

Comparing Fundamental Matrices with a Normalized F -Statistic

Nathan D. Cahill
 Eastman Kodak Company
 343 State Street, Rochester, NY, USA 14650-1816
 nathan.cahill@kodak.com

Abstract

This article introduces the Normalized F -Statistic (NFS), a measure for comparing estimates of the fundamental matrix. As opposed to other measures commonly used for comparison, such as the Frobenius norm of the difference of the estimates, and the RMS or mean squared error between points and their corresponding epipolar lines, the NFS yields a measure of confidence that one estimation technique is superior to another. This confidence measure is useful in comparing the performance of two techniques with respect to a specific image or to draw general conclusions about their relative performances.

1. Introduction

1.1 Overview

This article introduces the Normalized F -Statistic (NFS), a measure for comparing estimates of the fundamental matrix (the 3×3 matrix that describes the epipolar geometry of two images). Section 1 will briefly describe epipolar geometry and mention some of the measures that have been used in previous research to compare fundamental matrix estimates. In Section 2, the NFS is derived using the classical assumption of independent and identically distributed normal errors in the location of corresponding points. Some computational considerations are detailed, and the derivation is extended to handle corresponding points with heteroscedastic errors. Section 3 illustrates the use of the NFS to compare four standard fundamental matrix estimation techniques applied to 102 pairs of real images. Finally, in Section 4, some conclusions about the efficacy of the NFS are drawn from the results of Section 3.

1.2 Epipolar Geometry

Epipolar geometry describes a natural constraint that arises when two images of the same scene are captured. Consider a real world point \mathbf{M} captured by a camera with optical center C and another camera with optical center C' , as shown in Fig. 1:

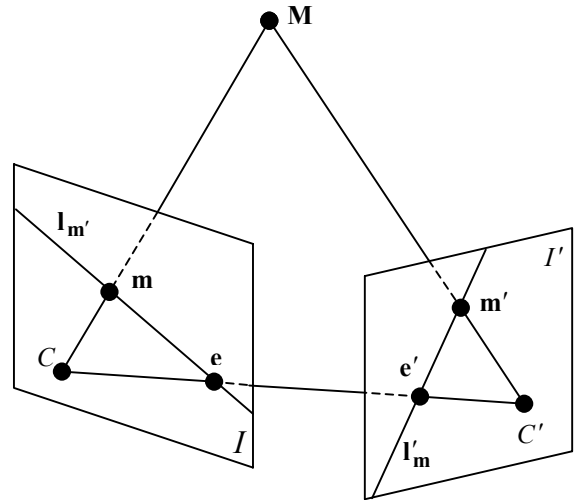


Figure 1: Epipolar geometry.

If \mathbf{M} is projected onto image I at location \mathbf{m} , the projection of \mathbf{M} onto I' (denoted \mathbf{m}') must fall on the epipolar line \mathbf{l}'_m (the intersection of I' with the plane containing \mathbf{m} , C and C'). Conversely, if \mathbf{M} is projected onto image I' at location \mathbf{m}' , the projection of \mathbf{M} onto I (denoted \mathbf{m}) must fall on the epipolar line \mathbf{l}_m (the intersection of I with the plane containing \mathbf{m}' , C and C'). Because the epipole \mathbf{e}' is the projection of C onto I' , all epipolar lines of points in I must pass through \mathbf{e}' . Similarly, all epipolar lines of points in I' must pass through the epipole \mathbf{e} .

The *fundamental matrix* is a 3×3 matrix \mathbf{F} that projects a point \mathbf{m} in homogeneous coordinates (i.e., $\mathbf{m} = (u, v, 1)^T$) to its epipolar line. That is, $\mathbf{l}'_m = \mathbf{F}\mathbf{m}$. Because the epipolar constraint ensures that \mathbf{m}' lies on \mathbf{l}'_m (or $\mathbf{m}'^T \mathbf{l}'_m = 0$), it must be true that

$$\mathbf{m}'^T \mathbf{F}\mathbf{m} = 0. \quad (1)$$

Transposing Eq. (1), we find that \mathbf{F}^T is the fundamental matrix when the role of the two cameras is reversed. By the geometry illustrated in Fig. 1, we can see that the

epipolar line l'_e of the epipole e reduces to a point; therefore, we have:

$$l'_e = Fe = \mathbf{0}. \quad (2)$$

and, conversely,

$$l_e = F^T e' = \mathbf{0}. \quad (3)$$

From Eqs. (2) or (3), it is clear that F has, at most, rank 2. (In general, the rank of F will be equal to 2.) It is also clear based on Eqs. (1)-(3) that F is unique only up to a scale factor. Hence, F has seven independent parameters. It is therefore possible to determine F exactly given seven (nondegenerate) sets of corresponding points. In many instances, more than seven point correspondences are known, so some estimation strategy that optimally fits the point correspondences must be applied. A variety of estimation techniques have been developed based on minimizing algebraic error, geometric error, covariance weighted error, and Sampson error [1]-[7]. If some of the point correspondences are mismatched, we can use robust statistical estimation techniques [8], [9].

1.3 Previous Work

Several measures have been used in the research literature to compare two estimates of the fundamental matrix (or to compare an estimate with the “true” fundamental matrix). Zhang [6] describes a few such measures, including the Frobenius norm of the difference between the two fundamental matrix estimates, the RMS distance between points and their corresponding epipolar lines, and an estimate of the average distances between randomly chosen points and their corresponding epipolar lines, computed by treating the two images symmetrically. Chesi et al. [1] use the mean distance between points and their epipolar lines. Luong and Faugeras [5] measure the relative error between coordinates of the epipoles of estimated and true fundamental matrices in order to quantify the sensitivity of the estimation technique.

As illustrated empirically by Zhang, it is clear that measuring the difference between fundamental matrix terms in the Frobenius norm (or in any norm) does not necessarily relate to a similar difference according to any geometric criterion. The methods that measure average or RMS distances between points and their corresponding epipolar lines do report a more geometrically meaningful number. However, making a direct comparison between two different fundamental matrix estimates, based on determining which estimate has the lesser measure, is akin to comparing two probability distributions solely based on which one has the lesser mean. A more appropriate comparison measure should include some notion of the confidence that one fundamental matrix estimate is superior to another.

2. The Normalized F -Statistic

2.1 Deriving the NFS

The benefit of the Normalized F -Statistic (NFS) over other measures of comparison is that the NFS relays a measure of the confidence that one fundamental matrix estimate is superior to another. To derive the NFS, we must first consider the distribution of the distance from points to their corresponding epipolar lines.

Slightly modifying the notation of Section 1.2, consider that N corresponding point pairs, $\{(\mathbf{m}_i, \mathbf{m}'_i), i = 1, \dots, N\}$ are observed, where \mathbf{m}_i is a point in image I , and \mathbf{m}'_i is the corresponding point in image I' , and all points are represented in homogeneous coordinates. Under the classical assumption, the observed corresponding points differ from the true corresponding points by i.i.d. normal random noise in the image plane; i.e., $\mathbf{m}_i = \tilde{\mathbf{m}}_i + \boldsymbol{\varepsilon}_i$ and $\mathbf{m}'_i = \tilde{\mathbf{m}}'_i + \boldsymbol{\varepsilon}'_i$, where $(\tilde{\mathbf{m}}_i, \tilde{\mathbf{m}}'_i)$ is the true corresponding point pair, and $\boldsymbol{\varepsilon}_i$ and $\boldsymbol{\varepsilon}'_i$ are i.i.d. with

$$\boldsymbol{\varepsilon}_i \sim Normal\left(\mathbf{0}, \sigma^2 \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}\right). \quad (4)$$

(The heteroscedastic case will be considered later in this section.)

If we consider F_1 and F_2 to be estimates of the fundamental matrix, the distance from a point in image I to its corresponding epipolar line is given by:

$$d(\mathbf{m}_i, F_j^T \mathbf{m}'_i) = \frac{|\mathbf{m}_i^T F_j^T \mathbf{m}'_i|}{\sqrt{(F_j^T \mathbf{m}'_i)_1^2 + (F_j^T \mathbf{m}'_i)_2^2}}. \quad (5)$$

Expanding and substituting, we have:

$$d(\mathbf{m}_i, F_j^T \mathbf{m}'_i) = k_{i,j} \left| \tilde{\mathbf{m}}_i^T F_j^T \tilde{\mathbf{m}}'_i + \boldsymbol{\varepsilon}_i^T F_j^T \tilde{\mathbf{m}}'_i + \tilde{\mathbf{m}}_i^T F_j^T \boldsymbol{\varepsilon}'_i + \boldsymbol{\varepsilon}_i^T F_j^T \boldsymbol{\varepsilon}'_i \right|, \quad (6)$$

where

$$k_{i,j} = \frac{1}{\sqrt{(F_j^T \tilde{\mathbf{m}}'_i)_1^2 + (F_j^T \tilde{\mathbf{m}}'_i)_2^2}}. \quad (7)$$

Therefore, a first order approximation yields:

$$d(\mathbf{m}_i, F_j^T \mathbf{m}'_i) \approx k_{i,j} \left| \tilde{\mathbf{m}}_i^T F_j^T \tilde{\mathbf{m}}'_i + \boldsymbol{\varepsilon}_i^T F_j^T \tilde{\mathbf{m}}'_i + \tilde{\mathbf{m}}_i^T F_j^T \boldsymbol{\varepsilon}'_i \right|. \quad (8)$$

From Eq. (4), and using properties of the multivariate normal distribution, we find that $\tilde{\mathbf{m}}_i^T F_j^T \tilde{\mathbf{m}}'_i + \boldsymbol{\varepsilon}_i^T F_j^T \tilde{\mathbf{m}}'_i + \tilde{\mathbf{m}}_i^T F_j^T \boldsymbol{\varepsilon}'_i$ is a normally distributed random variable with mean $\tilde{\mathbf{m}}_i^T F_j^T \tilde{\mathbf{m}}'_i$ and variance

$$\sigma^2 \omega_{i,j}^2 = \sigma^2 \left[(F_j \tilde{\mathbf{m}}_i)_1^2 + (F_j \tilde{\mathbf{m}}_i)_2^2 + (F_j^T \tilde{\mathbf{m}}'_i)_1^2 + (F_j^T \tilde{\mathbf{m}}'_i)_2^2 \right]. \quad (9)$$

Hence, $d(\mathbf{m}_i, \mathbf{F}_j^T \mathbf{m}'_i)^2 / (\omega_{i,j} k_{i,j})^2$ has approximately a noncentral chi-square distribution with one degree of freedom and noncentrality parameter

$$\lambda_{i,j} = (\tilde{\mathbf{m}}_i^T \mathbf{F}_j^T \tilde{\mathbf{m}}'_i)^2 / \omega_{i,j}^2. \quad (10)$$

Now, if we consider separate pairs of corresponding points to be independent of one another,

$$S_j = \sum_{i=1}^N \frac{(\mathbf{m}_i^T \mathbf{F}_j^T \mathbf{m}'_i)^2}{\omega_{i,j}^2} \quad (11)$$

has approximately a noncentral chi-square distribution with N degrees of freedom and noncentrality parameter

$$\lambda_j = \sum_{i=1}^N \lambda_{i,j}. \quad (12)$$

The ratio S_2/S_1 therefore has approximately a doubly noncentral F -distribution [10] with parameters $(N-1, N-1, \lambda_2, \lambda_1)$. We use this result to define the *NFS*:

$$NFS_{i,j} = H_{N-1, N-1, \lambda_j, \lambda_i}(S_j/S_i), \quad (13)$$

where $H_{n_1, n_2, d_1, d_2}(x)$ is the cumulative distribution function of the doubly noncentral F -distribution with parameters (n_1, n_2, d_1, d_2) , evaluated at x .

The *NFS* can be interpreted as a confidence level at which we can reject the hypothesis that both fundamental matrix estimates yield similar errors. Therefore, the closer the *NFS* $\nu_{1,2}$ is to one, the more probable it is that \mathbf{F}_1 is the better estimate. The closer the *NFS* is to zero, the more probable it is that \mathbf{F}_2 is the better estimate. The closer the *NFS* is to 0.5, the more probable it is that neither estimate is better than the other.

2.2 Computational Considerations

Clearly, the *NFS* cannot be computed exactly, because Eqs. (9) and (10) require the unknown locations of true corresponding points. Without knowledge of the 3-D positions of points in the original scene, the best one can do is approximate the true image positions. We suggest three different approximations, each with their own advantages and disadvantages.

One possibility is to approximate the true locations with the observed locations; i.e., $\tilde{\mathbf{m}}_i \approx \mathbf{m}_i$ and $\tilde{\mathbf{m}}'_i \approx \mathbf{m}'_i$. This approximation does have precedent (S_j becomes exactly the gradient criterion described in Luong and Faugeras [5]); however, in forcing the assumption that all observed locations are true, the *NFS* is no longer a valid measure because the variance of the observation errors is zero by definition.

Another possibility is to apply the Gold Standard method of fundamental matrix estimation [4] to the observed point correspondences. The Gold Standard

method, in addition to estimating the fundamental matrix, also yields estimates of the true corresponding point locations. The disadvantage to this approximation is that it assumes the Gold Standard produces a reliable fundamental matrix estimate (of which there is no guarantee), at least with respect to the estimates under comparison.

A third possibility is to approximate the true point locations with the orthogonal projection in the image plane of the observed point onto its corresponding epipolar line:

$$\begin{aligned} \tilde{\mathbf{m}}_i &\approx (\mathbf{I} - \mathbf{U}_{\mathbf{m}'_i} - \mathbf{u}_{\mathbf{m}'_i} \mathbf{u}_{\mathbf{m}'_i}^T) \mathbf{m}_i, \\ \tilde{\mathbf{m}}'_i &\approx (\mathbf{I} - \mathbf{U}'_{\mathbf{m}_i} - \mathbf{u}'_{\mathbf{m}_i} \mathbf{u}'_{\mathbf{m}_i}{}^T) \mathbf{m}'_i, \end{aligned} \quad (14)$$

where

$$\begin{aligned} \mathbf{u}_{\mathbf{m}'_i} &= \frac{\mathbf{l}_{\mathbf{m}'_i}}{\sqrt{\mathbf{l}_{\mathbf{m}'_i}^T \mathbf{l}_{\mathbf{m}'_i}}}, \quad \mathbf{u}'_{\mathbf{m}_i} = \frac{\mathbf{l}'_{\mathbf{m}_i}}{\sqrt{\mathbf{l}'_{\mathbf{m}_i}{}^T \mathbf{l}'_{\mathbf{m}_i}}}, \\ \mathbf{U}_{\mathbf{m}'_i} &= \begin{pmatrix} (\mathbf{u}_{\mathbf{m}'_i})_3^2 & 0 & 0 \\ 0 & (\mathbf{u}_{\mathbf{m}'_i})_2^2 & 0 \\ -(\mathbf{u}_{\mathbf{m}'_i})_1 (\mathbf{u}_{\mathbf{m}'_i})_3 & -(\mathbf{u}_{\mathbf{m}'_i})_2 (\mathbf{u}_{\mathbf{m}'_i})_3 & 0 \end{pmatrix}, \end{aligned}$$

and

$$\mathbf{U}'_{\mathbf{m}_i} = \begin{pmatrix} (\mathbf{u}'_{\mathbf{m}_i})_3^2 & 0 & 0 \\ 0 & (\mathbf{u}'_{\mathbf{m}_i})_2^2 & 0 \\ -(\mathbf{u}'_{\mathbf{m}_i})_1 (\mathbf{u}'_{\mathbf{m}_i})_3 & -(\mathbf{u}'_{\mathbf{m}_i})_2 (\mathbf{u}'_{\mathbf{m}_i})_3 & 0 \end{pmatrix}.$$

The disadvantage to this approach is that “true” point locations depend on the fundamental matrix estimate, so they will be different for the estimates under comparison. However, this approach does have the advantage that it centralizes the doubly noncentral F -distribution, because $\lambda_{i,j}$ becomes identically zero for all i and j . This, in turn, guarantees that the *NFS* satisfies $NFS_{i,j} = 1 - NFS_{j,i}$, which is a very useful property because it enables a symmetric interpretation.

2.3 The *NFS* with Heteroscedastic Data

In the derivation of the *NFS*, we assumed that the noise added to true corresponding point locations was independent and identically distributed, according to Eq. (4). Frequently, this assumption is too simple. As a result of a variety of factors, such as depth of points in the scene, nonstationary blurring of the image, orientation of points along edges or corners in the scene, etc., it may be more prudent to model the errors as normal random variates, each with their own covariance matrix. That is, let $\boldsymbol{\varepsilon}_i$ be multivariate normal with mean $\mathbf{0}$ and covariance matrix Σ_i , and let $\boldsymbol{\varepsilon}'_i$ be multivariate normal with mean $\mathbf{0}$ and covariance matrix Σ'_i . (We will still assume that $\boldsymbol{\varepsilon}_i$ and $\boldsymbol{\varepsilon}'_i$ are independent).

Under this general model, the variance $\omega_{i,j}^2$ from Eq. (9) becomes:

$$\omega_{i,j}^2 = \tilde{\mathbf{m}}_i^T \mathbf{F}_j^T \Sigma'_i \mathbf{F}_j \tilde{\mathbf{m}}_i + \tilde{\mathbf{m}}_i'^T \mathbf{F}_j \Sigma_i \mathbf{F}_j^T \tilde{\mathbf{m}}_i'. \quad (15)$$

Therefore, S_j is computed as in (11), but with $\omega_{i,j}^2$ given by (15). The NFS is then computed as in (13).

3. Results

3.1 Case 1: Church

The *church* sequence, shown in Fig. 2, and developed at INRIA [11], comprises 12 images of a church, the first ten of which are captured at approximately 8° intervals around the church by a handheld camera, and the remaining two at some spurious positions.

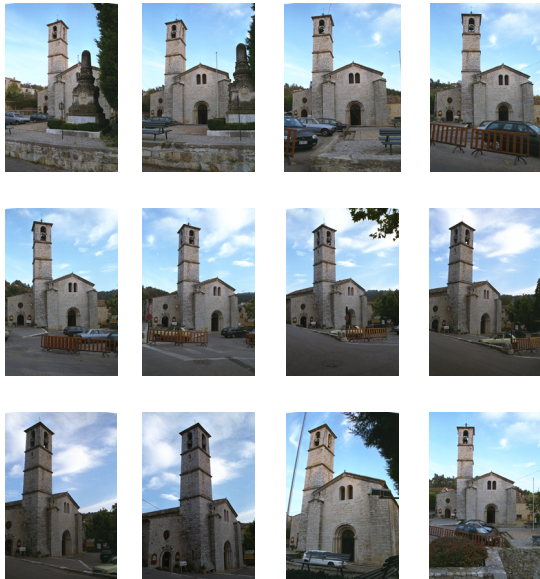


Figure 2: Church sequence of images.

Typically, an arbitrary pair of images from this sequence differs by some amount of rotation of the camera about the church. In the induced epipolar geometry, both epipoles should lie outside the borders of each image. For example, Fig. 3 shows church images #6 and #8 with epipolar lines that were computed after estimating the fundamental matrix from a collection of manually supplied correspondence points.



Figure 3: Church images #6 and #8 with epipolar lines.

As can be seen from Fig. 3, the epipole, or intersection of the epipolar lines in the image plane, in image #6 lies to the left of image #6, and the epipole in image #8 lies to the right of image #8.

For each of the 66 combinations of pairs of church images, corresponding points were manually labeled. Using these corresponding points, the fundamental matrix was estimated via four standard techniques: Hartley's Eight-Point Algorithm [3], minimization of the distance to epipolar lines [5], minimization of the gradient criterion [5], and Least Median of Squares [6]. The NFS is computed in order to compare each of the six pairs of estimation techniques for each pair of images, and the results are plotted in the following figures. The weight used is that of Eq. (9), and the approximation to the true corresponding points is that of Equation (14).

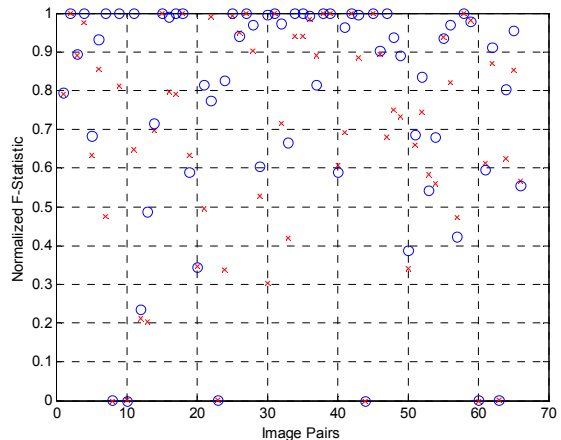


Figure 4: NFS (Hartley, DIST): Red X
NFS (Hartley, GRAD): Blue O

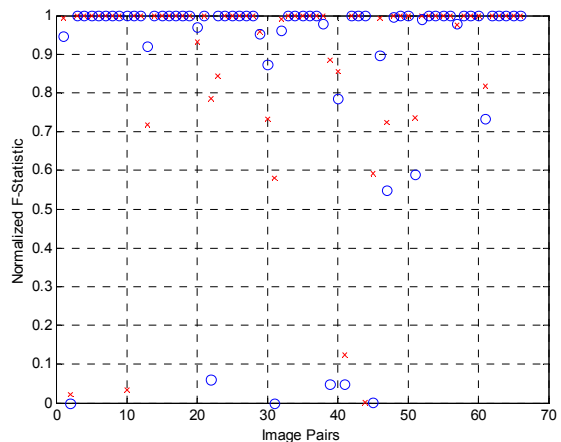


Figure 5: NFS (Hartley, LMS): Red X
NFS (DIST, LMS): Blue O

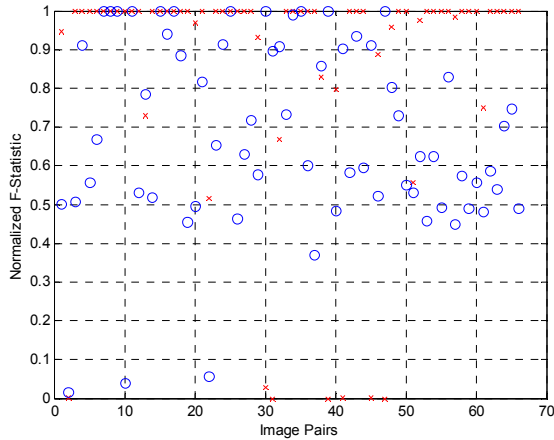


Figure 6: NFS (GRAD, LMS): Red X
NFS (DIST, GRAD): Blue O

3.2 Case 2: Corridor

The *corridor* sequence, shown in Fig. 7 and developed at INRIA [11], comprises nine images of the interior of a corridor, captured at three sets of various distances and focal lengths.



Figure 7: Corridor sequence of images.

In the corridor sequence, arbitrary pairs of images typically differ by a movement of the camera in the direction of the corridor, with possibly an additional slight rotation and zoom. Figure 8 shows corridor images #2 and #9 with some epipolar lines, again computed from a fundamental matrix estimate that used manually supplied correspondence points. In these images, it can easily be seen that the epipoles lie within the boundaries of each image.

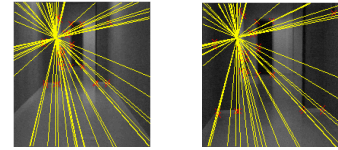


Figure 8: Corridor images #2 and #9 with epipolar lines.

As in the previous subsection, for each of the 36 combinations of pairs of corridor images, corresponding points were manually labeled. The NFS is computed in order to compare each of the six pairs of estimation techniques for each pair of images, and the results are plotted in the following figures.

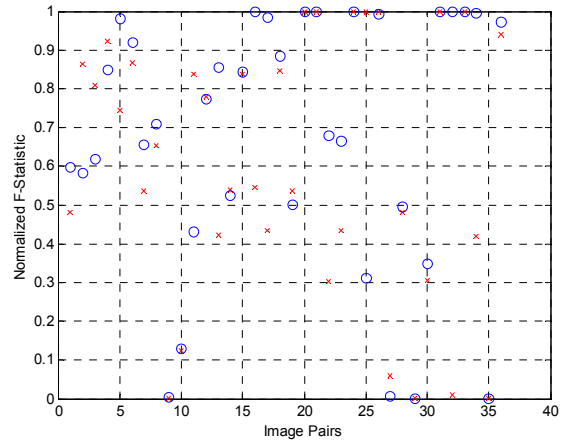


Figure 9: NFS (Hartley, DIST): Red X
NFS (Hartley, GRAD): Blue O

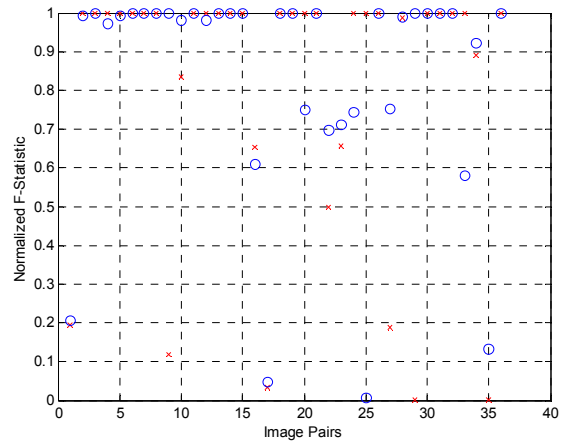


Figure 9: NFS (Hartley, LMS): Red X
NFS (DIST, LMS): Blue O

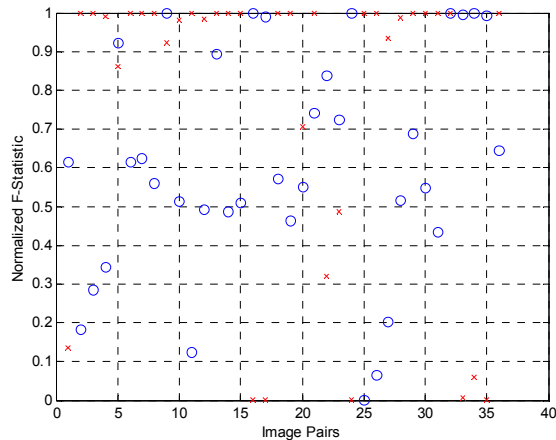


Figure 10: NFS (GRAD, LMS): Red X
NFS (DIST, GRAD): Blue O

4. Discussion

As can be seen in Figs. 4-6 and 8-10, we can gain some insight into how each fundamental matrix estimation technique performs relative to the other techniques for each individual image, and in general. For example, Figs. 4 and 8 seem to indicate that the DIST and GRAD techniques provide better estimates than Hartley's algorithm. However, a few select pairs of church and corridor images do yield a better estimate with Hartley's method. Perhaps this is due to the nonlinear techniques getting stuck in local minima.

Figures 5 and 9 lead us to believe that the DIST and GRAD criteria are equally useful in generating a good fundamental matrix estimate. These figures and Figs. 6 and 10 would also indicate that the LMS technique is superior to all others in the vast majority of image pairs. This conclusion is the intuitive one: an estimate that minimizes the median squared error will tend to make half of the terms of Eq. (11) small. However, a handful of pairs of church and corridor images do show that the LMS technique does NOT perform as well as the other techniques.

These results indicate the usefulness of the NFS as a measure of the confidence that one fundamental matrix estimation technique performs better than another, both in general, and on specific pairs of images.

Acknowledgements

Thanks are due to Tom Brooks, Stan Holmes, and Chris Wrzos-Lucacci for their help in manually labeling corresponding points for this research.

References

- [1] G. Chesi, A. Garulli, A. Vicino, and R. Cipolla, Estimating the Fundamental Matrix via Constrained Least-Squares: A Convex Approach, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 24(3), pp. 397-401, March 2002.
- [2] O. Faugeras and Q.-T. Luong, *The Geometry of Multiple Images*. MIT Press, 2001.
- [3] R. Hartley, In Defense of the Eight-Point Algorithm, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 19(6), June 1997.
- [4] R. Hartley and A. Zisserman, *Multiple View Geometry in Computer Vision*. Cambridge University Press, 2000.
- [5] Q.-T. Luong and O. Faugeras, The Fundamental Matrix: Theory, Algorithms, and Stability Analysis, *International Journal of Computer Vision*, 17, pp. 43-75, 1996.
- [6] Z. Zhang, Determining the Epipolar Geometry and its Uncertainty: A Review, *INRIA Research Report No. 2927*, July 1996.
- [7] Z. Zhang and C. Loop, Estimating the Fundamental Matrix by Transforming Image Points in Projective Space, *Computer Vision and Image Understanding*, 82, pp. 174-180, 2001.
- [8] P. Torr and D. Murray, The Development and Comparison of Robust Methods for Estimating the Fundamental Matrix, *International Journal of Computer Vision*, 24(3), pp. 271-300, 1997.
- [9] Z. Zhang, R. Deriche, O. Faugeras and Q.-T. Luong. A Robust Technique for Matching Two Uncalibrated Images Through the Recovery of the Unknown Epipolar Geometry, *INRIA Research Report No. 2273*, May 1994.
- [10] M. Paoletta and R. Butler, Calculating the Density and Distribution Function for the Singly and Doubly Noncentral F, *Statistics and Computing*, 12(1), pp.9-16, 2002.
- [11] <http://www.inria.fr>