

Computing the Uncertainty of the 8 Point Algorithm for Fundamental Matrix Estimation

Frédéric SUR, Nicolas NOURY, Marie-Odile BERGER

INRIA Nancy Grand-Est / INPL

MAGRIT project-team, LORIA

Campus Scientifique BP 239, 54506 Vandoeuvre-lès-Nancy cedex, France

Abstract

Fundamental matrix estimation is difficult since it is often based on correspondences that are spoiled by noise and outliers. Outliers must be thrown out via robust statistics, and noise gives uncertainty. In this article we provide a closed-form formula for the uncertainty of the so-called 8 point algorithm, which is a basic tool for fundamental matrix estimation via robust methods. As an application, we modify a well established robust algorithm accordingly, leading to a new criterion to recover point correspondences under epipolar constraint, balanced by the uncertainty of the estimation.

1 Introduction and motivation

The fundamental matrix is the basic object that represents the geometric information between two views in the pinhole camera model [6, 8]. Considering two image points (one in each view) that correspond to the same 3D point, the fundamental matrix is the 3×3 matrix F that makes the homogeneous coordinates \mathbf{x} and \mathbf{x}' of these points satisfy the relation $\mathbf{x}'F\mathbf{x} = 0$. Estimating F is crucial for structure and motion problems where information (location of the camera and 3D scene) has to be retrieved from several images.

The fundamental matrix is most of the time evaluated from point correspondences. Since F is defined up to a scale factor, it can be retrieved from eight correspondences. In fact, because F has rank 2, seven points are sufficient. Estimating F is a difficult task since point location is noisy and the correspondences are spoiled by outliers. Noise calls for estimating F over the largest possible set of correspondences but outliers make it awkward. A popular solution is to use tools from robust statistics such as RANSAC [7], LMedS or M-estimators [18]. RANSAC, and methods inspired from it, consists in iterating two steps: 1) estimating a model (here F) from the smallest possible set of data (here 7 or 8 randomly chosen correspondences), then 2) build a so-called consensus set among the data that fit the model. The shortcoming of RANSAC is the specification of the consensus set, and a large literature has been dedicated to this problem, see for example [3, 11, 17]. Once a consensus set has been built, it is possible to estimate the fundamental matrix from it. See [18] for a large survey on fundamental matrix estimation.

RANSAC-like methods make it possible to detect outliers, but the inaccuracy in the image point location is rarely modeled. Seminal contributions w.r.t. this are MLESAC [17] and MAPSAC [16] where point location is endowed with a Gaussian model. Such methods require to first estimate not only the variance of the point image location, but also outlier distribution and prior probabilities. Besides, they focus on model fitting and do not take

into account the uncertainty of the fundamental matrix estimation from the minimal set although it is also spoilt by inaccuracy.

The uncertainty of the fundamental matrix F is characterized by its two first moments, that is, its mean and its covariance which measures the spreading of F . It has already been computed for algorithms based on a (large) sample, free from outliers (e.g. on a consensus set or a set of manually selected points). Csurka et al. [4] as well as Hartley and Zisserman [8] explain how to derive it when F is estimated as the minimizer of the non-linear criterion:

$$C(F) = \sum_{(\mathbf{x}, \mathbf{x}')} d(F\mathbf{x}, \mathbf{x}') + d(x, F^T \mathbf{x}') \quad (1)$$

where n correspondences $(\mathbf{x}, \mathbf{x}')$ are given¹. The geometrical meaning of the quantity $d(F\mathbf{x}, \mathbf{x}')$ is the distance of \mathbf{x} to the epipolar line $F\mathbf{x}'$, which is 0 if $(\mathbf{x}, \mathbf{x}')$ is an actual correspondence. Let us also name [15], where the authors compute F as a maximum likelihood estimation and give a bound on its covariance, and [19] with a dedicated parameterization. However, in these papers the estimation is also based on an outlier-free sample. The present article is also related to [13] where covariance estimation is used in the guided matching context.

There are usually two ways of estimating F covariance: Monte-Carlo simulations, and derivation of a closed-form formula. Monte-Carlo suffers from its low convergence rate (which is all the lower since one estimates a vector of \mathbb{R}^9 , due to the so-called *curse of dimensionality*) while closed-form formula are only first order approximations. Let us also cite covariance estimation via the unscented transformation [9].

The main contribution of this article is to give the explicit computation of the uncertainty of the 8 point algorithm (Sec. 2). This is important since the 8 point algorithm is the basis of many RANSAC-like methods. We then derive the uncertainty of the epipolar lines from the uncertainty on F . This is a classical derivation, however we plug the result in a state-of-the-art robust algorithm and build a consistent probabilistic model (Sec. 3). A proof-of-concept experiment is in Sec. 4.

2 Uncertainty of the 8 point algorithm

The 8 point algorithm gives the fundamental matrix F which is consistent with 8 correspondences between two views. We propose here to characterize the uncertainty of the fundamental matrix estimation from the uncertainty of the image points. Let us notice that the uncertainty of F is inherent in the estimation algorithm, and that another algorithm would lead to another uncertainty computation.

2.1 Reminder: 8 point algorithm

For this section we refer to [18] and references therein. We have here two images representing the same scene from two different viewpoints, and n point correspondences between these two views. Let us give some notations:

- $F = (F_{ij})_{1 \leq i, j \leq 3}$ is the fundamental matrix between two views.
- f is the column $(F_{11}, F_{12}, F_{13}, F_{21}, F_{22}, F_{23}, F_{31}, F_{32}, F_{33})^T$.

¹by abuse of notation we identify the image points and their homogeneous coordinates.

- $\mathbf{x} = (x, y, 1)$ denotes an image point as well as its homogeneous coordinates (here (x, y) is the pixel coordinate).
- $(\mathbf{x}_j, \mathbf{x}'_j)$ is the j -th correspondence between the two views ($1 \leq j \leq n$),
- M is the $n \times 9$ matrix:

$$M = \begin{pmatrix} x'_1 x_1 & x'_1 y_1 & x'_1 & y'_1 x_1 & y'_1 y_1 & y'_1 & x_1 & y_1 & 1 \\ x'_2 x_2 & x'_2 y_2 & x'_2 & y'_2 x_2 & y'_2 y_2 & y'_2 & x_2 & y_2 & 1 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ x'_n x_n & x'_n y_n & x'_n & y'_n x_n & y'_n y_n & y'_n & x_n & y_n & 1 \end{pmatrix}$$

The epipolar geometry constraint (that is, $\forall j \in [1, n], \mathbf{x}'_j F \mathbf{x}_j = 0$) simply leads to the matrix equation $Mf = 0$. The case $n = 8$ corresponds to the 8-points algorithm. The trivial solution $f = 0$ is of course not acceptable, f is thus a singular vector associated to the singular value 0. Moreover, F has rank 2 (because the epipoles lie in its left and right kernels), which has to be imposed afterward because of numerical stability of f computation. A common way of handling the whole problem is to

1. Compute a singular value decomposition (SVD) of M and consider a singular vector associated to the smallest singular value (ideally 0); this is equivalent to searching an eigenvector associated to the smallest eigenvalue of $M^T M$ and also equivalent to the least square formulation, that is $\min_F \sum_{j=1}^n (x'^T_j F x_j)^2$ (subject to $\|F\|$ constant).
2. Then put the smallest singular value of F to zero, which is equivalent to substituting F with the closest rank 2 matrix w.r.t. the Frobenius norm.

Step 1 makes it necessary to choose F in a (generally) 1-dimensional singular or eigen space. Popular choices are either imposing $\|F\| = 1$ or setting one of the coefficient of F to 1. This latter choice is theoretically quite touchy since it endangers the numerical stability of F estimation if this coefficient should vanish. Although it is still a questionable point, we choose to set $F_{33} = 1$ for the sake of computation simplicity. Matrix F is thus given as a solution of the linear system $\tilde{M}\tilde{f} = c$, where:

1. $\tilde{f} = (F_{11}, F_{12}, F_{13}, F_{21}, F_{22}, F_{23}, F_{31}, F_{32})^T$,
- 2.

$$\tilde{M} = \begin{pmatrix} x'_1 x_1 & x'_1 y_1 & x'_1 & y'_1 x_1 & y'_1 y_1 & y'_1 & x_1 & y_1 \\ x'_2 x_2 & x'_2 y_2 & x'_2 & y'_2 x_2 & y'_2 y_2 & y'_2 & x_2 & y_2 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ x'_n x_n & x'_n y_n & x'_n & y'_n x_n & y'_n y_n & y'_n & x_n & y_n \end{pmatrix}$$

3. and $c = -(11111111)^T$.

In the case of interest, $n = 8$ and \tilde{M} is generically non-singular. We get F by solving

$$\tilde{f} = \tilde{M}^{-1} c, \quad (2)$$

and then imposing the rank 2 constraint.

2.2 Fundamental matrix uncertainty

The question of interest is, supposing the pixel coordinates of the image points are spoiled by Gaussian noise, how to estimate the uncertainty of the fundamental matrix F ?

In the sequel, we make use of the following classical theorem.

Theorem 1 (propagation property) Let v be a random vector in \mathbb{R}^m with mean \bar{v} and covariance matrix Σ , and $g : \mathbb{R}^m \rightarrow \mathbb{R}^n$ be an affine map such as $g(v) = g(\bar{v}) + A(v - \bar{v})$. Then $g(v)$ is a random vector in \mathbb{R}^n with mean $g(\bar{v})$ and covariance matrix $A\Sigma A^T$.

When g is not linear, a first order approximate gives an approximation of the covariance matrix by replacing A by the Jacobian matrix $J(\bar{v})$.

This property leads us to estimate the Jacobian of the whole transformation from $X = (x_1, y_1, x'_1, y'_1, x_2, y_2, x'_2, y'_2, \dots, x_8, y_8, x'_8, y'_8)$ to F . We first study the Jacobian of $X \mapsto \tilde{M}^{-1}c$, then we will take care of the rank 2 constraint.

Jacobian of \tilde{f} computation. Let $\Theta : A \mapsto A.c$, $\Psi : A \mapsto A^{-1}$, and $\Phi : X \mapsto \tilde{M}$. One has $\tilde{f} = \Theta \circ \Psi \circ \Phi(X)$, and the Jacobian matrix of the transform is, according to the chain rule:

$$J_{\Theta \circ \Psi \circ \Phi}(X) = J_{\Theta}(\Psi \circ \Phi(X)) \cdot J_{\Psi}(\Phi(X)) \cdot J_{\Phi}(X). \quad (3)$$

We now derive the Jacobian matrices.

- $J_{\Phi}(X)$ is a 64x32 matrix: first row consists in the 32 derivatives of the (1,1) coefficient of M (that is $x'_1 x_1$) w.r.t. the 32² components of X , second row consists in the 32 derivatives of the (2,1) coefficient of M (that is $x'_2 x_2$) and so on.
- $J_{\Psi}(A)$ is a 64x64 matrix. Since $d(A^{-1}) = -A^{-1}dA A^{-1}$ (differentiate $AA^{-1} = I_8$), one has $\frac{\partial(A^{-1})}{\partial E_{i,j}} = -A^{-1}E_{i,j}A^{-1}$ where $E_{i,j}$ is the canonical base matrix with entries equal to 0 except in position (i, j) which is 1. Now, $\frac{\partial(A^{-1})}{\partial E_{i,j}}$ gives the 64 coefficients of the column $i + 8 * (j - 1)$ of $J_{\Psi}(A)$.
- J_{Θ} is a 8x64 constant matrix (since Θ is a linear transform) and $E_{i,j}.c$ gives the 8 coefficients of the column $i + 8 * (j - 1)$ of J_{Θ} .

In the end, the Jacobian of the transformation which gives \tilde{f} from X is the 8x32 matrix:

$$J_X = J_{\Theta} \cdot J_{\Psi}(\Phi(X)) \cdot J_{\Phi}(X). \quad (4)$$

Since X has covariance matrix $\Sigma_X = \sigma^2 I_{32}$, the covariance of \tilde{f} is the 8x8 matrix

$$\Sigma_{\tilde{f}}(X) = J_X \Sigma_X J_X^T = \sigma^2 J_X J_X^T. \quad (5)$$

Jacobian of SVD. Let us denote $\tilde{F} = \begin{pmatrix} \tilde{f}_1 & \tilde{f}_2 & \tilde{f}_3 \\ \tilde{f}_4 & \tilde{f}_5 & \tilde{f}_6 \\ \tilde{f}_7 & \tilde{f}_8 & 1 \end{pmatrix}$.

As explained earlier, one has to impose $\text{rank}(F) = 2$ by SVD. Let $\tilde{F} = UDV^T$ be the SVD decomposition of F , where U and V are orthogonal matrices and D is the diagonal matrix made of the so-called singular value, sorted in increasing order. Imposing $\text{rank}(F) = 2$ simply consists in computing

$$F = UD \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} V^T \quad (6)$$

²There is a typo here in the proceedings of BMVC'08.

The 9x9 Jacobian matrix of this transformation is then computed. The derivatives:

$$\frac{\partial F}{\partial E_{i,j}} = \frac{\partial U}{\partial E_{i,j}} D \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} V^T + U \frac{\partial D}{\partial E_{i,j}} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} V^T + U D \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \frac{\partial V}{\partial E_{i,j}}^T \quad (7)$$

yield the column $i + 3 * (j - 1)$ of Jacobian matrix $J_{SVD}(\tilde{F})$.

Since coefficient \tilde{F}_{33} is fixed, one gets in the end the covariance matrix of F :

$$\Sigma_F = J_{SVD} \begin{pmatrix} \Sigma_{\tilde{f}} & \mathbf{0}_{8,1} \\ \mathbf{0}_{1,8} & 0 \end{pmatrix} J_{SVD}^T. \quad (8)$$

Papadopolu and Iourakis [14] give an elegant computation of the derivatives $\frac{\partial U}{\partial E_{i,j}}$, $\frac{\partial D}{\partial E_{i,j}}$ and $\frac{\partial V}{\partial E_{i,j}}$. We do not reproduce it here for the sake of conciseness. Let us notice that the Jacobian of the SVD was seemingly first derived in [12]. Let us also remark that the Jacobian of the SVD makes it possible to derive the uncertainty of the least-square estimate of F from $n > 8$ points. In this case, Eq. 2 would indeed become $\tilde{f} = \tilde{f}^+ c$.

3 Epipolar line uncertainty and application to robust fundamental matrix estimation

3.1 Epipolar line uncertainty

We give in this section a digest of the theory from [8], with explicit formulas. We consider the map $\phi: (F, \mathbf{x}) \mapsto \mathbf{l} = \frac{F\mathbf{x}}{\|F\mathbf{x}\|}$ which associate the normalized vector \mathbf{l} representing an epipolar line with a fundamental matrix and an image point. The propagation property (up to a first order approximate) involves that \mathbf{l} has mean $\bar{\mathbf{l}} = \bar{F}\bar{\mathbf{x}}/\|\bar{F}\bar{\mathbf{x}}\|$ and covariance matrix $\Sigma_l = J_F \Sigma_F J_F^T + J_{F,\mathbf{x}} J_{F,\mathbf{x}}^T$. Assuming that F and \mathbf{x} are independent³, one obtains:

$$\Sigma_l = J_F \Sigma_F J_F^T + \sigma^2 J_{\mathbf{x}} J_{\mathbf{x}}^T \quad (9)$$

where J_F and $J_{\mathbf{x}}$ are the Jacobian of ϕ with regards to F and \mathbf{x} .

To permit reproducibility of our work, we give here the explicit formulations of J_F and $J_{\mathbf{x}}$. Let us note $F\mathbf{x} = (L_1, L_2, L_3)^T$ and $N = \|F\mathbf{x}\|^2$. Then

$$J_F = \frac{1}{\|F\mathbf{x}\|} \begin{pmatrix} x - L_1^2 x/N & y - L_1^2 y/N & 1 - L_1^2/N & -L_1 L_2 x/N & -L_1 L_2 y/N & -L_1 L_2/N & -L_1 L_3 x/N & -L_1 L_3 y/N & -L_1 L_3/N \\ -L_1 L_2 x/N & -L_1 L_2 y/N & -L_1 L_2/N & x - L_2^2 x/N & y - L_2^2 y/N & 1 - L_2^2/N & -L_2 L_3 x/N & -L_2 L_3 y/N & -L_2 L_3/N \\ -L_1 L_3 x/N & -L_1 L_3 y/N & -L_1 L_3/N & -L_2 L_3 x/N & -L_2 L_3 y/N & -L_2 L_3/N & x - L_3^2 x/N & y - L_3^2 y/N & 1 - L_3^2/N \end{pmatrix} \quad (10)$$

$$J_{\mathbf{x}} = \frac{1}{\|F\mathbf{x}\|} \begin{pmatrix} F_{21} - L_2(L_1 F_{11} + L_2 F_{21} + L_3 F_{31})/N & F_{22} - L_2(L_1 F_{12} + L_2 F_{22} + L_3 F_{32})/N \\ F_{11} - L_1(L_1 F_{11} + L_2 F_{21} + L_3 F_{31})/N & F_{12} - L_1(L_1 F_{12} + L_2 F_{22} + L_3 F_{32})/N \\ F_{31} - L_3(L_1 F_{11} + L_2 F_{21} + L_3 F_{31})/N & F_{32} - L_3(L_1 F_{12} + L_2 F_{22} + L_3 F_{32})/N \end{pmatrix} \quad (11)$$

Now $(\mathbf{l} - \bar{\mathbf{l}})^T \Sigma_l^{-1} (\mathbf{l} - \bar{\mathbf{l}})$ follows a χ^2 distribution with r degrees of freedom (where r is the rank of Σ_l). It is a consequence of the classical property that $\Sigma_l^{-1/2} (\mathbf{l} - \bar{\mathbf{l}}) \sim \mathcal{N}(0, I_r)$; the random variable is thus the norm of a normalized Gaussian vector. Here $r = 2$ since $\mathbf{l} = F\mathbf{x}/\|F\mathbf{x}\|$ has 2 degrees of freedom.

As shown in [8], with probability α a line \mathbf{l} seen as a realization of the random process $F\mathbf{x}/\|F\mathbf{x}\|$ lies within the hyperbola defined by

³this makes sense if \mathbf{x} is not taken into account when estimating F .

$$C_k = \bar{\mathbf{\Pi}}^T - k^2 \Sigma_1 \quad (12)$$

where $F_2(k^2) = \alpha$ and F_2 is the cumulative χ_2^2 distribution. See [1] for a discussion on the bias introduced by the first order propagation theorem in the estimation of C .

3.2 Adapting Moisan and Stival's ORSA

We derive now from what precedes a robust estimation of F by sorting out inliers from a dataset of correspondences. It is based on Moisan and Stival's ORSA⁴ [11], which is another robust estimation algorithm of the fundamental matrix. ORSA has been shown to outperform existing methods such as M-estimators, LMedS and RANSAC. The main difference with RANSAC-like methods is that the consensus is not build on "hard thresholds", but a measure (called *meaningfulness*) represents a trade-off between the number of detected inliers and their fitness to the model. It is based on a so-called "a contrario" model (see the books [2] and [5] and references therein for a comprehensive account). The probability that a group of correspondences meets the epipolar constraint "by chance" is estimated. The lower this probability, the more likely the correspondences are not casual and are actually due to epipolar constraint.

Let us be more specific. A set of n correspondences $\mathcal{M} = (\mathbf{x}_j, \mathbf{x}'_j)_{1 \leq j \leq n}$ and a subset \mathcal{S} of cardinality $k \geq 8$ being given, a fundamental matrix F is estimated from 8 points of \mathcal{S} . We then estimate the probability that $\max \{d(F\mathbf{x}, \mathbf{x}'), (\mathbf{x}, \mathbf{x}') \in \mathcal{S}\}$ is below a given value δ , under the hypothesis that 1) the \mathbf{x} are fixed, and 2) the \mathbf{x}' are independent and uniformly distributed in the second view. This assumption constitutes the statistical background process. All probabilities in the sequel are computed w.r.t. this process. At the end of the section, we see how to modify this background process in order to take into account uncertainty of F .

Proposition 1 $\Pr\left(\frac{2D}{A} \max \{d(F\mathbf{x}, \mathbf{x}'), (\mathbf{x}, \mathbf{x}') \in \mathcal{S}\} \leq \delta\right) \leq \delta^{k-8}$

where D and A are respectively the diameter and the area of the second view.

Proof: Successively :

$$\begin{aligned} \Pr\left(\frac{2D}{A} \max \{d(F\mathbf{x}, \mathbf{x}'), (\mathbf{x}, \mathbf{x}') \in \mathcal{S}\} \leq \delta\right) &= \Pr\left(\forall (\mathbf{x}, \mathbf{x}') \in \mathcal{S}, \frac{2D}{A} d(F\mathbf{x}, \mathbf{x}') \leq \delta\right) \\ &= \left(\Pr\left(\frac{2D}{A} d(F\mathbf{x}, \mathbf{x}') < \delta\right)\right)^{k-8} \end{aligned}$$

The latter equation stands because $d(F\mathbf{x}, \mathbf{x}') = 0$ for the 8 correspondences used in F estimation and the \mathbf{x}' s are independent.

Now a simple geometrical argument gives $\Pr(d(F\mathbf{x}, \mathbf{x}') < \delta) \leq \frac{2D\delta}{A}$ (see [11]). ■

The meaningfulness of (\mathcal{S}, F) is then defined, motivated by Prop. 2.

Definition 1 (\mathcal{S}, F) is ε -meaningful if $(n-8) \binom{n}{k} \binom{k}{8} \alpha^{k-8} \leq \varepsilon$

with $\alpha = \frac{2D}{A} \max_{(\mathbf{x}, \mathbf{x}') \in \mathcal{S}} d(F\mathbf{x}, \mathbf{x}')$, and k and n defined as above.

Proposition 2 The expected number of ε -meaningful sets in the background process is less than ε .

⁴stands for Optimized Random Sample Algorithm

Proof: The proof is classical (though the presentation here slightly differs from the original one in [11] where the 7 point method is used). Denoting by $N = (n-8) \binom{n}{k} \binom{k}{8}$ and by χ_i the characteristic function of the event “group Nr i is ε -meaningful”, one has: $\mathbb{E}(\chi_i) = \Pr(\alpha \leq (\varepsilon/N)^{1/(k-8)}) \leq \varepsilon/N$ thanks to Prop. 1. Now, N groups are tested (there are $n-8$ choices for $k \geq 8$, $\binom{n}{k}$ choices for the set S , $\binom{k}{8}$ choices for the 8 correspondences to estimate F), thus $\sum_i \mathbb{E}(\chi_i) \leq \varepsilon$. ■

ORSA algorithm consists in randomly choosing 8 correspondences, estimating the corresponding F , building up the most meaningful (i.e. with the lowest ε) consensus set \mathcal{S} based on this F , and iterating. It returns the most meaningful consensus set, which has $\varepsilon \ll 1$. It means that this set cannot be the result of the background process. Def. 1 means that this algorithm realizes a trade-off between the size of the consensus set and the fitting to the model. Let us notice that ORSA has no parameter to be tuned by the user.

Adapting ORSA to the presented probabilistic model. In original ORSA, the background process simply states that the points of the second view are independent and uniformly distributed. Let us adapt it to our framework. We suppose now that the points \mathbf{x}_j and \mathbf{x}'_j ($1 \leq j \leq n$) are independent random variables with normal density $f_{\bar{\mathbf{x}}, \sigma}$ of mean $\bar{\mathbf{x}} = \bar{\mathbf{x}}_j$ (resp. $\bar{\mathbf{x}} = \bar{\mathbf{x}}'_j$) and covariance $\sigma^2 I_2$.

One shall compute for every \mathbf{x} and \mathbf{x}' the probability that a point \mathbf{x}' lies beneath a given distance δ from the epipolar line $F\mathbf{x}$. We now give a precise meaning to the distance d so that $\Pr(d(F\mathbf{x}, \mathbf{x}') \leq \delta)$ can actually be computed. In the original ORSA algorithm, d is simply $2D/A$ multiplied by the Euclidean distance between \mathbf{x}' and the epipole line $F\mathbf{x}$. In fact, as can be seen from the proofs, all what we need is the property $\Pr(d(F\mathbf{x}, \mathbf{x}') \leq \delta) \leq \delta$, which holds in the original ORSA because of this particular choice of d and of the uniform distribution of the points \mathbf{x}' . In our situation \mathbf{x} , \mathbf{x}' , and F are random variables. The probability can be expressed as:

$$\Pr(d(F\mathbf{x}, \mathbf{x}') \leq \delta) = \int_{\mathbf{x}' \in \mathbb{R}^2} f_{\bar{\mathbf{x}}, \sigma}(\mathbf{x}') \Pr(d(F\mathbf{x}, \mathbf{x}') \leq \delta | \mathbf{x}') d\mathbf{x}'. \quad (13)$$

From Eq. 12, we know that with probability $F_2(k^2(\mathbf{x}'))$ (such that $\mathbf{x}'^T C_{k(\mathbf{x}')} \mathbf{x}' = 0$) $F\mathbf{x}$ lies within the conic C_k . We therefore set $d(F\mathbf{x}, \mathbf{x}') = F_2(k^2(\mathbf{x}'))$. Then,

$$\Pr(d(F\mathbf{x}, \mathbf{x}') \leq \delta | \mathbf{x}') = \Pr(F_2(k^2(\mathbf{x}')) | \mathbf{x}') \leq \delta) \leq \delta. \quad (14)$$

This latter inequality stands because when \mathbf{x}' is fixed, $k^2(\mathbf{x}')$ is a realization of a χ_2^2 process and F_2 is precisely the cumulative distribution function of χ_2^2 . As a consequence, $\Pr(d(F\mathbf{x}, \mathbf{x}') \leq \delta) \leq \int_{\mathbf{x}' \in \mathbb{R}^2} f_{\bar{\mathbf{x}}, \sigma}(\mathbf{x}') \delta d\mathbf{x}' = \delta$, and the desired inequality holds.

All what we need is therefore to replace α in Def. 1 by $\alpha' = F_2(k^2(\mathbf{x}'))$. Proposition 2 still holds, the expected number of ε -meaningful sets in the background process⁵ is $< \varepsilon$. Groups with $\varepsilon \ll 1$ are not likely to be due to the background model, independence assumption is indeed not valid and correspondences are due to the epipolar constraints.

Remark that the criterion at hand is $F_2(k^2)$ where

$$k^2 = \frac{(\mathbf{1}^T \mathbf{x}')^2}{\mathbf{x}'^T \Sigma_1 \mathbf{x}'} = \frac{(\mathbf{x}'^T F \mathbf{x})^2}{\|F \mathbf{x}'\|^2 \mathbf{x}'^T \Sigma_1 \mathbf{x}'}. \quad (15)$$

Compared to the classical linear least square estimator (minimize $\Sigma(\mathbf{x}'^T F \mathbf{x})^2$), the constraint is relaxed when the confidence on the epipolar line $F\mathbf{x} = \mathbf{1}$ is poorer. Note that

⁵here, independent points in each view following laws of the kind $\mathcal{N}(\bar{\mathbf{x}}, \sigma^2 I_2)$.

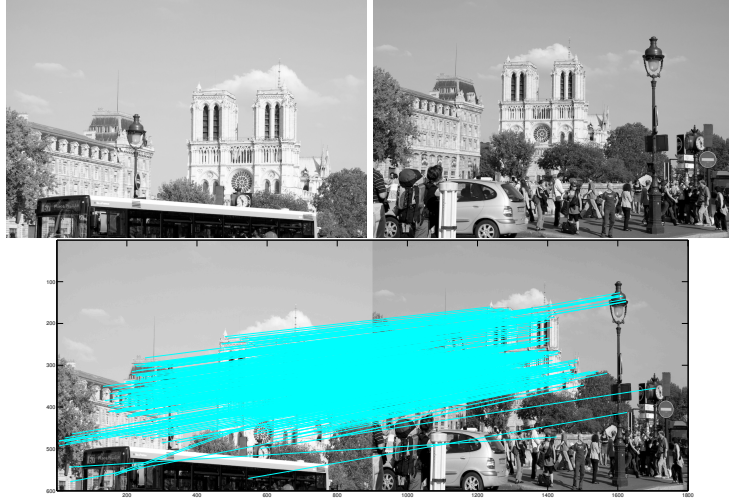


Figure 1: Notre-Dame cathedral seen from two different viewpoints (top) and the 491 Sift correspondences (bottom, correspondences are marked by a straight line between the left and the right points of interest). We use the popular criterion which consists in matching a point of left image with a point of right image if this latter point is the nearest neighbor of the former, and if the ratio between their distance and the distance to the second nearest is upon some threshold [10]. Note that most actual matches belong to the cathedral and the building (in the background), and some of them belong to the tree on the left (in the foreground). There are approximately 10% of wrong matches.

the criterion of [15] is (with our notations) $\frac{(\mathbf{x}'^T F \mathbf{x})^2}{f'^T \Sigma_M f}$. While our criterion is based on the uncertainty on the epipolar lines, this latter criterion directly comes from the Mahalanobis distance between correspondences, that is, the uncertainty on the image points.

Remark that the whole process is symmetrizable w.r.t. the two views, by considering

$$k^2 = \frac{(\mathbf{x}'^T F \mathbf{x})^2}{\|F \mathbf{x}'\|^2 \mathbf{x}'^T \Sigma_I \mathbf{x}'} + \frac{(\mathbf{x}^T F \mathbf{x}')^2}{\|F^T \mathbf{x}\|^2 \mathbf{x}^T \Sigma_I \mathbf{x}} \quad (16)$$

which follows a χ^2 distribution with 4 degrees of freedom. We do not discuss it further.

4 Experimental assessment

We propose in this section a proof-of-concept experiment of the preceding modified-ORSA algorithm. We do not discuss the algorithmic choices here, and simply use the same crafty algorithm as in [11]. Let us recall that the genuine ORSA is based on the 7 point algorithm instead of the 8 point version used here but the difference is not noticeable. See captions of the figures for the discussion.

Note that the aim of what is proposed is not to increase the accuracy of F estimation, but to select a set of inliers as large as possible and not concentrated on e.g. a dominant plane. This is crucial in structure and motion problems: F estimation should be proceed over a set of correspondences that represents well the overall motion. The output of our algorithm is intended as the entry of a non-linear reestimation or a guided matching stage.

References

- [1] S. Brandt. On the probabilistic epipolar geometry. *Image and Vision Computing*, 26(3):405–414, 2006.
- [2] F. Cao, J.L. Lisani, J.-M. Morel, P. Musé, and F. Sur. *A theory of shape identification*. Number 1948 in Lecture Notes in Mathematics. Springer, 2008.
- [3] O. Chum and J. Matas. Matching with PROSAC - progressive sample consensus. In *Proceedings of the International Conference on Computer Vision and Pattern Recognition (CVPR)*, volume 1, pages 220–226, 2005.
- [4] G. Csurka, C. Zeller, Z. Zhang, and O. Faugeras. Characterizing the uncertainty of the fundamental matrix. *Computer Vision and Image Understanding*, 68(1):18–36, 1997.
- [5] A. Desolneux, L. Moisan, and J.-M. Morel. *From Gestalt theory to image analysis: a probabilistic approach*. Interdisciplinary applied mathematics. Springer, 2008.
- [6] O. Faugeras, Q.-T. Luong, and T. Papadopoulos. *The Geometry of Multiple Images*. MIT Press, 2001.
- [7] M. Fischler and R. Bolles. Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography. *Communications of the ACM*, 24(6):381–395, 1981.
- [8] R. Hartley and A. Zisserman. *Multiple View Geometry in Computer Vision*. Cambridge University Press, 2000.
- [9] S. Julier, J. Uhlmann, and H. Durrant-Whyte. A new approach for filtering nonlinear systems. In *Proceedings of the American Control Conference*, volume 3, pages 1628–1632, 1995.
- [10] D. Lowe. Distinctive image features from scale-invariant keypoints. *International Journal of Computer Vision*, 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. *International Journal of Computer Vision*, 57(3):201–218, 2004.
- [12] H. Neudecker and S. Liu. The density of the Moore-Penrose inverse of a random matrix. *Linear algebra and its applications*, 237-238:123–126, 1996.
- [13] B. Ochoa and S. Belongie. Covariance propagation for guided matching. In *Proceedings of the Workshop on Statistical Methods in Multi-Image and Video Processing (SMVP)*, 2006.
- [14] T. Papadopoulos and M.I.A. Lourakis. Estimating the Jacobian of the Singular Value Decomposition: theory and applications. In *Proceedings of the European Conference on Computer Vision (ECCV)*, volume 1, pages 554–570. Springer, 2000.
- [15] Y. Sugaya and K. Kanatani. High accuracy computation of rank-constrained fundamental matrix. In *Proceedings of the British Machine Vision Conference (BMVC)*, volume 1, pages 282–291, 2007.
- [16] P. Torr. Bayesian model estimation and selection for epipolar geometry and generic manifold fitting. *International Journal of Computer Vision*, 50(1):35–61, 2002.
- [17] P. Torr and A. Zisserman. MLESAC: A new robust estimator with application to estimating image geometry. *Computer Vision and Image Understanding*, 78:138–156, 2000.
- [18] Z. Zhang. Determining the epipolar geometry and its uncertainty: a review. *International Journal of Computer Vision*, 27(2):161–195, 1998.
- [19] Z. Zhang and C. Loop. Estimating the fundamental matrix by transforming image points in projective space. *Computer Vision and Image Understanding*, 82(2):174–180, 2001.

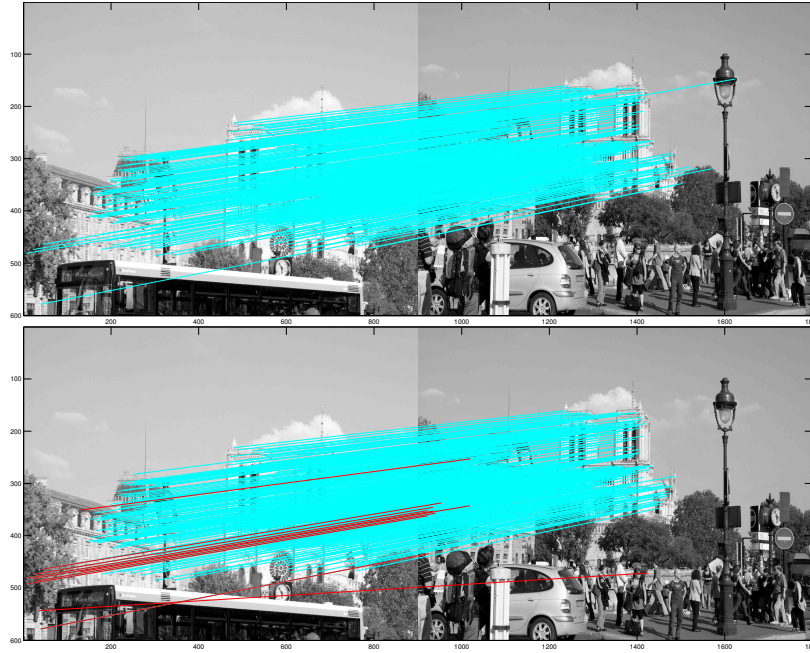


Figure 2: Top: Moisan and Stival's 8 point ORSA; 426 correspondences consistent with the epipolar constraint. Bottom: proposed modified-ORSA: 424 consistent correspondences. The 10 red matches on the bottom pair of images correspond to the most uncertain epipolar lines. All of them are indeed marginal compared with the mass. The fundamental matrix is estimated on the points belonging to the background; the correspondences on the tree on the foreground lead consequently to more uncertain epipolar lines. This is assessed by considering the trace of the covariance matrix of these epipolar lines. The 10 highest value correspond to correspondences between points on the tree (in red). One wrong correspondence can be seen; however Fig. 4 shows that the distance to the epipolar is quite low, such a case simply cannot be resolved from two views.

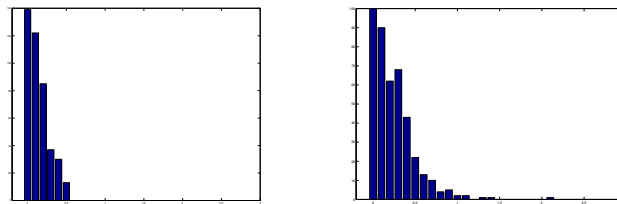


Figure 3: Histograms of the distances from the image points in the right image to the epipolar line computed from the corresponding image point in the left image. Left: ORSA, right: modified-ORSA. All image points from the ORSA consensus set lie less than half a pixel from the corresponding epipolar. The proposed modified-ORSA is less conservative w.r.t. this point and enables distances up to 2 pixels. Indeed, our criterion does not minimize the distance to the epipolar and is relaxed w.r.t. the uncertainty of the estimated epipolar lines. Note that these histograms show that our method cannot be reduced to simply increasing the distance threshold.