A Probabilistic Criterion to Detect Rigid Point Matches Between Two Images and Estimate the Fundamental Matrix

LIONEL MOISAN AND BÉRENGER STIVAL
CMLA, Ecole Normale Supérieure de Cachan, 61 avenue du président Wilson, 94235 Cachan cedex, France
moisan@cmla.ens-cachan.fr
stival@cmla.ens-cachan.fr

Abstract. The perspective projections of n physical points on two views (stereovision) are constrained as soon as $n \geq 8$. However, to prove in practice the existence of a rigid motion between two images, more than 8 point matches are desirable in order to compensate for the limited accuracy of the matches. In this paper, we propose a computational definition of rigidity and a probabilistic criterion to rate the meaningfulness of a rigid set as a function of both the number of pairs of points (n) and the accuracy of the matches. This criterion yields an objective way to compare, say, precise matches of a few points and approximate matches of a lot of points. It gives a yes/no answer to the question: "could this rigid points correspondence have occurred by chance?", since it guarantees that the expected number of meaningful rigid sets found by chance in a random distribution of points is as small as desired. It also yields absolute accuracy requirements for rigidity detection in the case of non-matched points, and optimal values of n, depending on the expected accuracy of the matches and on the proportion of outliers. We use it to build an optimized random sampling algorithm that is able to detect a rigid motion and estimate the fundamental matrix when the set of point matches contains up to 90% of outliers, which outperforms the best currently known methods like M-estimators, LMedS, classical RANSAC and Tensor Voting.

Keywords: stereovision, point matches, structure from motion, rigidity detection, fundamental matrix, meaningful event

1. Introduction

Stereovision, and more generally structure from motion, is based on the fact that two images of a static scene taken by a pinhole camera from two different viewpoints are redundant. A classical approach to use this redundancy is to find the position of several physical parts of the scene (typically points, or lines) in each image. This feature matching approach allows to recover, up to a scale factor, not only the 3D position of the points before their projection in each image, but also the camera motion and its intrinsic parameters (Longuet-Higgins, 1981; Faugeras et al., 1987; Hartley, 1997). All these methods use the fact that two per-spective projections of a static scene are related by the so-called *epipolar geometry*, which is itself described by the fundamental matrix. This outlines the importance of estimating the fundamental matrix from point matches, and explains the large attention it received in the last two decades (Faugeras and Luong, 2001; Luong and Faugeras, 1996; Oliensis and Genc, 2001; Salvi et al., 2001; Torr and Murray, 1997; Zhang et al., 2001; Zhang, 1998).

The epipolar geometry relies on a rigidity constraint. Indeed, in the camera referential, changing the point of view amounts to apply an Euclidean motion (3D rotation and translation) to the scene, or, in other words, to move it like a rigid object. Thus, stereovision is a

particular case of rigid motion estimation, the latter being more general since several rigid motions may occur at the same time.

In this paper, we propose a probabilistic criterion to detect the existence of a rigid motion between two sets of point matches, that permits to decide whether these points are independent or if they are correlated by a rigid motion. Since a rigid motion only involves a small number of parameters (a 3D translation followed by a 3D rotation), a large number of point correspondences between the two images cannot in general be explained by such a simple transform, so that true point correspondences can be detected as large deviations from randomness.

Detecting the existence of a rigid motion between two images has several interesting applications, like e.g. the detection and segmentation of rigid moving objects. Our approach can also be used for stereovision and structure from motion, since it provides a very efficient and robust way to estimate the fundamental matrix from a set of point matches.

Given a set of point matches, we define its rigidity as the least maximum epipolar distance achieved for all epipolar geometries defined from 7 sub-pairs. Then, we define the meaningfulness of a set as the expected number of similar sets (that is, having the same size and at least the same rigidity) in a random uniform distribution of points. Such a detection against a random uniform assumption had been proposed by Lowe (1985) and Stewart (1995), before a systematic simple formulation, based on expectation (instead of probability), was developed in Desolneux (2000). In the case we consider here, the criterion we obtain not only gives a yes/no answer to the question "could it have happened by chance that this set is so rigid?", but also provides a way to compare the rigidity of sets *with different sizes*. Indeed, it combines the measurement of several parameters (relative and total number of points, rigidity, etc.) into a single intuitive "expected number of false alarms". Such a parameter reduction has important consequences on the identification of outliers in the fundamental matrix estimation problem. Compared to McReynolds and Lowe's approach (McReynolds and Lowe, 1996), our method has several advantages. Among them are the fact that it has zero parameter (the algorithm that computes the rigidity is not iterative and does not require multiple thresholds to define a "stopping criterion for convergence" (McReynolds and Lowe, 1996) and the fact that it does not require any assumption on the

camera motion or preliminary estimates of the noise variance.

As we mentioned above, the criterion we define allows to build an efficient algorithm for robust estimation of the fundamental matrix. For a small number n of point matches (say $n \leq 25$), a deterministic algorithm (systematic search) may be used, but for the general case we use a stochastic algorithm following the random sampling consensus (RANSAC), introduced in the context of image analysis by Fischler and Bolles (1981). The rigidity detection criterion allows to dramatically improve the performance of classical RANSAC algorithms. Our approach permits to detect rigidity and to provide a good estimate of the fundamental matrix when the initial set of point matches contains up to 90% of outliers. This outperforms the best currently known algorithms like M-estimators, LMedS and RANSAC (Salvi et al., 2001; Torr and Murray, 1999), which typically break down around 50% of outliers. As far as we know, even the recent Tensor Voting technique (Tang et al., 2001) has not been reported to work with more than 65% of outliers.

This paper is organized as follows. In Section 2, we recall some basics of epipolar geometry. In Section 3, we give a computational definition of rigidity, before we derive probabilistic criteria to rate the meaningfulness of a given rigid set of point matches. We also discuss the case of non-matched points, and prove the existence of an absolute accuracy requirement for rigidity detection, and of an optimal value of n, depending on the accuracy of the matches and on the proportion of outliers. In Section 4, we describe the deterministic and stochastic algorithms mentioned above, then we present some experiments in Section 5.

2. Epipolar Geometry

2.1. *The Epipolar Constraint*

Epipolar geometry describes in an elegant way the constraints satisfied by two perspective projections $m = (xy)^T$ and $m' = (x'y')^T$ of a physical 3D point M on two different images. For pinhole cameras (with known or unknown internal parameters), the relation

$$(x' \, y' \, 1)F(x \, y \, 1)^T = 0 \qquad (1)$$

holds, where F, the fundamental matrix, is a 3×3 matrix with rank 2 that depends on the rigid motion between the two image planes and on the camera parameters

(position of the optical center, and relative pixel size along both axes). Equation (1) has the following interpretation: the point m' of the second image plane belongs to the epipolar line Fm whose equation $ax' + by' + c = 0$ is obtained from $m = (x, y)$ by

$$(a\, b\, c)^T = F(x\ y\ 1)^T.$$

This epipolar line is nothing but the projection in the second image of the optical ray going from M to the optical center of the first camera. The position of m' on this line is related to the (unknown) depth of M, which is not involved in (1). The intersection of all epipolar lines is called the epipole: it is the projection in the second image of the optical center of the first camera.

The fundamental matrix can be factorized under the form $F = C'EC^{-1}$, where C and C' are 3×3 internal calibration matrices depending on the cameras only and E is the essential matrix defined from the motion parameters $T = (t_x\, t_y\, t_z)^T$ (3D translation vector) and R (3D rotation matrix) by

$$E = \begin{pmatrix} 0 & -t_z & t_y \\ t_z & 0 & -t_x \\ -t_y & t_x & 0 \end{pmatrix} R.$$

The essential matrix has, like F, a null singular value, but in addition its two other singular values are equal (see Faugeras and Luong, 2001).

2.2. Estimating the Fundamental Matrix from Point Matches

For cameras with calibrated internal parameters, only the essential matrix has to be recovered and the epipolar geometry depends on 5 parameters: 2 for the 3D translation (the unknown scale factor lets us impose $\|T\| = 1$ for example), and 3 for the 3D rotation (matrix R). Since each point correspondence produces one equation like (1), a minimum of 5 matchings is required to recover the epipolar geometry. As shown by Faugeras (2001), 5 matchings are sufficient but the computation is rather complex (and unstable) and provides in general 10 solutions.

For that reason, and because more points are needed in the case of cameras with uncalibrated internal parameters, the much simpler linear 8-points algorithm is preferred in general. Introduced by Longuet-Higgins (1981), this algorithm has been refined in several ways. The classical formulation is the following: for each point match (m_i, m'_i), Eq. (1) can be rewritten as a linear equation depending on the 9 coefficients of F, viewed as a 9-dimensional vector f. Hence, 8 point correspondences yield a 8×9 linear system

$$Af = 0,$$

where the i-th line of A is obtained from the coordinates of m_i and m'_i (as noticed by Hartley (1997), these coordinates must be normalized to improve the conditioning of the matrix A). Since f is defined up to a scale factor, this system permits to recover f. This can be done by looking for a unit norm vector in the (generally 1-dimensional) null space of A. This method presents a small drawback: the solution F obtained is not a rank 2 matrix in general, which means that the "epipolar geometry" underlying F may not be consistent with a physical realization (epipolar lines may not all intersect at the epipole). This problem can be solved by using the Singular Value Decomposition of F and forcing the smallest eigenvalue to 0, or with some kind of parallax estimation (see Boufama and Mohr, 1995).

The **7-points algorithm** avoids this issue by considering only 7 point matches while enforcing the rank 2 constraint. The 7×9 linear system $Af = 0$ obtained with 7 point correspondences yields in general a 2-dimensional space of solutions. If (F_1, F_2) is a basis of this space (whose elements are now written as 3×3 matrices), then the fundamental matrix is obtained as a solution $F = \alpha F_1 + (1 - \alpha)F_2$ of $\det F = 0$. This condition requires, in order to determine α, the computation of the roots of a real cubic polynomial, and may yield 1 or 3 solutions. The fundamental matrices obtained with this method define a true epipolar geometry, where all epipolar lines pass through a unique point (the epipole) while each of the 7 point correspondences are realized exactly.

When dealing with more than 7 or 8 points, several methods exist. The classical Least Square Minimization (LMS) looks for F that minimizes $\sum_i (m'^T_i F m_i)^2$. The solution, easily computed using liner algebra, can then be improved by minimizing a more geometric criterion (Luong and Faugeras, 1996), the sum of the squared distances from each point to the corresponding epipolar line, $\sum_i r_i^2$, where $r_i = \text{dist}(m'_i, Fm_i)$. This criterion can be symmetrized into $\sum_i r_i^2 + r'^2_i$, with $r'_i = \text{dist}(m_i, F^T m'_i)$. These methods, and similar ones, heavily suffer from badly located points or false matches. This is the reason why robust methods have been introduced, that are able to reject outliers in the optimization process.

2.3.   Robust Methods

Among the most efficient robust methods are M-estimators, LMedS and RANSAC (Salvi et al., 2001; Torr and Murray, 1997) and the recent Tensor Voting (Tang et al., 2001). M-estimators (see Zhang et al., 1994 for example) try to reduce the influence of outliers by applying an appropriate weighting of the distances. The criterion to minimize becomes $\sum_i \rho(r_i)$, where ρ grows more slowly than the square function. As pointed out by Faugeras and Luong, "*it seems difficult to select a ρ-function for general use without being rather arbitrary*" (Faugeras and Luong, 2001). RANSAC methods use a set of fundamental matrices obtained from random sets of 7 point matches (with the 7-points algorithm), to find the largest consensus set, that is the largest set of point matches achieving a given accuracy of the matches. Like M-estimators, RANSAC suffers from an arbitrary choice, since an accuracy threshold has to be preset. The Tensor Voting method, which relies on the scalar selection of a 8-dimensional hyperplane, suffers from the same kind of problem.

We shall see that the criterion we propose solves this difficulty by giving a way to combine the set size and the matching accuracy into a single value, which in addition has an absolute meaning. The algorithm we describe in Section 4 is a generalization of the RANSAC method that takes advantage of this absolute criterion.

3.   A Criterion for Rigidity Detection

3.1.   Measuring Rigidity

The question we ask is the following: let $m_i = (x_i, y_i)$ and $m'_i = (x'_i, y'_i)$ be the n points found in the two images, how can we decide that each pair (m_i, m'_i) represents the projection of a 3D point in two different image planes? If this was exactly true, the epipolar constraint $(x'_i\ y'_i\ 1)F(x_i\ y_i\ 1)^T = 0$ should be satisfied for all i and for some 3×3 matrix F with rank 2, but in practice some location errors will occur and we have to measure how bad a correspondence is. We choose to measure the rigidity of a matching set $S = \{(m_i, m'_i)\}_{i=1..n} \in (I \times I')^n$ (I and I' represent the (convex) domains of the 2 images) with the following physical criterion.

Definition 1.   Let F be a 3×3 matrix with rank 2. The F-rigidity of a set S of point matches between two images is

$$\alpha_F(S) = \frac{2D'}{A'} \max_{(m_i m') \in S} \text{dist}(m', Fm) \qquad (2)$$

where Fm is the epipolar line associated by F to m, dist the Euclidean distance, A' the area of the second image domain and D' its diameter.

We shall justify later the normalization coefficient $\frac{2D'}{A'}$, but we can notice that it yields a scale-independent definition of rigidity. Contrary to classical criteria, we use the l^∞ norm (maximum) instead of the l^2 norm (sum of squares). This has the advantage of being extremely selective against outliers, and simplifies most of the computations that will follow. However, our approach could also be applied to a more usual l^2 criterion by using the Central Limit Theorem. As usual, (2) may be symmetrized by using:

$$\tilde{\alpha}_F(S) = \max_{(m,m') \in S} \max \left(\frac{2D'}{A'} \text{dist}(m', Fm), \frac{2D}{A} \text{dist}(m, F^T m') \right) \qquad (3)$$

(A is the area of the first image domain, D its diameter).

Now, we could define a meaningful point correspondence as a matching set S for which $\alpha_F(S)$ is too small (for some F) to be reasonably explained by randomness. Let us precise this: if we suppose that the points of S are randomly, uniformly, and independently distributed in $(I \times I')^n$, then we can define the probability $q(t) = P[\inf_F \alpha_F(S) \leq t]$. A meaningful point correspondence could then be defined as a matching set S for which $q(\inf_F \alpha_F)$ is very small, say less than 10^{-3}. However, the computation of the *probability q* seems very difficult, and this is why we develop a more feasible approach based on *expectation*, following (Desolneux et al., 2000). Instead of controlling the probability of a false detection (a false detection is a rigid set detected in a random distribution of points), we shall measure the expected number of false detections. In other words, our criterion will be designed in order to ensure that the number of detections for random data is as low as desired. As we shall see, this requires a quantization of the number of tests (candidates for F) performed.

3.2. A Computational Definition of Rigidity

According to the previous remark, the rigidity of a matching set S should be defined by $\inf_F \alpha_F(S)$, that is to say by minimizing the matching error among all possible F. However, this definition is too difficult to manage with and we adopt a more computational point of view.

Definition 2. A matching set $S = \{(m_i, m'_i)\}_{i=1..n}$ is α-rigid if there exists a fundamental matrix F associated to a subset of 7 matchings of S such that $\alpha_F(S) \leq \alpha$.

Note that we could have used as well $\tilde{\alpha}_F(S)$ in this definition. From this definition, the *rigidity* of S could be defined as the least α for which S is α-rigid. This implies the computation of up to $3\binom{n}{7}$ fundamental matrices (since, as we saw previously, the 7-points algorithm may produce up to 3 solutions). The fact that this computation becomes rapidly unfeasible when n becomes larger than 25 (see Table 1) justifies the use of less systematic estimates, allowed by Definition 2.

In comparison with the ideal approach that would consist in estimating the best F among all possibilities (and not only among solutions of the 7-points algorithm), we have reduced the number of possible F in the spirit of the RANSAC paradigm. This is not too restrictive, though, since if there exists a good correspondence between the point matches of S, it is likely that the 7-points algorithm will produce a good approxima-

Table 1. Number of transforms that would be required to define the exact rigidity of a set of n point matches, and the associated computation time on a recent PC laptop (1 GHz). The rapidly increasing computational cost justifies the less systematic Definition 2 of α-rigidity. Our implementation of the 7-points algorithm is able to handle around 10,000 7-points matches (that is, up to 30,000 fundamental matrices) per second (see Section 4).

n	Number of transforms, $3\binom{n}{7}$	Expected computation time
8	24	0.0008 s
10	360	0.012 s
15	19,305	0.64 s
20	232,560	7.7 s
25	1,442,100	48 s
30	6,107,400	3 min 30 s
50	$3 \cdot 10^8$	2h45
100	$4.8 \cdot 10^{10}$	18 days

tion of the "ideal" F (that is, minimizing $\alpha_F(S)$) for some subset of 7 point matches of S. Moreover, this quantization of the possible F will have very important consequences in the next section.

3.3. Meaningful Rigid Sets

3.3.1. The Ideal Case (Checking Rigidity). Now we have given a definition of approximate rigidity, we use the notion of ε-meaningful event introduced in Desolneux et al. (2000) to reduce the set parameters (n, the size of S, and α, the rigidity threshold) to a single intuitive parameter, an expected number of false alarms.

Definition 3. Let $\mathcal{E}$ be a discrete set of random variables taking values in an ordered set. The event "$E \leq t$" is ε-meaningful if its associated number of false alarms, defined by $|\mathcal{E}| \cdot P[E \leq t]$, is less than ε.

This definition ensures that the expected number of ε-meaningful events among a realization of $\mathcal{E}$ is less than ε (in average, we have less than ε false detections). In the following, we will use this simple criterion to define meaningful rigid matching sets. The underlying model will be the following: the points m_i and m'_i are uniformly and independently distributed on I and I'.

Proposition 1. *A matching set S of size $n(n \geq 8)$ is ε-meaningful as soon as it is α-rigid with*

$$\varepsilon_1(\alpha, n) := 3 \cdot \binom{n}{7} \cdot \alpha^{n-7} \leq \varepsilon. \tag{4}$$

The number ε_1 measures the meaningfulness (from the rigidity viewpoint) of a given set of point matches. Hence, Proposition 1 not only provides an intuitive threshold for rigidity detection, but also gives a way to compare the meaningfulness of two sets: since ε_1 measures the expected number of α-rigid sets of size n for random points, the smaller ε_1, the better the accuracy of S is. We can also notice that Proposition 1 remains true when α-rigidity is defined with $\tilde{\alpha}_F$ (instead of α_F) since $\tilde{\alpha}_F \leq \alpha_F$. We keep using α_F in the following because we did not see how to improve significantly the meaningfulness thresholds in the symmetric case (in particular because some F are such that $\tilde{\alpha}_F = \alpha_F$).

Proof of Proposition 1: Let $T \subset S$ be a set of 7 point matches, and F one of the associated rank 2
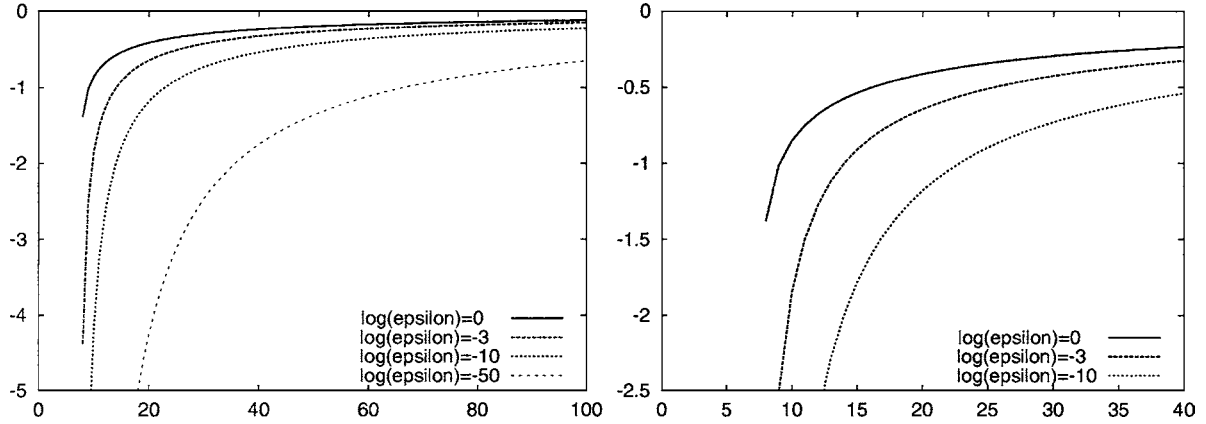
Figure 1. Detectability curves in the $(n, \log \alpha)$ according to Eq. (4) for several meaningfulness values ε_1. A rigidity value of $\log \alpha = -2$ corresponds to a maximum epipolar distance equal to 3.5 pixels on a 1000×1000 image (all logarithms taken in base 10).

fundamental matrix. If (m_i, m_i') belongs to T, then the distance from m_i' to the epipolar line $l = Fm_i$ is zero by construction. If it belongs to $S - T$, then m_i' and l are independent and since the intersection of l with I' has length less than D, the t-dilated of l in I' has area less than $2tD$ and the probability that dist $(m_i', l) \leq t$ is less than $2tD/A$. Consequently, the probability that $\alpha_F(S) \leq \alpha$ is less than α^{n-7}. Now, the number of subsets T of S with 7 elements is $\binom{n}{7}$, and since each of these subsets can produce at most 3 matrices F, we have $|\mathcal{E}| \leq 3 \cdot \binom{n}{7}$. Hence, for every α satisfying (4), any α-rigid set of n point matches is ε-meaningful. $\square$

Equation (4) is very encouraging for the detection of rigidity in the case of already matched points. Indeed, as n grows, the left term goes quickly to zero and a set of point matches is not required to be α-rigid with α very small to be detected as rigid, provided that it contains enough points. The link between α, n and ε is illustrated on Fig. 1. We can see, for example, that even with a low accuracy (3.5 pixels on a 1000×1000 image), 10 points are enough to define a meaningful rigid set ($\varepsilon = 10^{-3}$).

More precisely, when n tends to infinity (4) writes

$$\log \alpha \leq -7\frac{\log n}{n} + \frac{C + \log \varepsilon}{n} + \underset{n \to +\infty}{O}\left(\frac{\log n}{n^2}\right), \quad (5)$$

where $C = \log(7!/3) = 3.22\ldots$ (all logarithms taken in base 10). The proof is given in Appendix.

3.3.2. The Case of Outliers. Up to now, we have studied the meaningfulness of the whole set of point

matches. This is not very realistic since in practice the presence of outliers (badly located points or false matches) cannot be avoided. Moreover, our definition of rigidity, defined as a maximum epipolar distance, is very sensitive to outliers. Hence, we need to consider the case when the set we consider for rigidity detection is not the whole set S of point matches but only a subset of it, S'. Compared to the previous case, we now have to count, in the expected number of false alarms, the number of possible choices for k (that is, $n - 7$) and the number $\binom{n}{k}$ of possible subsets of S with size k.

Proposition 2. *A set $S' \subset S$ of k point matches among n is ε-meaningful as soon as it is α-rigid with*

$$\varepsilon_2(\alpha, n, k) := 3(n-7) \cdot \binom{n}{k} \cdot \binom{k}{7} \cdot \alpha^{k-7} \leq \varepsilon. \quad (6)$$

Figure 2 illustrates the relation between n, $p = 1 - k/n$ and α for $\varepsilon = 10^{-3}$. On Fig. 3, the meaningfulness achieved for fixed values of n and α is presented as a function of p.

As before, we can derive asymptotic estimates from (6). If $p = 1 - k/n$ is the (fixed) proportion of outliers, then assuming $0 < p < 1$ and $n \to \infty$, we can rewrite (6) into

$$\log \alpha \leq -A_p - \frac{15 \log k}{2k} + \frac{C + \log \varepsilon + B_p - 7A_p}{k}$$
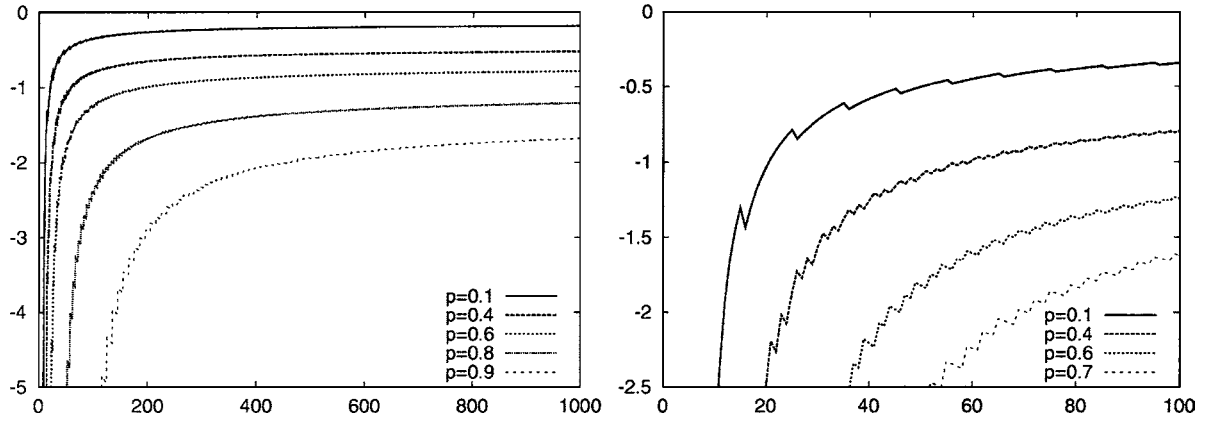$$+ \underset{n \to +\infty}{O}\left(\frac{\log n}{n^2}\right), \quad (7)$$

Figure 2. Detectability curves ($\varepsilon_2 = 10^{-3}$) in the (n, log α) plane for different proportions p of outliers (see Eq. (6)). The curves present small oscillations because $p = 1 - k/n$ cannot be exactly achieved in reason of the discrete nature of k and n (this could be avoided by interpolating the binomial coefficients with the Gamma function).
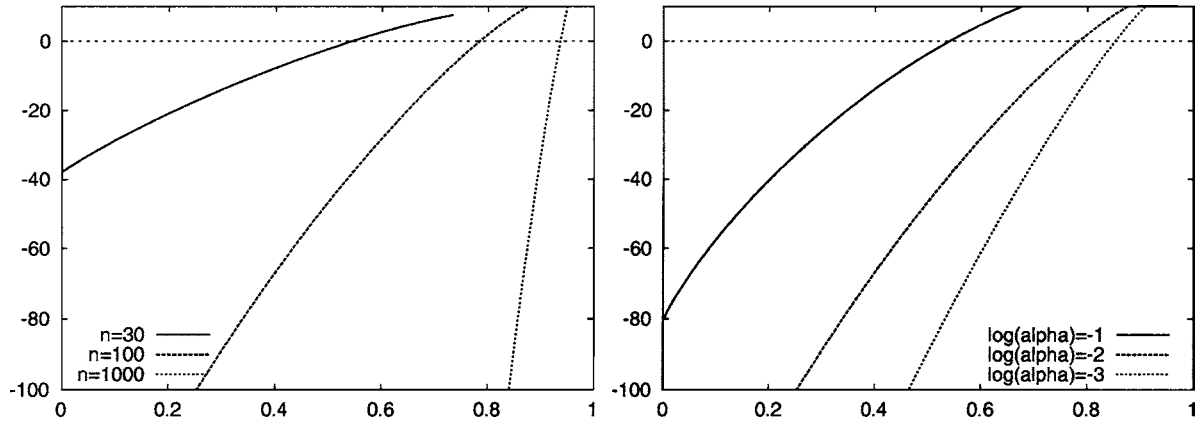

Figure 3. Log-meaningfulness log(ε_2) in function of the proportion p of outliers (Eq. (8)). Left: $\alpha = 10^{-2}$ and n varies. Right: $n = 100$ and α varies.

with $A_p = -\log(1 - p) - \frac{p}{1-p} \log p$ and $B_p = -\frac{1}{2}\log(2\pi p(1 - p)^2)$. This estimate is proven in Appendix. Compared to (5), the main difference lies in the leading term $-A_p$ that forces a fixed gain (depending on p) in the accuracy of the matches. On Fig. 4, we can see that discovering a rigid set among 75% (resp. 90%) of outliers requires a gain of 1 (resp. 1.4) for log α, which comes down to dividing the maximum allowed epipolar distance by 10 (resp. 25).

3.3.3. The Case of Non-Matched Points.

We now consider a less structured situation: we suppose that some characteristic points have been detected on the 2 images but we do not know how to match them together. In this case, we have to take into account the fact that
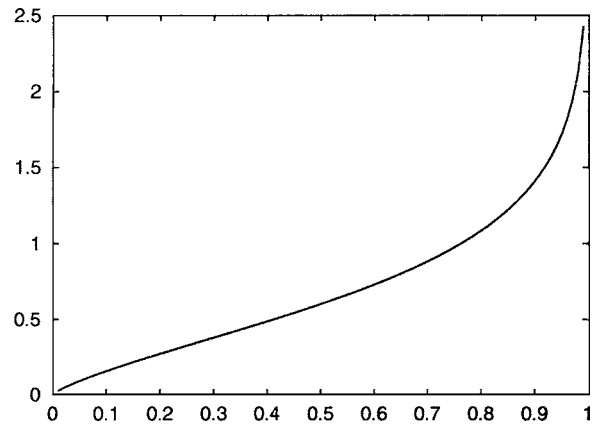

Figure 4. The required rigidity gain A_p in function of the proportion of outliers (p).

two subsets of *k* points have to be chosen, and that *k*! possible permutations remain possible to define point matches between these two subsets.

Proposition 3. *Let* $P = \{m_i\}_{i=1..n}$ *and* $P' = \{m'_i\}_{i=1..n}$ *the points detected on each image. A matching set S of size k formed by disjoint pairs taken from* $P \times P'$ *is ε-meaningful as soon as it is α-rigid with*

$$\varepsilon_3(\alpha, n, k) := 3(n-7) \cdot \binom{n}{k}^2 \cdot k! \cdot \binom{k}{7} \cdot \alpha^{k-7} \leq \varepsilon.$$

(8)

We have assumed for simplicity that $|P| = |P'|$, but in the general case where $|P| = n$ and $|P'| = n'$, we would obtain the condition

$$3 \min(n-7, n'-7) \cdot \binom{n}{k} \cdot \binom{n'}{k} \cdot k! \cdot \binom{k}{7} \cdot \alpha^{k-7} \leq \varepsilon.$$

The condition (8) is very different from the previous case (6), because the matching pairs now have to be found. In particular, even for $n = k$ (all points matched), the function $n \mapsto \varepsilon_3(\alpha, n, n)$ is increasing when *n* grows to infinity. This means that for a fixed rigidity α, no meaningful rigid set can be found beyond a certain value of *n*. The following proposition states this more precisely (see Appendix for the proof).

Proposition 4. *If $\varepsilon_3(\alpha, n, n) < 1$, then*

$$\alpha < \frac{1}{32} \quad and \quad n < \frac{e}{\alpha} + \underset{\alpha \to 0}{O}(\log \alpha),$$

and the minimum of $n \mapsto \varepsilon_3(\alpha, n, n)$ is obtained for a unique $n = \bar{n}$ such that

$$\bar{n} = \frac{1}{\alpha} - 8 + \underset{\alpha \to 0}{O}(\alpha).$$

Moreover, one has

$$\log \varepsilon_3(\alpha, \bar{n}, \bar{n}) = -\frac{\log e}{\alpha} - \frac{3}{2} \log \alpha - C' + \underset{\alpha \to 0}{O}(\alpha),$$

(9)

with $C' = \log \frac{7!}{3} - \frac{1}{2} \log(2\pi) = 2.82\ldots$ and $\log e = 0.434\ldots$

This proposition calls for several comments. First, it yields a universal threshold (1/32) on the rigidity coefficient α, above which a set cannot be meaningfully rigid. Second, for any expected rigidity α it suggests an optimal number $n = \bar{n}(\alpha)$ of characteristic points to be found on each image, and guarantees that rigidity detection is possible for that *n*, since the right term of (9) is negative for α small enough.

The situation is similar in the case of outliers. On Fig. 5 (right), we can see that given a precision α and a meaningfulness threshold (e.g. $\varepsilon_3 = 10^{-3}$), the maximum allowed amount *p* of outliers is not an increasing function of *n*: it increases up to an optimal value of *n* and then decreases for larger values of *n*. For *p* fixed, the required rigidity threshold α behaves in a similar way, and the less restrictive threshold is obtained for a particular value of *n* (Fig. 5, left).

For a practical[1] value of $\alpha = 10^{-2.5}$, we can see on Fig. 6 (left) or on Fig. 5 (right) that the maximum allowed amount of outliers is around $p = 0.25$, and
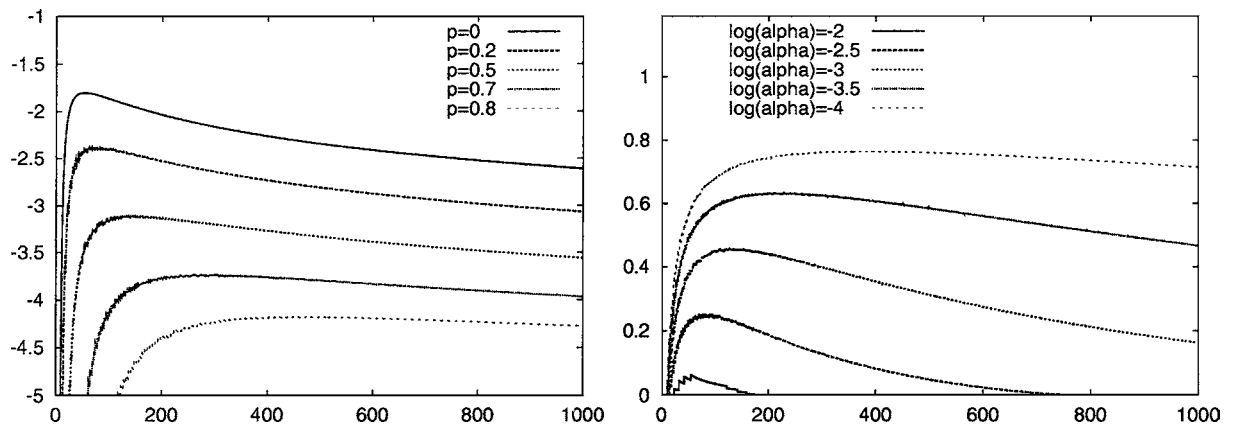

Figure 5.   Detectability curves ($\varepsilon_3 = 10^{-3}$) in the $(n, \log \alpha)$ plane for several values of *p* (left) and in the (n, p) plane for several values of α (right).
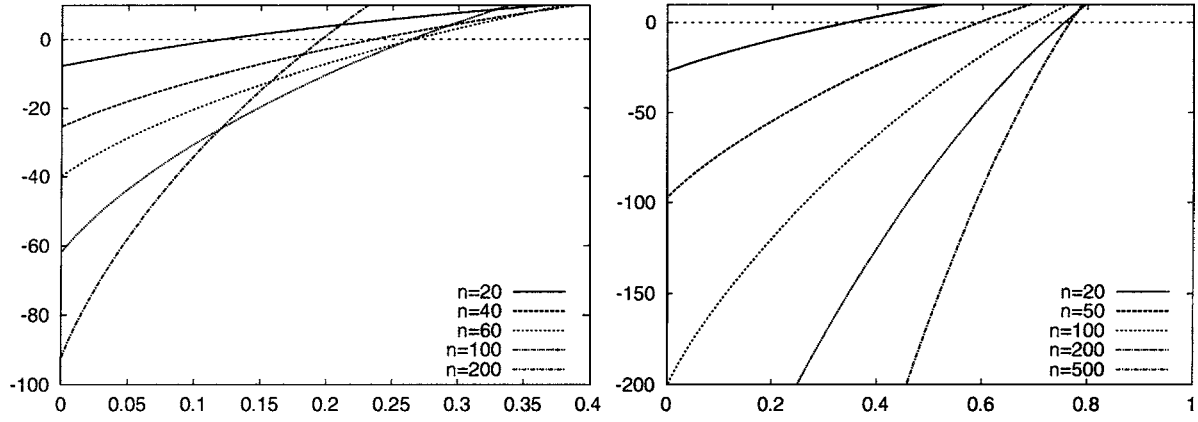
Figure 6. Log-meaningfulness $\log(\varepsilon_3)$ in function of the proportion p of outliers, for several values of n. Left: $\alpha = 10^{-2.5}$. Right: $\alpha = 10^{-4}$.

is attained around $n = 80$. Large values of p can be attained for $\alpha = 10^{-4}$ for example (see Fig. 6, right), but this would require an irrealistic precision (0.035 pixel in a 1000×1000 image) in the localization of the characteristic points. This outlines an absolute limit of rigidity detection in the case of non-matched points, and suggests that in most situations rigidity detection is not possible, since a large amount of outliers is to be expected if no matching criterion is involved in the detection of characteristic points.

However, small values of α (say $\alpha = 10^{-4}$ and even less) may be attained by considering what we could call *colored rigidity*. If we attach to each point some quantity c that is expected to be the same in each view (for example, a kind of dominant color around the point, or any other measure not a priori related to the location of the point), then the meaningfulness of a rigid set can be improved by constraining not only a good match between each pair of points (α small), but also between the values of c measured at these points. If c is a scalar quantity that is uniformly distributed in $\{1..M\}$, then constraining $|c(m_i) - c(m_i')| \leq \delta$ for each pair (m_i, m_i') of matching points allows to multiply the meaningfulness coefficient ε_3 by $(\frac{2\delta+1}{M})^k$, which practically amounts to multiply α by $\frac{2\delta+1}{M}$. In this paper, we shall not investigate further this possibility of colored rigidity, but we think this could yield an interesting way to compensate for the tight rigidity thresholds involved in the case of non matched points.

Another reason for considering colored rigidity is the huge computational cost required to search for sets of 7 point matches. In the case of non-matched points, there are $7!\binom{n}{7}^2$ possible sets of 7 point matches, which is to

be compared to the (only) $\binom{n}{7}$ possible sets in the case of already matched points. The large ratio $7!\binom{n}{7} \simeq n^7$ suggests that a systematic (or even stochastic) search is irrealistic, and this is why we shall not produce experiments for the ε_3 criterion in Section 5. However, we keep thinking that the ε_3 criterion may be useful, for example if additional information (colored rigidity mentionned above, or local image comparison near the points to match) allows to reduce the search for matches to a limited number of cases. This extension would probably require a study by itself.

3.4. A Few Remarks

Before we describe the algorithms and present experiments, let us discuss the validity of the 3 models we presented. The first model (ε_1) concerns only the case when all point matches are to be kept in the rigidity detection. Even if it is known that no outliers can be present, the use of this model is questionable (that is, the second model should be preferred) since the less precise point matches may reduce the meaningfulness of the whole set and compromise the precision of F. Hence, this first model has mainly a pedagogic value.

The second model (ε_2) is applicable when point matches are available between the two images, provided that no location criterion was involved in the matching process. If, for example, each point m_i has been matched to the nearest m_j' in the second image (for some distance that may not be the usual Euclidean distance), which supposes that some a priori assumption is made on the relative camera positions, then a meaningful rigid set may be found by accident with

the ε_2 criterion, since the uniform model against which rigid sets are detected is not valid any more. In this case, or when point matches are not known, the third model (ε_3) should be used.

The meaningfulness thresholds we derived should be distinguished from the algorithms used for rigidity detection. Even if the meaningfulness thresholds count *all* possible transforms F obtained from 7 point matches, which may represent a huge number (especially for ε_3), a selective exploration of these transforms can be used in practice to detect rigidity. If, for example, some restrictions are known about the epipolar geometry (this may be the case for a stereo pair, for example), then this knowledge might be very useful to speed up the rigidity detection. All kinds of optimized strategies for rigidity detection may be used, provided that all possible transforms are counted in the derivation of the meaningfulness thresholds, as we did for ε_3.

We also would like to point out that even if our model uses a uniform distribution of points to derive the meaningfulness coefficient, this is not an assumption to be realized by the observed data since we detect rigid sets *against* this uniform model. If, for example, all points of the second image are concentrated in a very small region, then this particular concentration (which is in some way meaningful) might raise the detection of a rigid set. If a symmetrical definition of rigidity is used, as suggested in Section 3.2, then this might only happen if this concentration also occurs in the first image. In this case, the existence of a rigid set still makes sense, though it would probably not be the *simplest explanation* of the observed distribution of points. This suggests that a similar approach should also be applied for the detection of simpler motions (given by homographies, affine or even Euclidean transforms), in order to allow the use of Occam's razor principle ("*the simplest explanation is the best*"), which could be realized for example with a Minimum Description Length (Rissanen, 1983, 1989) criterion. Another possibility would be to detect rigidity against a more general model of points distribution allowing a certain amout of clustering, but this is not clear that the new obtained thresholds would be significantly different.

4. Algorithms

4.1. Combinatorial Search

We implemented the 7 points algorithm in C using the Singular Value Decomposition algorithm from Numer-

ical Recipes (Press et al., 1988) and a classical explicit solver for the third degree equation raised by the rank 2 constraint. Our algorithm is able to test 10,000 7-points matches per second (that is, between 10,000 and 30,000 fundamental matrices) on a recent PC laptop (1 GHz).

When considering a fundamental matrix F given by a subset T of 7 point matches of S, we compute for each of the $n - 7$ remaining pairs the normalized distance α_i between m_i' and the epipolar line Fm_i. We then sort these distances in the increasing order and obtain for each value of k between 8 and n the subset $S_k(F)$ of size k such that $T \subset S_k \subset S$ and $\alpha_F(S_k)$ is minimal. Since $\varepsilon_2(\alpha, n, k)$ is increasing with respect to α (see Eq. (6)), $S_k(F)$ is also the most meaningful (with respect to F) rigid set of size k containing T. Then we find, among the sets $(S_k(F))_{k=8..n'}$ the most meaningful one (that is the one minimizing ε_2), written $\bar{S}(F)$.

The deterministic algorithm works as follows: for each subset T of 7 point matches of S and each of the (possibly 3) fundamental matrices F associated to T, we compute $\bar{S}(F)$. The result of the algorithm is simply the most meaningful set $\bar{S}$ encountered in this systematic search. We also compute, for each value of k, the most meaningful set *with size k*, that is the most meaningful set S_k among the $S_k(F)$. By construction, this function of k is minimal when $S_k = \bar{S}$.

4.2. Random Sampling Algorithm

The algorithm above must be abandoned when $n \geq 25$ in favor of a stochastic version, as shown by Table 1. Following the random sampling consensus, we run similar computations of rigidity and meaningfulness as above, but only for certain fundamental matrices F obtained from randomly sampled 7-points matches. The algorithm can be written as follows:

```
set ε̄ = +∞
repeat
    find a random set T of 7 point matches
    for each fundamental matrix F associated to T
        compute the most meaningful rigid set
            S̄ = S̄(F) associated to F
        if ε₂(S̄) < ε̄, set ε̄ = ε₂(U) and U = S̄
    end
until the number of trials T exceeds N
return U and ε̄
```

Notice that once U (the set of inliers) has been found, one may apply as a final step a classical linear optimization (e.g. Least Mean Squares) to obtain the optimal estimation of F. In practice, this final step may be rather useless since the estimation of F performed by our algorithm (using sets of 7 points) is generally excellent (this will be confirmed by Section 5.1 and Fig. 8).

If we assume that among n point matches, k are correct (that is, we have $n - k$ outliers), then the probability of selecting at least one set of 7 correct point matches in N trials is

$$q = 1 - \left(1 - \frac{\binom{k}{7}}{\binom{n}{7}}\right)^N = 1 - \left(1 - \prod_{i=0}^{6} \frac{k-i}{n-i}\right)^N.$$

For $q = 95\%$ (that is, $\log_e(1 - q) \simeq -3$), we can see that the number of trials required is approximately

$$N \simeq 3\left(\frac{n}{k}\right)^7, \tag{10}$$

which allows a proportion $p = 1 - k/n$ of outliers of about 70% to stay in reasonable computation time limits (10,000 trials). A proportion of 90% outliers would require $N = 30{,}000{,}000$ trials, which is unfeasible. However, we shall see that this estimate of N is questionable, and that we manage to find a rigid set among more than 90% outliers in much less than 30,000,000 trials.

How do we compare to classical random sampling algorithms? In the RANSAC algorithm (see Torr and Murray, 1997 for example), a threshold has to be preset arbitrary to decide whether a set of point matches is "compatible" with a fundamental matrix F. Then, using random sampling, the largest compatible set is selected. Our algorithm works in a similar way, except that no arbitrary threshold has to be selected: we use the ε_2 criterion to rate the meaningfulness of a set of point matches in function of its size (k) and its rigidity (α). This way, we allow large sets of fair point matches as well as small sets of accurate ones, which probably explains why we obtain better results (see Section 5). Even better, we know in the end if the most meaningful set found by the algorithm can be considered as meaningfully rigid (that is $\varepsilon_2 \ll 1$) or if it could have been that rigid by accident. Last, the selection of the most meaningful rigid set allowed by the ε_2 criterion permits to define a very efficient optimization step, as shown below.

4.3. Optimized Random Sampling Algorithm (ORSA)

There is a simple way to improve the stochastic algorithm. It relies on the idea that, on average, the proportion of outliers should be smaller than p among the most meaningful rigid sets. This suggests that the final optimization step below should be added at the end of the previous Random Sampling Algorithm (just before the "return U and $\bar{\varepsilon}$" line).
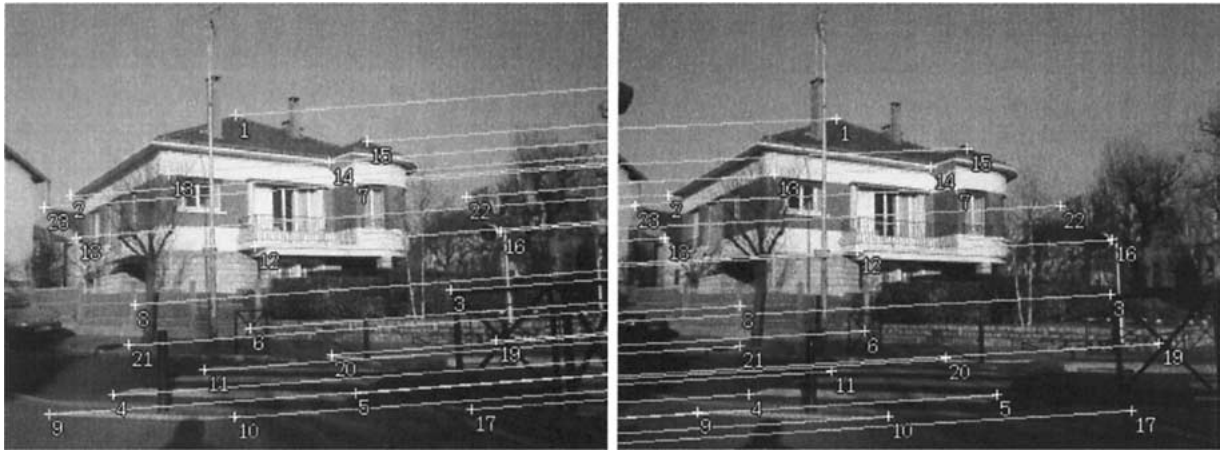

Figure 7. The first stereo pair with 23 point matches, and the best epipolar geometry recovered by the systematic exploration of all 7-points transforms.
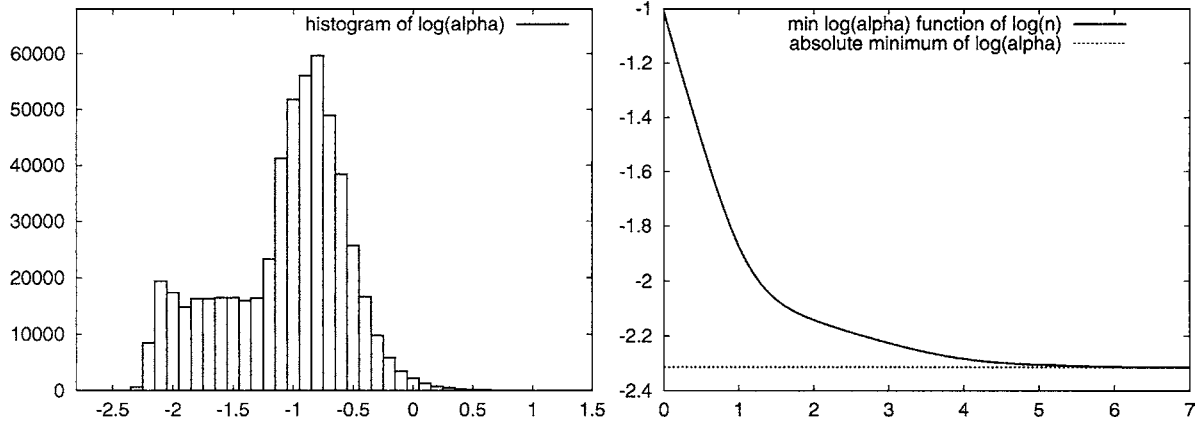
Figure 8. Left: the histogram of $\log \alpha_F(S)$ for all possible F obtained from sets of 7 point matches taken out from S. From this histogram (measured at a much finer scale), we computed the expectation of $\min_{i=1..n} \log \alpha_{Fi}(S)$ when F_i are obtained from independent random sets of 7 of point matches. This expectation is represented as a function of $\log n$ on the right. We can see that the systematic exploration of all sets of 7 point matches is not really required: an excellent rigidity coefficient ($\log \alpha = -2.23$) would be obtained in about 1000 random trials, which is to be compared to the 550,000 trials needed to obtain the best possible rigidity coefficient $\log \alpha = -2.31$.

set $\bar{\varepsilon} = \varepsilon_2(U)$
repeat
    find a random set T of 7-point matches among U
    for each fundamental matrix F associated to T
        compute the most meaningful rigid set
          $\bar{S} = \bar{S}(F)$ associated to F
        if $\varepsilon_2(\bar{S}) < \bar{\varepsilon}$, set $\bar{\varepsilon} = \varepsilon_2(\bar{S})$ and $U = \bar{S}$
    end
until the number of trials T exceeds N_{opt}

If U is absolutely meaningful (that is, $\varepsilon_2(U) < 1$), this optimization step converges very quickly and N_{opt} does not require to be large. In any case, it cannot be harmful since $\bar{\varepsilon}$ is constrained to decrease. In practice, we set $N_{opt} = \frac{N}{10}$ and apply this optimization step to the first absolute meaningful set found by random sampling, or, if none has been found after $N - N_{opt}$ trials, to the most meaningful set found so far. In this way, we ensure that the total number of trials cannot exceed N (but can be as small as $N_{opt} + 1$), while improving dramatically the detection of outliers, as shown in Section 5. We shall refer to this algorithm with the ORSA acronym.

5. Experiments

5.1. Checking All Matchings

We took two images (resolution 800×600) of the same scene with a digital camera and we recorded by hand a set S of 23 point matches on them, with about one pixel accuracy (see Fig. 7). However, by taking a snapshot

of a regular grid we noticed that geometric distortions were between 0 and 5 pixels, so we could expect a rigidity coefficient between 0 and

$$\alpha = \frac{2 \cdot 5 \cdot \sqrt{800^2 + 600^2}}{800 \cdot 600} \simeq 10^{-1.68}.$$

Using the systematic combinatorial search, we obtained $\log(\alpha) = -2.31$ with the best 7-points transform, while the Least Mean Squares optimization applied to the 23×9 linear system yielded a transform F such that $\log \alpha_F(S) = -2.32$ (note that the Least
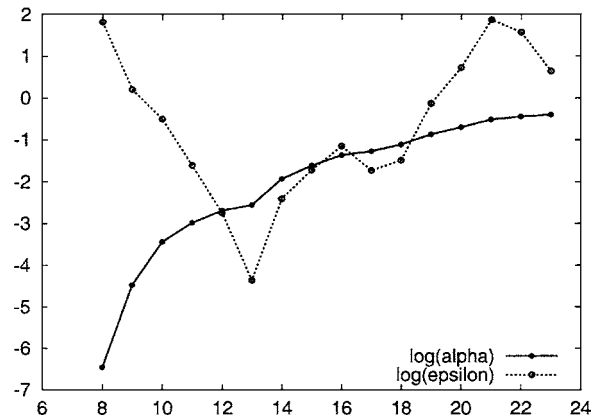

Figure 9. Rigidity error $\bar{\alpha}(k)$ and meaningfulness $\bar{\varepsilon}(k)$ of the most rigid subset of size k (log scale). Whereas $\bar{\alpha}(k)$ can hardly be used alone to detect outliers, $\bar{\varepsilon}(k)$ presents a strong minimum which permits to detect the 13 inliers easily.
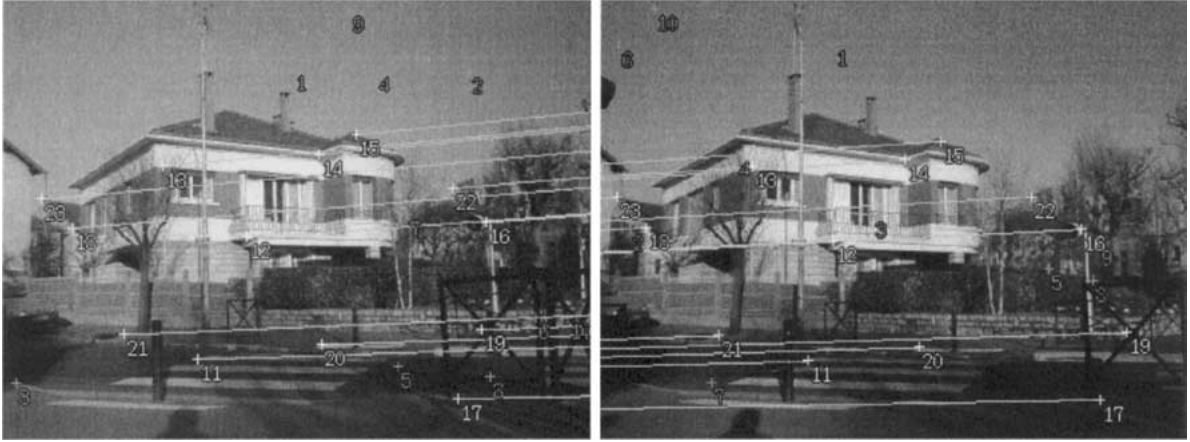
Figure 10. When we replace the first 10 point matches (among 23) by random points, the use of the ε_2 criterion allows to identify the meaningfully rigid set of 13 inliers and to recover the right epipolar geometry, as shown above.
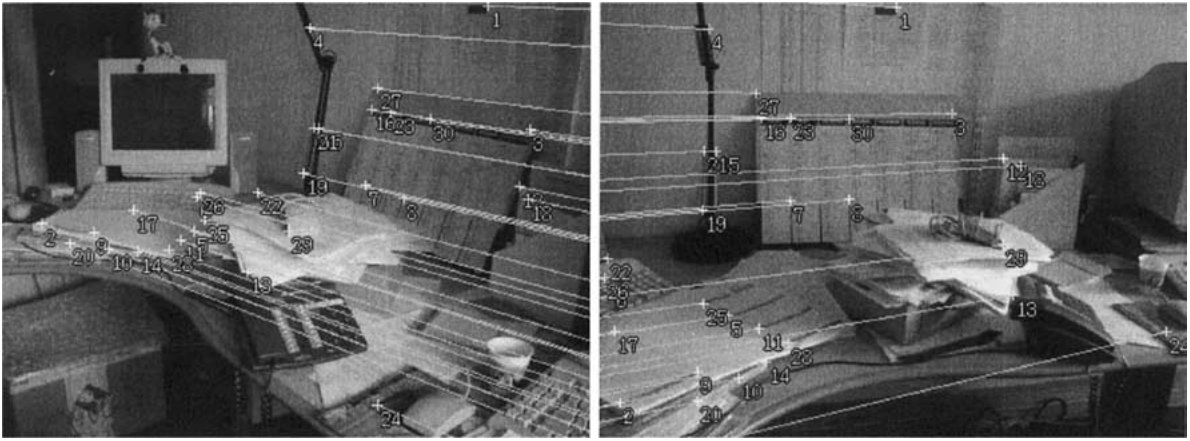

Figure 11. The second stereo pair with 70 point matches (only 30 of them are shown here for readability). The epipolar geometry is recovered in no time an yields $\log \alpha = -1.92$ and $\log \varepsilon = -110$.

Means Squares optimization does not significantly improve the estimation of F, compared to the systematic combinatorial search). As expected, the meaningfulness of the rigidity test was very good:

$$\varepsilon_1 = 3 \cdot \binom{23}{7} 10^{-2.31(23-7)} \simeq 10^{-31}.$$

One could argue that testing $\binom{23}{7} \simeq 250,000$ configurations takes time (25 seconds), but in fact a nearly optimal accuracy is achieved by testing only a small number of random 7-points transforms, as shown on Fig. 8. This justifies the systematic use of the random sampling algorithm we presented in Section 4.

5.2. Detecting Outliers

We simulated outliers by replacing the first 10 point matches (among 23) by random points. Then we computed, for each fundamental matrix F yielded by a 7-point transform, the most rigid set size k, $S_k(F)$ (see Section 4.1), as well as its rigidy $\alpha^k(F) = \alpha_F(S_k(F))$ and its meaningfulness $\varepsilon^k(F) = \varepsilon_2(\alpha_k(F), 23, k)$. On Fig. 9 are represented the two functions $k \mapsto \bar{\alpha}(k) = \min_F \alpha^k(F)$ and $k \mapsto \bar{\varepsilon}(k) = \varepsilon^k(F)$. The function $\bar{\alpha}(k)$ is increasing with respect to k (by construction), and presents only a small gap between $k = 13$ and $k = 14$, from which it would be difficult (and uncertain) to detect directly the outliers. On
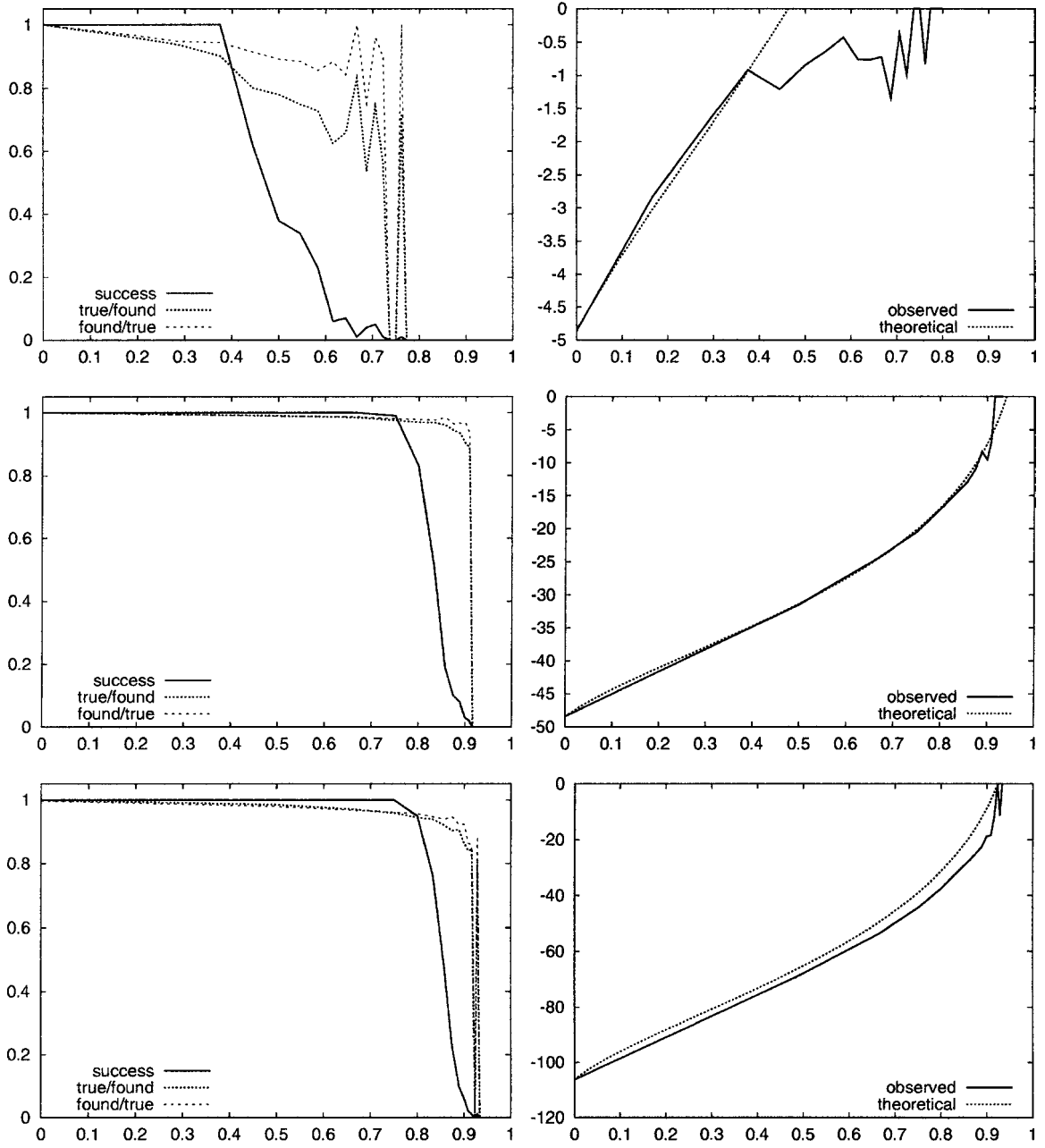
Figure 12. Performance of the Optimized Random Sampling Algorithm (10,000 trials) in function of the proportion p of outliers (horizontal axis). Top row: 10 inliers. Middle row: 30 inliers. Bottom row: 70 inliers. Left: empirical probability of success (detection of an absolutely meaningful rigid set), average proportion of true inliers among the most rigid set found (in case of success), and average proportion of points found among true inliers (still in case of success). Right: observed and theoretical average meaningfulness $\langle \varepsilon_2 \rangle$ in case of success.

the contrary, the function $\bar{\varepsilon}(k)$ presents a strong minimum for $k = 13$, which identifies without doubt the true set of 13 point matches and the 10 outliers. The epipolar geometry is well recovered too, as shown on Fig. 10.

5.3. Evaluation of the Optimized Random Sampling Algorithm

In some situations, the number of point matches n may be large (say $n \geq 25$) and a systematic exploration of

all sets of 7 point matches is not possible (see Table 1). In this case, we use the Optimized Random Sampling Algorithm (ORSA) presented in Section 4.3. In fact, it is interesting to use this algorithm even for small values of n, as shown on Fig. 8. In this section, we propose a systematic evaluation of ORSA and we prove in particular the usefulness of the final optimization step we introduced in Section 4.3.

We took a second couple of 800×600 images, and recorded by hand 70 point matches on them (see Fig. 11). Then we selected a set T_k of k point matches (among these 70) for three values of k ($k = 10$, $k = 30$ and $k = 70$), and added $kp/(1 - p)$ random outliers for several values of p before we ran ORSA on these data with $N = 10,000$. We repeated this 100 times for each pair (k, p), so that we could compute:

- the empirical probability of success (by success, we mean that an absolute meaningful set has been found, that is $\bar{\varepsilon} < 1$);
- the average proportion of true points (that is, belonging to T_k) among all points found in case of success;
- the average proportion of points found among true points in case of success;
- the average meaningfulness $\langle \varepsilon_2 \rangle$ in case of success.

The results are shown on Fig. 12. We can notice on these experiments that for $N = 10,000$ (which is not very large), the detection thresholds are good, and in case of success the recovery of the true set of inliers is almost perfect (notice that by adding random points, we may also add new pairs that match well the true

epipolar transform). With a probability of success of 50% (which would become $1 - 0.5^5 = 97\%$ for $N = 50,000$), we can allow a proportion of outliers equal to $p = 0.5$ for $n = 10$, to $p = 0.83$ for $n = 30$ and to $p = 0.86$ for $n = 70$. This last case means that we are able to detect 70 inliers among 500 points, and proves that (10) underestimates by far the performance of ORSA.

Another interesting point is that the ε_2 criterion allows us to predict the performance of the algorithm. From the rigidity coefficient α obtained with $p = 0$ (no outliers), we could compute a theoretical estimate of $\langle \varepsilon_2 \rangle$ with

$$\varepsilon_2 \left(\frac{k}{1 - p}, k, \alpha \right).$$

This number measures the meaningfulness of U in function of the proportion p of outliers, U being the most meaningful rigid set found for $p = 0$. On Fig. 12 (right), we can see that this estimate is generally very accurate, which shows that our rigidity detection algorithm is efficient (the rigidity α of the most meaningful set it finds is quite insensitive to the proportion of outliers) and that we are able to predict the value of p for which it breaks down, as the number p_0 for which $\varepsilon_2(\frac{k}{1-p_0}, k, \alpha) = 1$.

The optimization step described in Section 4.3 appeared to be very efficient. By using pure (non-optimized) random sampling instead of ORSA, we obtained much worse results both for the probability of success and for the precision of the estimate of F, as
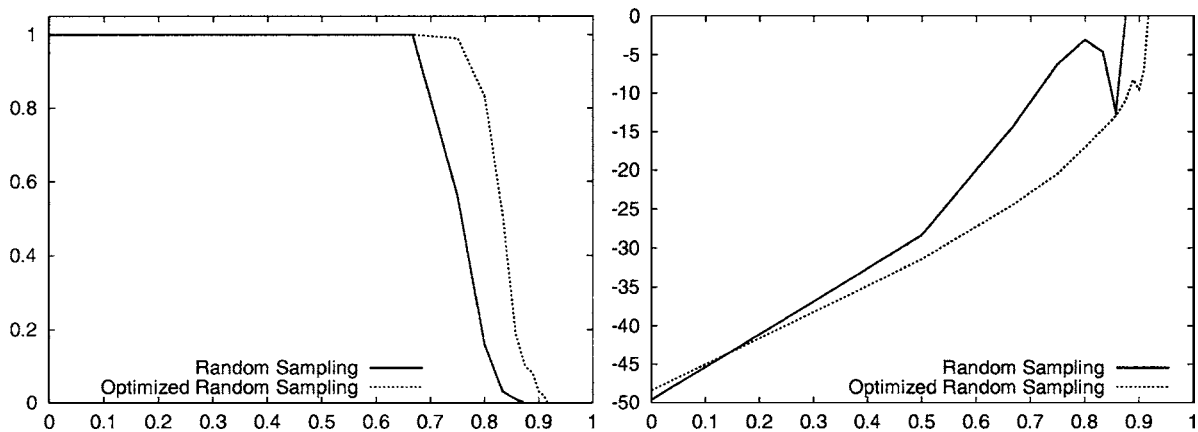

Figure 13. Performance of non-optimized random sampling, compared to ORSA (30 inliers, 10,000 trials). Left: empirical probability of success. Right: average meaningfulness $\langle \varepsilon_2 \rangle$. This shows the efficiency of the optimization step described in Section 4.3: instead of being able to detect (with probability $1/2$ for 10,000 trials) the 30 inliers among 175 point matches ($p = 0.83$), the non-optimized random sampling algorithm reaches its limit for 115 point matches ($p = 0.74$), and the accuracy of the fundamental matrix it computes is far from being optimal (right).

shown by the loss of meaningfulness (see Fig. 13). Hence, our optimization process not only speeds up the algorithm, but often permits to find a meaningful rigid set when none has been found in $N - N_{\text{opt}}$ trials, by choosing sets of 7 point matches only from the most meaningful rigid set found so far.

6. Conclusion

The a contrario model we presented in this paper yields absolute thresholds for rigidity detection and allows the design of an optimized random sampling algorithm that is able to estimate the epipolar geometry between two images, even for a very large proportion of outliers. It can be applied to stereo pairs, or in the context of Structure From Motion, since no assumption is required on the magnitude of the camera motion. It could also be used to detect several rigid motions in an image sequence (see Torr et al., 1995 for an example), though we did not make any experiment for that application. The case of non-matched points revealed the existence of an absolute accuracy threshold (1/32) in the matching process and of an optimal value for the number of points to be matched, depending on the expected accuracy and on the expected proportion of outliers. It suggests that additional measures should be used for point correspondence, like color or other perspective-invariant qualities. Our criterion does not measure the uncertainty of the fundamental matrix, and it might happen that a very meaningful rigid set be found with a very uncertain associated fundamental matrix. Hence, computing this uncertainty (see Zhang, 1998 for a review) may still be useful for practical applications. The approach we presented could also be applied to simpler motions, as mentioned in Section 3.4. The case of lines matches (Faugeras et al., 1987) or of simultaneous points and lines matches could probably be treated as well. The extension to more than two views and to the detection of several simultaneous rigid motions could also have interesting applications.

Appendix

Proof of (5):   We have, when $n \to +\infty$,

$$\log \binom{n}{7} = 7 \log n - \log(7!) + O\left(\frac{1}{n}\right),$$

so that if $3\binom{n}{7}\alpha^{n-7} = \varepsilon$, then

$$
\begin{aligned}
\log \alpha &= \frac{1}{n-7}\left(\log \varepsilon + \log \frac{7!}{3} - 7 \log n + O\left(\frac{1}{n}\right)\right) \\
&= -7\frac{\log n}{n} + \frac{C + \log \varepsilon}{n} + O\left(\frac{\log n}{n^2}\right) \quad (11)
\end{aligned}
$$

with $C = \log(7!/3)$.   □

Proof of (7):   If $k = (1 - p)n$ (p fixed, $0 < p < 1$) and $n \to +\infty$, Stirling's formula yields

$$
\begin{aligned}
\log \binom{n}{k} &= -n(p \log p + (1-p)\log(1-p)) - \frac{1}{2}\log n \\
&\quad - \frac{1}{2}\log(2\pi p(1-p)) + O\left(\frac{1}{n}\right), \quad (12)
\end{aligned}
$$

so that if $(n-7)\binom{n}{k} \cdot 3\binom{k}{7}\alpha^{k-7} = \varepsilon$, then (12) and (11) can be combined into

$$
\begin{aligned}
\log \alpha &= -\frac{1}{k-7}\left(\log(n-7) + k A_p - \frac{1}{2}\log n \right. \\
&\quad \left. - \frac{1}{2}\log(2\pi p(1-p)) + O\left(\frac{1}{n}\right)\right) \\
&\quad - 7\frac{\log k}{k} + \frac{C + \log \varepsilon}{k} + O\left(\frac{\log k}{k^2}\right),
\end{aligned}
$$

with $A_p = -\log(1-p) - \frac{p}{1-p}\log p$. Hence, writing $B_p = -\frac{1}{2}\log(2\pi p(1-p)^2)$, we get

$$
\begin{aligned}
\log \alpha &= -A_p - \frac{15 \log k}{2k} + \frac{C + \log \varepsilon + B_p - 7A_p}{k} \\
&\quad + O\left(\frac{\log n}{n^2}\right).
\end{aligned}
$$

□

Proof of Proposition 4:   Let us write $u_n = \varepsilon_3(\alpha, n, n) = 3(n - 7)n!\binom{n}{7}\alpha^{n-7}$. We have

$$\frac{u_n}{u_{n-1}} = \frac{n-7}{n-8} \cdot n \cdot \frac{n}{n-7} \cdot \alpha = \frac{n^2 \alpha}{n-8},$$

so that

$$u_n \le u_{n-1} \Leftrightarrow \alpha n^2 - n + 8 \le 0.$$

If $\alpha \ge 1/32$, this last inequality can never be realized so that u_n is increasing and

$$u_n \ge u_8 = 3 \cdot 8 \cdot 8! \cdot \alpha \ge 6 \cdot 7! \gg 1,$$

which is impossible since we assumed $\varepsilon_3(\alpha, n, n) < 1$. Thus, we necessarily have $\alpha < 1/32$, and since one root of $n^2 - n + 8$ is in $]0, 8[$ and the other $(\bar{n})$ larger than 8, u_n is decreasing for $n \leq \bar{n}$ and increasing for $n \geq \bar{n}$, with

$$\bar{n} = \frac{1 + \sqrt{1 - 32\alpha}}{2\alpha} = \frac{1}{\alpha} - 8 + \underset{\alpha \to 0}{O}(\alpha).$$

Moreover, when $n \to \infty$ one has

$$\log u_n = n \log \frac{\alpha n}{e} + O(\log n),$$

so that the condition $u_n \to 1$ imposes

$$\frac{\alpha n}{e} = 1 + O\left(\frac{\log n}{n}\right),$$

which means that

$$n = \frac{e}{\alpha} + \underset{\alpha \to 0}{O}(\log \alpha).$$

Last, we have

$$\log u_n = -\log \frac{7!}{3} + \log n + n \log \frac{n}{e} + \frac{1}{2}\log(2\pi n)$$
$$+ 7\log n + (n - 7)\log \alpha + \underset{n \to +\infty}{O}\left(\frac{1}{n}\right),$$

and since $\alpha = \frac{1}{\bar{n}}(1 - \frac{8}{\bar{n}})$,

$$\log u_{\bar{n}} = \frac{3}{2}\log \bar{n} - (\bar{n} + 8)\log e - \log \frac{7!}{3} + \frac{1}{2}\log 2\pi$$
$$+ O\left(\frac{1}{\bar{n}}\right)$$
$$= -\frac{\log e}{\alpha} - \frac{3}{2}\log \alpha - C' + \underset{\alpha \to 0}{O}(\alpha)$$

with $C' = \log \frac{7!}{3} - \frac{1}{2}\log 2\pi$.   □

Acknowledgments

Note

1. $\alpha = 10^{-2.5}$ means a precision of about 1 pixel on a 1000×1000 image (see Eq. (2)). Achieving smaller values is not impossible, but requires an excellent compensation of the camera distortions, and an accurate sub-pixel registration algorithm.

References

Boufama, B. and Mohr, R. 1995. Epipole and fundamental matrix estimation using virtual parallax. In *Proc. Int. Conf. Vision*, Boston, pp. 1030–1036.

Desolneux, A. Moisan, L., and Morel, J.-M. 2000. Meaningful alignments, *Int. J. of Computer Vision*, 40(1):7–23.

Faugeras, O., Lustman, F., and Toscani, G. 1987. Motion and structure from motion from point and line matches. In *Proc. 1st Int. Conf. on Computer Vision*, pp. 25–34.

Faugeras, O. and Luong, Q.-T., 2001. *The Geometry of Multiple Images*. MIT Press.

Fischler, M. and Bolles, R. 1981. Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography. *Communications of the ACM*, 24:381–385.

Hartley, R.I. 1997. Kruppa's equations derived from the fundamental matrix. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 19(2):133–135.

Hartley, R.I. 1997. Self-calibration of stationary cameras. *Int. J. of Computer Vision*, 22(1):5–24.

Longuet-Higgins, H. 1981. A computer algorithm for reconstructing a scene from two projections. *Nature* 293:133–135.

Lowe, D. 1985. *Perceptual Organization and Visual Recognition*. Kluwer Academic Publishers.

Luong, Q.-T. and Faugeras, O. 1996. The fundamental matrix: Theory, algorithms and stability analysis. *Int. J. of Computer Vision*, 17(1):43–76.

McReynolds, D.P. and Lowe, D.G. 1996. Rigidity checking of 3D point correspondences under perspective projection. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 18(12):1174–1185.

Oliensis, J. and Genc, Y. 2001. Fast and accurate algorithms for projective multi-image structure from motion. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 23(6):546–559.

Press, W.H., Teukolsky, S.A., Vetterling, W.T., and Flannery, B.P. 1988. *Numerical Recipes in C*. Cambridge University Press.

Rissanen, J. 1983. A universal prior for integers and estimation by Minimum Description Length. *Annals of Statistics*, 11(2):416–431.

Rissanen, J. 1989. *Stochastic Complexity in Statistical Inquiry*. World Scientific Press.

Salvi, J., Armangué, X., and Pages, J. 2001. A survey addressing the fundamental matrix estimation problem. *IEEE Int. Conf. on Image Processing*, Thessaloniki (Greece).

Stewart, C.V. 1995. MINPRAN: A new robust estimator for computer vision. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 17:925–938.

Tang, C.K., Medioni, G., and Lee, M.S. 2001. N-dimensional tensor voting, and application to epipolar geometry estimation. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 23(8):829–844.

Torr, P.H.S. and Murray, D.W. 1997. The development and comparison of robust methods for estimating the fundamental matrix. *Int. J. of Computer Vision*, 24(3):271–300.

Torr, P.H.S., Zisserman, A., and Murray, D.W. 1995. Motion clustering using the trilinear constraint over three views. In *Workshop on Geometrical Modeling and Invariants for Computer Vision*, Xidian University Press.

Zhang, Z., Deriche, R., Faugeras, O., and Luong, Q.T. 2001. Estimating the fundamental matrix by transforming image points

in projective space. *Computer Vision and Image Understanding* 82:174–180.

Zhang, Z. 1998. Determining the epipolar geometry and its uncertainty: A review. *Int. J. of Computer Vision*, 27(2):161–195.

Zhang, Z., Deriche, R., Faugeras, O., and Luong, Q.-T. 1994. A robust technique for matching two uncalibrated images through the recovery of the unknown epipolar geometry. *AI Journal* 78:87–119.