms of inlier ratios and feature distribution. We provide in the first column of Table 2 a rough estimation (*Rat**) of the inlier ratio, which was obtained by taking as inlier any correspondence with error (w.r.t. the groundtruth fundamental matrix) below 3 pixels; this value is just a proxy for the inlier ratio (an unclear concept when dealing with real data) and it was not used for further evaluation.

We tried to reduce the effect of random sampling by averaging the results over 200 runs of each method. We show the main output given by the considered methods in columns 3 to 6 in Table 2. We added to RANSAC the sampling

strategy from ORSA (see beginning of Section 2), so that the number of iterations was comparable. Note that the optimization criteria are different for each method, and so the thresholds are hardly comparable.

Evaluation errors (see Section 4) are shown in the last 4 columns of Table 2. Pair (0,5) is not reasonably solved on average by any of the methods, although occasionally all of them found a good solution. We highlight in bold the best results among the three non-parametric methods (ORSA, Hiso, Hani) excluding RANSAC, but at the same time we keep the RANSAC output in the table for the reader to see the sometimes dramatic effect of parameter selection. According to our preliminary experiments, ORSA gives slightly better average results than our proposals for medium-to-high inlier ratios, being worst on average otherwise, depending on the feature distribution – see for instance results on pair (0,4).

Fig. 2. Pairs (0,1) on top and (0,4) on bottom of the used Strecha's Castle-P19 sequence. For visualization purposes, the data matches were classified as inliers (green) and outliers (red) using an arbitrary threshold of 3 pixels on groundtruth errors.

6 Conclusion

The two proposed non-parametric robust methods for computation of the fundamental matrix consist in modifying the ORSA method by using an empirical reference distribution of the image features. Preliminary experiments (only a sample of them are included in this paper) show that these methods improve ORSA for low inlier ratios in terms of reprojection error and relative motion precision. Incorporating chromatic information (see e.g. [14]) could lead to further improvement.

castle-P19	**Method**	**#it**	**Thr**	**#in**	RMSE	Err_1	Err_2	R_{err}	T_{err}
Pair (0,1)	RANSAC	1001.6	1.00	1550.13	0.14	1.06	0.46	0.14	0.26
2570 matches		1001.3	2.00	1628.81	0.23	2.29	0.96	0.26	0.58
Rat* = 0.66		1001.3	3.00	1659.90	0.36	5.61	2.35	0.35	1.32
	ORSA	1001.1	0.70	1493.42	0.12	**0.97**	**0.39**	**0.11**	**0.25**
	Hiso	1001.1	1.20	1588.92	0.62	1.45	0.50	0.20	0.44
	Hani	1001.1	1.37	1585.06	0.59	1.64	0.59	0.21	0.49
Pair (0,2)	RANSAC	1002.3	1.00	1099.76	0.16	1.93	0.54	0.27	1.04
2112 matches		1001.8	2.00	1169.30	0.24	1.91	0.57	0.32	0.92
Rat* = 0.58		1001.6	3.00	1194.65	0.31	1.90	0.55	0.38	0.95
	ORSA	1001.1	0.77	1067.75	0.14	1.80	0.48	**0.26**	**0.96**
	Hiso	1001.2	0.99	1122.86	1.24	1.89	0.45	0.35	0.98
	Hani	1001.3	0.94	1115.74	1.16	**1.65**	**0.45**	0.33	1.00
Pair (0,3)	RANSAC	1002.8	1.00	810.10	0.18	15.34	2.34	2.68	9.30
1818 matches		1001.9	2.00	918.29	0.27	3.05	0.59	0.59	2.15
Rat* = 0.53		1001.7	3.00	947.54	0.35	3.41	0.68	0.64	2.48
	ORSA	1001.2	1.07	833.39	0.19	13.22	2.55	2.36	7.44
	Hiso	1001.4	18.52	923.50	1.28	**3.05**	0.52	**0.61**	**2.51**
	Hani	1001.4	1.77	918.85	1.01	3.47	**0.51**	0.62	2.97
Pair (0,4)	RANSAC	9984.2	1.00	62.15	21.12	39.38	10.18	6.13	18.60
1070 matches		9716.8	2.00	132.41	1.84	11.57	2.49	1.80	6.54
Rat* = 0.15		7962.1	3.00	146.35	0.52	15.60	3.77	2.68	7.97
	ORSA	1028.7	0.16	26.14	0.03	127.53	92.11	115.76	37.69
	Hiso	1086.0	8.35	172.69	8.61	25.09	9.58	17.90	11.15
	Hani	1100.4	6.91	144.67	8.36	**13.63**	**3.63**	**3.50**	**4.83**
Pair (0,5)	RANSAC	10000.0	1.00	45.26	31.67	104.97	101.47	93.21	44.33
1055 matches		10000.0	2.00	58.54	31.23	100.78	97.75	94.34	41.66
Rat* = 0.10		10000.0	3.00	67.31	29.68	100.33	77.98	86.98	43.35
	ORSA	1030.3	0.01	15.50	0.00	137.34	126.96	128.27	41.89
	Hiso	1134.1	43.37	91.33	2.23	187.73	225.49	133.83	63.37
	Hani	1589.3	13.62	125.43	3.43	110.73	113.46	93.45	41.14

Table 2. Average results on Strecha's castle-P19 data over 200 runs with $S = 10000$ maximal trials each for the tested methods; *Rat** being a rough estimation of the inlier ratio (see text). For each pair and method, we show the (average) number of iterations (**#it**), the threshold (**Thr**) either being a RANSAC's input or a method's output, RMSE optimal reprojection error of the found inliers (Section 4.1), the evaluation errors Err_1, Err_2 (in pixels, Section 4.2), and the relative motion errors R_{err} (in degrees), and T_{err} (in %), explained in Section 4.3.

Acknowledgments. Part of this work was carried out in IMAGINE, a joint research project between École des Ponts ParisTech (ENPC) and the Scientific and Technical Centre for Building (CSTB) and was funded by the Agence Nationale de la Recherche, Callisto project (ANR-09-CORD-003). Pedro Delicado, Universitat Politècnica de Catalunya, is acknowledged for his advice on non-parametric density estimation.

References

1. O. Chum, T. Werner, and J. Matas. Two-view geometry estimation unaffected by a dominant plane. In *Proc. CVPR*, 2005.
2. A. Desolneux, L. Moisan, J.-M. Morel. *From Gestalt Theory to Image Analysis. A Probabilistic Approach.* Springer-Verlag, collection "Interdisciplinary Applied Mathematics", vol. 34, 2008.
3. M. A. Fischler and R. C. Bolles. Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography. *Comm. ACM*, 24(6):381–395, 1981.
4. J.-M. Frahm and M. Pollefeys. Ransac for (quasi-)degenerate data (QDEGSAC). In *Proc. CVPR*, 2006.
5. L. Goshen and I. Shimshoni. Balanced exploration and exploitation model search for eficient epipolar geometry estimations. In *Proc. ECCV*, 2006.
6. B. Grosjean and L. Moisan. A-contrario detectability of spots in textured backgrounds. *J. Math. Imag. Vis.* 33(3):313–337, 2009.
7. R. Hartley and A. Zisserman. *Multiple View Geometry in Computer Vision.* Cambridge University Press, second edition, 2004.
8. K. Kanatani, Y. Sugaya, and H. Niitsuma. Triangulation from two views revisited: Hartley-Sturm vs. optimal correction. In *Proc. BMVC*, 2008.
9. K. Kanatani and Y. Sugaya. Fundamental matrix computation: Theory and practice. *Memoirs of the Faculty of Engineering, Okayama University*, 42:18–35, 2008.
10. D.G. Lowe. Distinctive image features from scale-invariant keypoints. *Int. J. Comput. Vis.*, 60(2):91–110, 2004.
11. L. Moisan and B. Stival. A probabilistic criterion to detect rigid point matches between two images and estimate the fundamental matrix. *Int. J. Comput. Vis.*, 57(3):201–218, 2004.
12. L. Moisan, P. Moulon and P .Monasse. Automatic Homographic Registration of a Pair of Images, with A Contrario Elimination of Outliers. *IPOL*, 2012.
13. P. Moulon, P. Monasse, and R. Marlet. Adaptive structure from motion with a contrario model estimation. In *Proc. ACCV*, 2012.
14. N. Noury, F. Sur, and M.-O. Berger. Fundamental matrix estimation without prior match In *Proc. ICIP*, 2007.
15. S. J. Sheather. Density estimation. *Stat. Sci.*, 19(4):588–597, 2004.
16. S. J. Sheather and M. C. Jones. A reliable data-based bandwidth selection method for kernel density estimation. *J. R. Stat. Soc. (Series B)*, 53:683–690, 1991.
17. P.H.S. Torr and A. Zisserman. MLESAC: A new robust estimator with application to estimating image geometry. *Comput. Vis. Image Understand.*, 78:138–156, 2000.
18. P.H.S. Torr, A. Zisserman, and S.J. Maybank. Robust detection of degenerate configurations while estimating the fundamental matrix. *Comput. Vis. Image Understand.*, 71(3):312–333, 1998.
19. Z. Zhang. Determining the epipolar geometry and its uncertainty: a review. *Int. J. Comput. Vis.*, 27(2):161–195, 1998.