# Approximation to the Fisher-Rao Metric for the Focus of Expansion

S.J. Maybank

10 April 2007

*School of Computer Science and Information Systems, Birkbeck College, Malet Street,*
*London, WC1E 7HX, UK.*
*sjmaybank@dcs.bbk.ac.uk*

**Abstract.** The Fisher-Rao metric for the focus of expansion is approximated, under the assumption that the focus is estimated from correspondences between two images taken by a translating camera. The approximation is accurate if the errors in the image correspondences are small. The parameter space for the focus of expansion is sampled at a finite set of points chosen such that every point in the space is near to at least one sample point. A focus of expansion is detected by checking each sample point in turn to see if it is supported by the image correspondences.

*Keywords:* false detection; Fisher-Rao metric; focus of expansion; Riemannian metric, translation.

## 1 Introduction

Many tasks in computer vision involve the detection of a structure using measurements obtained from one or more images. The structure could be a simple line or a curve or it could be more abstract, for example the projective transformation defined by the pairs of corresponding points in two images of a line. The range of structures of a given type is described by a parameter space $T$, where each point of $T$ specifies a unique structure.

The volume of $T$ is a measure of the difficulty of detecting the structure in question: if the volume is large, then it is difficult to search the whole of $T$ to find all the structures compatible with a given set of measurements. At first sight this idea of using volume to measure the difficulty of structure detection appears to be unworkable because the same range of structures can be described by different parameter spaces with different volumes. For example, the chords in the unit disk can be parameterised by their mid points. The resulting parameter space, considered as a subset of $\mathbb{R}^2$, has an area (or two-dimensional volume) $\pi$. An alternative parameterisation is obtained by extending each chord to a line in $\mathbb{R}^2$ and taking the intercept and the gradient of each line. The area of this second parameter space is infinite.

The difficulty is resolved by calculating the volume of $T$ using the Fisher-Rao metric [3,4] on $T$. The Fisher-Rao metric is directly related to structure detection. If two points of $T$ are close together in the Fisher-Rao metric, then the corresponding structures are difficult to distinguish using the measurements. This observation leads to a simple algorithm for detecting structures. A set $G$ of points is chosen from $T$, such that each point of $T$ is near to at least one point of $G$ under the Fisher-Rao metric. Then each point of $G$ is checked in turn to see if the presence of the corresponding structure is supported by the measurements. If the volume of $T$ is finite and if $T$ has a "reasonable" shape then the number $|G|$ of points in $G$ is finite and proportional to the volume of $T$.

A major obstacle to finding $G$ is that there are in general very few cases in which the Fisher-Rao metric can be obtained in a closed form or even approximated numerically in a practical way. It is shown in [12] that the Fisher-Rao metric can be approximated by a more tractable metric which can in some cases be obtained in closed form. The approximation is accurate if the measurement noise is low. The approximating metric has been obtained for lines in the unit disc [10], lines in a rectangular image [14], projective transformations of the line [9,11] and ellipses [13]. In the current paper the approximating metric is obtained for the focus of expansion associated with the images taken by a single camera as it undergoes a pure translational motion without rotation. The focus of expansion is the point in the image plane that defines the direction of motion of the camera. If the camera is moving forward then the scene points appear to diverge from the focus of expansion. If the camera is moving backwards then the scene points appear to converge on the focus of expansion [6].

The literature on estimating the focus of expansion from image sequences is vast. For an early review see, for example, [2]. The current paper is the first one to investigate the Fisher-Rao metric for the focus of expansion. The approximation to the Fisher-Rao metric obtained in this paper is important because *i)* it is the basis of a simple and reliable search algorithm for detecting all the foci of expansion compatible with a given set of image correspondences; *ii)* it can be used to assess any algorithm for detecting foci of expansion, for example to see whether the algorithm in question has missed any foci; *iii)* it may in future provide the information needed to design neural nets or networks of processors for detecting foci of expansion: the optimal spacing between adjacent nodes or between adjacent processors can be estimated using the approximating metric.

The approximating metric is obtained in Section 2. A numerical investigation of the approximating metric is carried out in Section 3. This investigation includes the calculation of the volume of the parameter space. An algorithm to find the focus of expansion is described in Section 4, the probabilities of false detection and false rejection are analyzed in Section 5 and a method for sampling the parameter space is described in Section 6. Experimental results are reported in Section 7 and some concluding remarks are made in Section 8.

Most of the symbolic calculations and all of the numerical work reported in this paper were carried out using Mathematica [18].

# 2   Mathematical Framework

## 2.1   Image Formation

It is assumed that the image points are obtained by projection to a plane in accordance with the pinhole camera model, as described in [5,6]. For mathematical convenience, it is assumed that the image itself is the unit disk $D$ centred at the origin. The plane containing $D$ is referred to as the image plane, even though a large part of this plane is outside the visible image.

Suppose that the camera translates from one position to another, without undergoing any rotation, and let $c$ be the projection of the translation vector. The point $c$ is known as the focus of expansion. If the camera is moving backwards, then the translation vector

points behind the camera. In this case the translation vector is reflected in the origin and then projected to the image plane. The focus of expansion remains fixed during the motion of the camera, and it may be inside $D$ or outside $D$.

Let a scene point project to a point $q_1$ in $D$ when the camera is in the first position and to a point $q_2$ in $D$ when the camera is in the second position. The points $q_1$, $q_2$ are said to be corresponding points, $q_1 \leftrightarrow q_2$. The three points $c$, $q_1$, $q_2$ are collinear because the motion of the camera is a pure translation. For further information, see [6], Section 10.1.3. Conversely, suppose that the three points $c$, $q_1$, $q_2$ are collinear. Then the ray projecting to $q_1$ in the first image is coplanar with the ray projecting to $q_2$ in the second image. If the two rays intersect, then the intersection is a scene point projecting to $q_1$ and to $q_2$. If the two rays are parallel, then it is convenient to introduce a scene point at infinity, which projects to $q_1$ and to $q_2$.

## 2.2 The joint image

Each pair $q_1$, $q_2$ of points in $D$, whether corresponding or not, defines a point $q = (q_1, q_2)$ in $D^2 \subset \mathbb{R}^4$. The space $D^2$ obtained in this way is referred to as the measurement space. It is closely related to the joint image space defined in [17]. The Euclidean metric in $D^2$ is compatible with the Euclidean metric in $D$, in that

$$\|q\|^2_{D^2} = \|q_1\|^2_D + \|q_2\|^2_D.$$

The subscripts for the various Euclidean metrics such as $\|.\|_{D^2}$, $\|.\|_D$ are omitted from now on. If $q_1 \leftrightarrow q_2$, then $q$ is on a hypersurface $H(c)$ in $D^2$. Let the Cartesian coordinates of $c$, $q_i$ be $(c_1, c_2)$ and $(q_{i1}, q_{i2})$ respectively. The equation of $H(c)$ is given by

$$f(q, c) \equiv \det \begin{pmatrix} q_{11} & q_{12} & 1 \\ q_{21} & q_{22} & 1 \\ c_1 & c_2 & 1 \end{pmatrix} = 0. \tag{1}$$

## 2.3 Conditional density for a measurement

The probability density function $p(q|c)$ is defined for a measurement $q$ in $D^2$, conditional on the focus of expansion $c$. It is assumed that $q$ arises as a perturbation of an error free measurement $\tilde{q}$ on $H(c)$ and that the probability density function for the error $q - \tilde{q}$ is Gaussian, of the form

$$(2\pi\sigma^2)^{-2} \exp\left(-\|q - \tilde{q}\|^2/(2\sigma^2)\right), \tag{2}$$

where $\sigma > 0$ is the standard deviation of the error in single component of $q$. Let the error free measurement, $\tilde{q}$, be distributed randomly on $H(c)$ with a probability measure $d\mu$. It follows from (2) that

$$p(q|c) = (2\pi\sigma^2)^{-2} \int_{H(c)} \exp\left(-\|q - \tilde{q}\|^2/(2\sigma^2)\right) d\mu(\tilde{q}).$$

## 2.4 The Fisher-Rao metric

The focus of expansion $c$ takes values in a parameter space $T$ which happens to coincide with the image plane $\mathbb{R}^2$ which contains $D$. The Fisher-Rao metric [3,4], $J(c)$, on $T$ is

defined by

$$J_{ij}(c) = -\int_{D^2} \left(\partial^2_{c_i c_j} \ln p(q|c)\right) p(q|c) dq, \qquad i = 1, 2 \text{ and } j = 1, 2.$$

The image $D$ has two metrics, namely the Euclidean metric for measuring distances between image points, and the Fisher-Rao metric, inherited from $T \supset D$, for measuring distances between possible choices for the focus of expansion.

## 2.5 Approximation to the Fisher-Rao metric

Under the assumption that $q$ is close to $H(c)$, let $w(q, c)$ be the signed distance from $q$ to $H(c)$. If $\sigma$ is small, then $q$ is close to $H(c)$ with a high probability, and the Fisher-Rao metric, $J(c)$, is approximated by the Riemannian Metric, $M(c)$, defined by

$$M_{ij}(c) = \frac{1}{2\sigma^2} \int_{H(c)} \left(\partial^2_{u_i u_j} w(q, u)^2\right)_{q=\tilde{q}, u=c} d\mu(\tilde{q}), \qquad i = 1, 2 \text{ and } j = 1, 2. \tag{3}$$

See [12] for further information. The approximation of $M(c)$ to $J(c)$ increases in accuracy as $\sigma$ tends to zero.

Let $q$ be a point near to $H(u)$ and let $q + a$ be the point on $H(u)$ nearest to $q$. On taking the Taylor expansion of $f(q + a, u)$ it follows that

$$0 = f(q + a, u) = f(q, u) + a.\partial_q f(q, u) + O_2(a),$$

where $O_2(a)$ is a term of order two in $a$. If $\partial_q f(q, u) \neq 0$ and if $\|a\|$ is sufficiently small, then

$$a = -\frac{f(q, u)\partial_q f(q, u)}{\|\partial_q f(q, u)\|^2} + O_2(a). \tag{4}$$

It follows from (4) that to leading order,

$$w(q, u)^2 = \|a\|^2 = \frac{f(q, u)^2}{\|\partial_q f(q, u)\|^2}, \tag{5}$$

It follows from (3) and (5) that

$$M_{ij}(c) = \frac{1}{\sigma^2} \int_{H(c)} \left(\frac{(\partial_{c_i} f(q, u))(\partial_{c_j} f(q, u))}{\|\partial_q f(q, u)\|^2}\right)_{q=\tilde{q}, u=c} d\mu(\tilde{q}), \qquad i = 1, 2 \text{ and } j = 1, 2. \tag{6}$$

The measure $d\mu$ is chosen proportional to the canonical measure [7] associated with the Riemannian metric induced on $H(c)$ by the Euclidean metric on the space $D^2$ containing $H(c)$.

## 2.6 The hypersurface $H(c)$

In order to evaluate the right-hand side of (6) it is necessary to describe the hypersurface $H(c)$ in more detail. Let $o$ be the origin in $D$ and let the polar coordinates of $c$ be $(r, \theta)$. Let $q = (q_1, q_2)$ be a point on $H(c)$, and let $\phi$ be the angle between the lines $\langle o, c \rangle$ and the
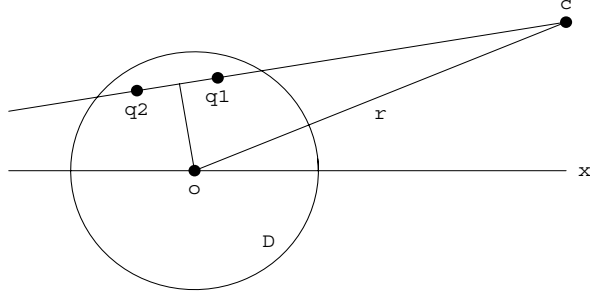
Figure 1: Position of the focus of expansion $c$.

perpendicular to the line $\langle c, q_2 \rangle$ through $o$. The positions of the various points are shown in Fig. 1 for the case in which $c$ is outside $D$.

Let $u_i$ be the signed distance from $c$ to $q_i$, $i = 1, 2$, measured from $c$ along the line $\langle c, q_2 \rangle$. It is assumed that $u_1$, $u_2$ have the same sign, or equivalently that $c$ is not between $q_1$ and $q_2$. If $c$ is outside the image, then $c$ cannot be between $q_1$ and $q_2$ for any choice of $q_1 \leftrightarrow q_2$ on $H(c)$. On the other hand, if $c$ is inside the image, then certain choices of $q_1 \leftrightarrow q_2$ are excluded. The excluded image correspondences are physically impossible, given that the movements of the points in the image are due to a pure translational motion of the camera. The sign of the $u_i$ is chosen such that the foot of the perpendicular from $o$ to $\langle c, q_2 \rangle$ corresponds to $u_i = r \sin(\phi)$.

If $c$ is outside the image, then the ranges of $u_1$, $u_2$, $\phi$ are

$$r \sin(\phi) - (1 - r^2 \cos^2(\phi))^{1/2} \leq u_i \leq r \sin(\phi) + (1 - r^2 \cos^2(\phi))^{1/2}, \qquad i = 1, 2, \qquad (7)$$

$$-\pi/2 \leq \phi \leq -\cos(1/r) \text{ and } \cos^{-1}(1/r) \leq \phi \leq \pi/2. \qquad (8)$$

If $c$ is inside the image then the ranges of $u_1$, $u_2$, $\phi$ are

$$
\begin{aligned}
0 &\leq u_1, u_2 \leq r \sin(\phi) + (1 - r^2 \cos^2(\phi))^{1/2}, \\
r \sin(\phi) - (1 - r^2 \cos^2(\phi))^{1/2} &\leq u_1, u_2 \leq 0, \\
-\pi/2 &\leq \phi \leq \pi/2.
\end{aligned}
\qquad (9)
$$

The coordinates of the $q_i$ are

$$q_i = r \cos(\phi)(\cos(\phi+\theta), \sin(\phi+\theta)) + (u_i - r \sin(\phi))(-\sin(\phi+\theta), \cos(\phi+\theta)), \quad i = 1, 2. \quad (10)$$

Let $L$ be the Riemannian metric induced on $H(c)$ by the Euclidean metric on the space $D^2$ which contains $H(c)$, and let $V(L, H(c))$ be the volume of $H(c)$ under the Riemannian metric $L$. As noted at the end of Section 2.5, the measure $d\mu$ on $H(c)$ is chosen to be proportional to the canonical measure associated with $L$,

$$d\mu(q(\phi, u_1, u_2)) = \frac{|\det(L)|^{1/2}}{V(L, H(c))} d\phi \, du_1 \, du_2. \qquad (11)$$

The matrix for $L$ in the $(\phi, u_1, u_2)$ coordinate system is obtained by solving

$$(d\phi, du_1, du_2) L \begin{pmatrix} d\phi \\ du_1 \\ du_2 \end{pmatrix} = \|q(\phi + d\phi, u_1 + du_1, u_2 + du_2) - q(\phi, u_1, u_2)\|^2 + O_3, \qquad (12)$$

5

where $O_3$ consists of terms of degree three or higher in $d\phi$, $du_1$, $du_2$. It follows from (10), (11) and (12) that

$$L = \begin{pmatrix} u_1^2 + u_2^2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \tag{13}$$

$$V(L, H(c)) = \int_{H(c)} |\det(L)|^{1/2}\, d\phi\, du_1 du_2,$$

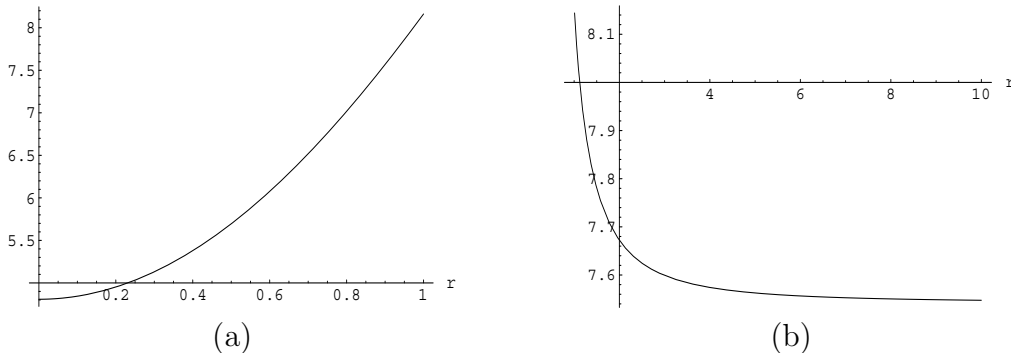$$= \int_{H(c)} (u_1^2 + u_2^2)^{1/2}\, d\phi\, du_1 du_2. \tag{14}$$



(a)

(b)

Figure 2: (a) the volume $V(L, H(c))$ as a function of $r = \|c\|$ for $0 \le r \le 1$; (b) the volume $V(L, H(c))$ as a function of $r$ for $1 \le r \le 10$.

Let $r$, $\theta$ be polar coordinates on $T$, with $c = (r, \theta)$. The volume, $V(L, H(c))$, is independent of $\theta$. Graphs of $V(L, H(c))$ as a function of $r$ are shown in Fig. 2. It is clear from Fig. 2 that $V(L, H(c))$ has a maximum at $r = 1$. The value of this maximum is

$$\max_{r \ge 0} V(L, H(c)) = \frac{32}{9}(\sqrt{2} + \sinh^{-1}(1)). \tag{15}$$

## 2.7 Approximating metric for the focus of expansion

It follows from (1),(6),(11) and (13) that the approximation to the Fisher-Rao metric on the parameter space $T$ is given in a Cartesian coordinate system by

$$M(c) = (\sigma^2 V(L, H(c)))^{-1} \times$$

$$\int_{H(c)} \frac{(u_1 - u_2)^2}{(u_1^2 + u_2^2)^{1/2}} \begin{pmatrix} \cos^2(\phi + \theta) & \cos(\phi + \theta)\sin(\phi + \theta) \\ \cos(\phi + \theta)\sin(\phi + \theta) & \sin^2(\phi + \theta) \end{pmatrix} d\phi\, du_1\, du_2.$$

Let $H$ be the Jacobian matrix for the transformation from Cartesian coordinates $(c_1, c_2)$ to polar coordinates $(r, \theta)$, where $c_1 = r\cos(\theta)$, $c_2 = r\sin(\theta)$. The matrix $H$ is defined by

$$H = \begin{pmatrix} \partial c_1/\partial r & \partial c_1/\partial \theta \\ \partial c_2/\partial r & \partial c_2/\partial \theta \end{pmatrix} = \begin{pmatrix} \cos(\theta) & -r\sin(\theta) \\ \sin(\theta) & r\cos(\theta) \end{pmatrix}.$$

6

The approximation $K(c)$ to the Fisher-Rao metric is given in polar coordinates by

$$K(c) = H^\top M(c)H,$$

$$= \frac{1}{\sigma^2 V(L, H(c))} \int_{H(c)} \frac{(u_1 - u_2)^2}{(u_1^2 + u_2^2)^{1/2}} \begin{pmatrix} \cos^2(\phi) & r\sin(2\phi) \\ r\sin(2\phi) & r^2\sin^2(\phi) \end{pmatrix} d\phi \, du_1 \, du_2. \qquad (16)$$

The Riemannian metric $K(c)$ is independent of $\theta$. Where convenient the notation $K(r)$ will be used in place of $K(c)$.

On making the substitutions $u_1 \mapsto -u_1$, $u_2 \mapsto -u_2$, $\phi \mapsto -\phi$, in (16), it can be seen that the diagonal entries $K_{12}(c)$ and $K_{21}(c)$ are zero for $r \leq 1$ and for $r \geq 1$.

# 3 Numerical investigation

The integrals defining $K_{11}(r)$ and $K_{22}(r)$ can be evaluated numerically for any given value of $r \geq 0$. The numerical integration is simplified by carrying out one of the integrations on the right hand side of (16) symbolically,

$$\int \frac{(u_1 - u_2)^2}{(u_1^2 + u_2^2)^{1/2}} \, du_1 = \left(\frac{1}{2}u_1 - 2u_2\right)(u_1^2 + u_2^2)^{1/2} + \frac{1}{2}u_2^2 \ln\left(u_1 + (u_1^2 + u_2^2)^{1/2}\right).$$

The remaining two integrations are carried out numerically. In the special case $r = 1$, all the integrals can be evaluated symbolically to give

$$K(1) = \frac{4 - 3\sqrt{2} + \sinh^{-1}(1)}{5\sigma^2(\sqrt{2} + \sinh^{-1}(1))} \begin{pmatrix} 1 & 0 \\ 0 & 4 \end{pmatrix}.$$

In sections 3.1 and 3.2, Taylor series approximations to $K(r)$ are given, and the volume, $V(K, T)$, of $T$ is estimated numerically.

## 3.1 Focus of expansion inside the image

The Taylor series approximation to $K(r)$, accurate at small values of $r$, is

$$K(r) = \frac{4 - 3\sqrt{2} + \sinh^{-1}(1)}{32\sigma^2(\sqrt{2} + \sinh^{-1}(1))} \begin{pmatrix} 16 - 18r^2 + 15r^4 & 0 \\ 0 & 2r^2(8 + 9r^2) \end{pmatrix} + O(r^5). \qquad (17)$$

The Taylor series (17), with the $O(r^5)$ error term omitted, is compared to the numerical estimate of $K(r)$ in Fig. 3.

Let $T_1$ be the subset of $T$ for which the focus of expansion is inside $D$,

$$T_1 = \{c, c \in \mathbb{R}^2 \text{ and } \|c\| \leq 1\}.$$

The volume, $V(K, T_1)$, of $T_1$ is estimated,

$$\begin{aligned} V(K, T_1) &= \int_{T_1} |\det(K(r))|^{1/2} \, dr \, d\phi, \\ &= \frac{0.3993...}{\sigma^2}. \end{aligned} \qquad (18)$$

7

## 3.2  Focus of expansion outside the image

The Taylor series approximation to $K(r)$, accurate at large values of $r$, is

$$K(r) = \frac{4}{15\sigma^2 r^4} \begin{pmatrix} 1/7 & 0 \\ 0 & r^4 - 19r^2/105 - 13/1573 \end{pmatrix} + O(r^{-5}). \qquad (19)$$

The Taylor series (19), with the $O(r^{-5})$ error term omitted, is compared to the numerical estimates of $K(r)$ in Fig. 4.
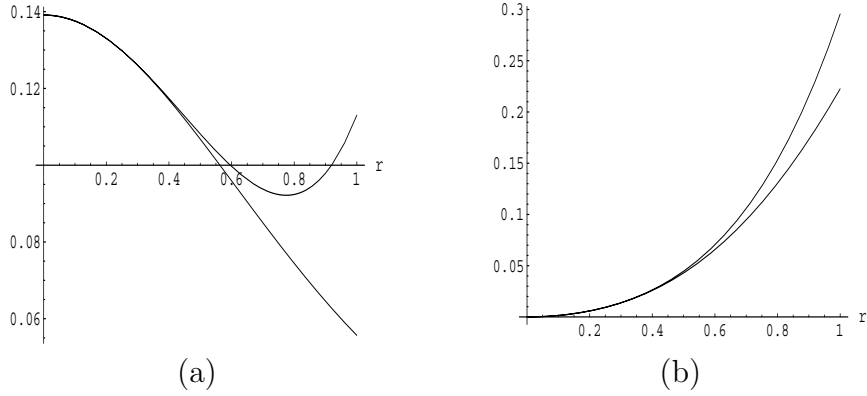


(a)           (b)

Figure 3: (a) Numerical approximation to $K_{11}(r)$, $r \leq 1$ (lower graph) and Taylor series expansion of $K_{11}(r)$ in powers of $r$ (upper graph); (b) numerical approximation to $K_{22}(r)$, $r \leq 1$ (lower graph) and Taylor series expansion of $K_{22}(r)$ in powers of $r$ (upper graph). In all cases, $\sigma = 1$.
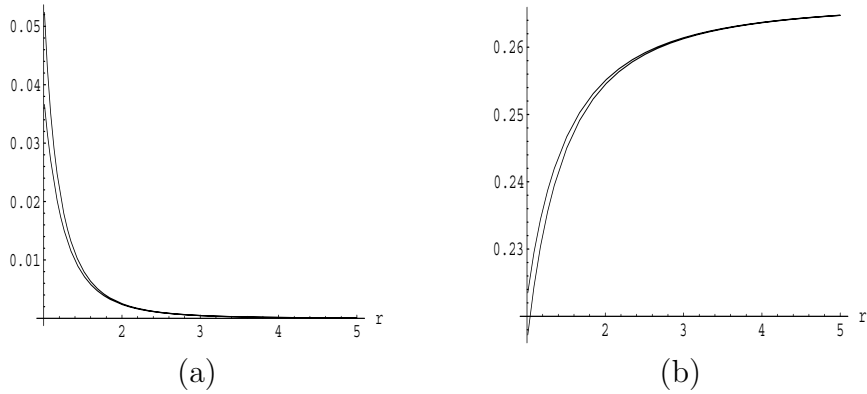


(a)           (b)

Figure 4: (a) Numerical approximation to $K_{11}(r)$, $r \geq 1$ (upper graph) and Taylor series expansion of $K_{11}(r)$ in powers of $r^{-1}$ (lower graph); (b) numerical approximation to $K_{22}(r)$, $r \geq 1$ (upper graph) and Taylor series expansion of $K_{22}(r)$ in powers of $r^{-1}$ (lower graph). In all cases, $\sigma = 1$.

Let $T_2$ be the subset of $T$ for which the focus of expansion is outside $D$,

$$T_2 = \{c, c \in \mathbb{R}^2 \text{ and } \|c\| \geq 1\}.$$

The volume, $V(K, T_2)$, of $T_2$ is estimated,

$$V(K, T_2) = \int_{T_2} |\det(K(r))|^{1/2} \, dr \, d\theta,$$

$$= \frac{0.6425\ldots}{\sigma^2}. \tag{20}$$

In order to evaluate the integral on the right hand side of (20), the range $[1, \infty)$ is divided into two parts, $[1, 5]$ and $[5, \infty)$. In $[1, 5)$ a numerical approximation to $K(r)$ is used. In $[5, \infty)$ the Taylor series (19) for $K(r)$ is used, with the $O(r^{-5})$ error term omitted.

It follows from (20) and (18) that the total volume of $T$ is

$$V(K, T) = V(K, T_1) + V(K, T_2) = \frac{1.0418\ldots}{\sigma^2}. \tag{21}$$

It can be shown using numerical integration that 99% of the volume of $T$ is contained within the disk centred at the origin and with Euclidean radius $1000\sigma$.

# 4 Algorithm for Detecting the Focus of Expansion

Many algorithms for detecting the focus of expansion $c$ have two stages. In the first stage, pairs of corresponding points are obtained from two images, and in the second stage the values of $c$ supported by the corresponding points are found. In the noise free case just two pairs $q(1) = (q_1(1), q_2(1))$, $q(2) = (q_1(2), q_2(2))$ of corresponding points are required, because

$$c = \langle q_1(1), q_2(1)\rangle \cap \langle q_1(2), q_2(2)\rangle.$$

The theoretical calculations in Sections 2 and 3 suggest the following strategy for estimating $c$ from a given set $S(N)$ of $N$ measurements of pairs of corresponding image points: sample $T$ to obtain a finite set $G$ of points such that every point in $T$ is near to at least one point of $G$,

$$\max_{c' \in T} \min_{c \in G} \operatorname{dist}_K(c', c) \le 1, \tag{22}$$

where $\operatorname{dist}_K(c', c)$ is the distance between $c'$ and $c$ under the Riemannian metric $K$ on $T$ [7]. Then check each $c \in G$ to see if $S(N)$ supports the choice of $c$ as a focus of expansion. The value 1 on the right hand side of (22) is suggested in [11], Section 4. It ensures that the average value of the log likelihood ratio, $\ln(p(q|c)/p(q|c'))$, is less than $2^{-1}$. It is not assumed that all the elements of $S(N)$ support the same or even any focus of expansion. The size $|G|$ of $G$ is given approximately by

$$|G| \approx \pi^{-1} V(T, K), \tag{23}$$

because the set of points $c'$ in $T$ such that $\operatorname{dist}_K(c', c) \le 1$ has volume under $K$ approximately equal to $\pi$.

The support of $S(N)$ for a focus of expansion $c$ is measured in the following way. Let $M$ be a positive integer with $M \le N$, let $\tau$ be a threshold and let $p_0$ be a background density on $D^2$. Then $S(N)$ supports $c$ if it contains a subset $S(M)$ of size $M$ such that

$$\min_{q \in S(M)} (p(q|c)/p_0(q)) > \tau. \tag{24}$$

9

The reason for choosing (24) is that it is not assumed a priori that all the measurements in $S(M)$ are sampled from $p(q|c)$. If (24) holds, then all the elements of $S(M)$ are near to $H(c)$, and it is plausible to assume that they are all sampled from $H(c)$.

In order to implement the above algorithm it is necessary to choose $M$ and $\tau$. The values of these two parameters depend on the performance required from the algorithm, or more specifically, the probability, $e_f$, of a false detection and the probability, $e_r$, of a false rejection. It this shown in Section 5 how to calculate $M$, $\tau$ from given values of $e_f$, $e_r$.

# 5 False Detection and False Rejection

A false detection of a focus of expansion $c$ occurs when $S(N)$ supports the choice $c$, even though there is no casual link between $S(N)$ and $c$. In such cases, the support for $c$ is a random effect, without any real significance. A false rejection occurs if there is a casual link between $S(N)$ and $c$, but the errors in the measurements are large enough to obscure the link, to the extent that $S(N)$ does not support the choice of $c$ as a focus of expansion.

An upper bound on the probability, $e_f$, of a false detection is obtained in Section 5.1 and an expression for the probability, $e_r$, of a false rejection is obtained in Section 5.2. The upper bound for $e_f$ is obtained under the assumption that the background density, $p_0$, is the uniform density on $D^2$.

The experimental evidence reported in Section 7 shows that the upper bound on $e_f$ obtained in this section is too low. In practice the probability of a false detection is larger than the proposed upper bound because the outlying image correspondences are not uniformly distributed in $D^2$. They often have regularities or patterns, which are not related to any image motions but which nevertheless suggest the presence of a focus of expansion. For example, if the first image contains a small textured region which "attracts" correspondences with points in a larger region of the second image then this will increase the probability of detecting a focus of expansion in the small textured region.

## 5.1 False detection

Let $A(c) \subset D^{2M}$ be the set on which (24) holds. The probability that $S(M)$ is in $A(c)$, given that the elements of $S(M) = \{q(1), \ldots, q(M)\}$ are sampled independently from $D^2$ using the background density $p_0$ is

$$\int_{A(c)} p_0(S(M)) \, dq(1), \ldots, dq(M). \tag{25}$$

The upper bound for the probability of obtaining a false detection from $S(N)$ is obtained by maximising (25) over $c \in G$ and then summing over all choices of $c$ and $S(M)$,

$$|G| \binom{N}{M} \max_{c \in G} \left\{ \int_{A(c)} p_0(S(M)) \, dq(1) \ldots, dq(M) \right\}. \tag{26}$$

The logarithm of the density $p(q|c)$ is approximated by

$$\ln(p(q|c)) = -\frac{w(q, c)^2}{2\sigma^2} - \ln(\sigma) + O(\sigma^0), \tag{27}$$

10

as described in [12]. The $O(\sigma^0)$ term in (27) is discarded. It follows from (24) and (27) that $S(M)$ supports $c$ if and only if

$$\max_{q \in S(M)} w(q, c)^2 < 2\sigma^2(-\ln(\sigma) - \ln(\tau)). \tag{28}$$

It is assumed that the right hand side of (28) is positive. Let $\rho$ be defined by

$$\rho = \sqrt{2}\sigma(-\ln(\sigma) - \ln(\tau))^{1/2}, \tag{29}$$

and let $C(M, 2\rho)$ be the hypercube in $\mathrm{I\!R}^M$, centred at the origin and with side length $2\rho$. It follows from (28) and (29) that $S(M)$ supports $c$ if and only if

$$(w(q(1), c), \ldots, w(q(M), c)) \in C(M, 2\rho).$$

Let $E$ be the Euclidean metric on $\mathrm{I\!R}^M$ and let $V(E, C(M, 2\rho))$ be the volume of $C(M, 2\rho)$ under $E$. It is assumed that the threshold $\tau$ has a comparatively high value, and that as a result $S(M)$ supports $c$ only if all the points $q(i)$ in $S(M)$ are close to the hypersurface $H(c)$, or equivalently only if $\rho$ is small. Each point $q(i)$ is associated with a unique point $\tilde{q}(i)$ on $H(c)$ such that $\|q(i) - \tilde{q}(i)\| = |w(q(i), c)|$. Under these assumptions,

$$\begin{aligned}
\int_{A(c)} p_0(S(M))dq(1)\ldots dq(M) &=& \pi^{-2M} \int_{A(c)} dq(1)\ldots dq(M), \\
&\approx& \pi^{-2M} V(E, C(M, 2\rho)) \int_{H(c)^M} d\mu(\tilde{q}(1))\ldots d\mu(\tilde{q}(M)) \\
&\approx& \pi^{-2M} V(E, C(M, 2\rho)V(L, H(c))^M. \tag{30}
\end{aligned}$$

It follows from (26), (29) and (30) that the following expression is an upper bound for $e_f$,

$$\pi^{-2M-1} \binom{N}{M} V(K, T) V(E, C(M, 2\rho)) \left(\max_{c \in G} V(L, H(c))^M\right). \tag{31}$$

On applying (15), (31) yields the upper bound

$$\pi^{-1} V(K, T) \binom{N}{M} \left(\frac{64\rho}{9\pi^2}(\sqrt{2} + \sinh^{-1}(1))\right)^M. \tag{32}$$

## 5.2   False Rejection

A false rejection occurs when a set of image correspondences is sampled from the density $p(q|c)$ but the sample is unusual, to the extent that it does not support $c$ as a focus of expansion. The probability, $e_r$, that $c$ is falsely rejected, given that the set $S(N)$ of image correspondences contains a subset $S(M) = \{q(1), \ldots, q(M)\}$ sampled from $p(q|c)$ is

$$e_r = 1 - \int_{A(c)} \prod_{i=1}^{M} p(q(i)|c)\, dq(1)\ldots dq(M). \tag{33}$$

The worst case is assumed, in which all the image correspondences in $S(N) \setminus S(M)$ are sampled from the uniform background density $p_0(q)$. This leaves open the possibility of

a false detection of $c$ based on $S(N) \setminus S(M)$, but the probability of this latter event is omitted.

Each $q(i)$ in $S(M)$ is obtained by applying a random perturbation to a point $\tilde{q}(i)$ in $H(c)$. The perturbation is along the normal to $H(c)$ at $\tilde{q}(i)$ and the signed magnitude $w(q(i), c)$ of the perturbation has a Gaussian distribution with expected value 0 and standard deviation $\sigma$. The points $\tilde{q}(i)$ are distributed independently on $H(c)$ in accordance with the probability measure $d\mu$ defined in (11). Let $\Phi$ be the cumulative distribution function for the Gaussian distribution with expected value 0 and standard deviation 1. It follows from (33) that

$$
\begin{aligned}
e_r &= 1 - \prod_{i=1}^{M} \left( \frac{1}{(2\pi\sigma^2)^{1/2} V(L, H(c))} \int_{H(c)} d\mu(\tilde{q}(i)) \int_{-\rho}^{\rho} \exp\left( -\frac{w(q(i), c)^2}{2\sigma^2} \right) dw(q(i), c) \right), \\
&= 1 - (2\Phi(\rho/\sigma) - 1)^M.
\end{aligned}
\tag{34}
$$

## 5.3   Relation between false detection and false rejection

In applications, $\sigma$, $e_f$, $e_r$ are often specified, the measurements $S(N)$ are obtained from two images and it is necessary to choose $M$ such that the probabilities of false detection and false rejection are no higher that $e_f$, $e_r$ respectively. At the same time $M$ should be as small as possible, to maximize the probability of finding a sufficient number of inliers.
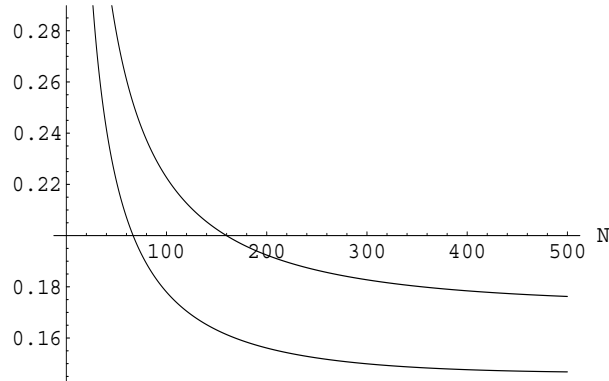


Figure 5: Upper curve: $N \mapsto M(N)/N$, $20 \le N \le 500$, with $\sigma = 10^{-2}$, $e_f = e_r = 10^{-2}$; lower curve: $N \mapsto M(N)/N$, $20 \le N \le 500$, with $\sigma = 10^{-2}$ and $e_f = e_r = 10^{-1}$.

As an example, let $\sigma = 10^{-2}$ and $e_f = e_r = 10^{-2}$. The value of $M$ as a function of $N$ is obtained by first solving (34) for $\rho$ and then finding the value of $M$ for which (32) is equal to $e_f$. The graph of $N \mapsto M(N)/N$ for $20 \le N \le 500$ is shown as the upper curve in Figure 5. The lower curve is the graph of $N \mapsto M(N)/N$, $20 \le N \le 500$ for $e_f = e_r = 10^{-1}$. The value of $\sigma$ is unchanged.

An inspection of the curves in Fig. 5 suggests that for large $M(N)$, $N$, the ratio $M(N)/N$ might tend to a limit as $N \to \infty$. In fact $M(N)/N$ does not tend to a limit as $N \to \infty$; instead $M(N)/N$ continues to decrease, but at an ever slower rate. To investigate further, it is assumed that $e_r$ is small and that $M$ is relatively large. It follows from (34) that

$$
\rho = \sigma \Phi^{-1} \left( 2^{-1}(1 + (1 - e_r)^{1/M}) \right).
\tag{35}
$$

12

On setting $e_f$ equal to (32), taking the $M$th root of both sides of the resulting equation, setting quantities of order $O(1)^{1/M}$ equal to 1 and setting $(\sigma^2 e_f)^{1/M}$ equal to 1, it follows that

$$1 = \frac{64}{9\pi^2}(\sqrt{2} + \sinh^{-1}(1))\sigma \left(\frac{N}{M}\right)^{1/M} \Phi^{-1}\left(2^{-1}(1 + (1 - e_r)^{1/M})\right). \qquad (36)$$

Let $\xi = M(N)/N$. A short calculation based on Stirling's formula [1] shows that if $M$, $N$ and $N - M$ are large, then

$$\left(\frac{N}{M}\right)^{1/M} \approx \xi^{-1}(1 - \xi)^{1-\xi^{-1}}.$$

It can be shown that

$$\frac{\partial}{\partial M}\Phi^{-1}(2^{-1}(1 + (1 - e_r)^{1/M})) = O(M^{-1}),$$

thus, when $M$ is large, $\Phi^{-1}(2^{-1}(1 + (1 - e_r)^{1/M}))$ increases slowly with $M$. It can be estimated by replacing $M$ by an upper bound such as $\sigma^{-2}$, which is an upper bound on the number of distinguishable points in $D$. It follows from (36) and these observations that $\xi$ can be estimated by solving

$$1 = \frac{64}{9\pi^2}(\sqrt{2} + \sinh^{-1}(1))\sigma\xi^{-1}(1 - \xi)^{1-\xi^{-1}}\Phi^{-1}\left(2^{-1}(1 + (1 - e_r)^{\sigma^2})\right). \qquad (37)$$

Some numerical values of $\xi$ are given in Table 1.

| $\sigma$ | $e_r = e_f$ | predicted from (37) | $M(N)/N$ |
|---|---|---|---|
| 0.01 | 0.1 | 0.18 | 0.15 |
| 0.01 | 0.01 | 0.20 | 0.18 |
| 0.001 | 0.1 | 0.024 | 0.029 |
| 0.001 | 0.01 | 0.025 | 0.035 |

Table 1. Third column: predicted values of $\xi = M(N)/N$ obtained from (37). Fourth column: values of $M(N)/N$ obtained from (32) and (33).

# 6   Method for Sampling the Parameter Space

An efficient method for sampling $T$ is required. The parameter space $T$ consist of the whole plane, $\mathbb{R}^2$, however the volume of $T$ under the Riemannian metric $K$ is finite. It is thus possible to find a bounded region of $\mathbb{R}^2$ which contains a large proportion of the total volume of $T$. As noted in Section 3.2, 99% of the volume of $T$ under $K$ is contained in the disk centred at the origin and with Euclidean radius $1000\sigma$. The set $G$ of sample points is chosen from this disk.

The points of $G$ are chosen on circles centred at the origin and with Euclidean radii $r_1, r_2, \ldots, r_m$. These radii are chosen such that $r_1 = 0$, $\text{dist}_K((r_i, 0), (r_{i+1}, 0)) = 1$, $r_m \leq 1000\sigma$ and $r_{m+1} > 1000\sigma$. Let $C_i$ be the circumference of the circle with Euclidean radius $r_i$, $1 \leq i \leq m$. The value of $C_i$ is

$$C_i = \int_0^{2\pi} (K_{22}(r_i))^{1/2}\, d\phi = 2\pi(K_{22}(r_i))^{1/2}. \qquad (38)$$

It follows from (19) and (38) that if $r_i$ is large, then $C_i$ approaches $4\pi/(\sqrt{15}\,\sigma)$.

Each circle is sampled by spacing points uniformly around it such that the arc length between two adjacent points, as measured using $K$, is at most $\sqrt{3}$. This choice of points is based on the following observation: let $a$, $b$ be two points in the plane such that $\|a - b\| = \sqrt{3}$. The circles of unit radius centered at $a$, $b$ respectively, intersect in two points, each one of which is at a distance of $1/2$ from the line $\langle a, b \rangle$. If the $i$th circle is sufficiently large then the set of points within a distance $1/2$ of a sample point on the circle, as measured in the $K$ metric, contains an annulus of width approximately 1, with $i$th circle running along the middle. The points of G are shown in Fig. 6 for $\sigma = 10^{-2}$.
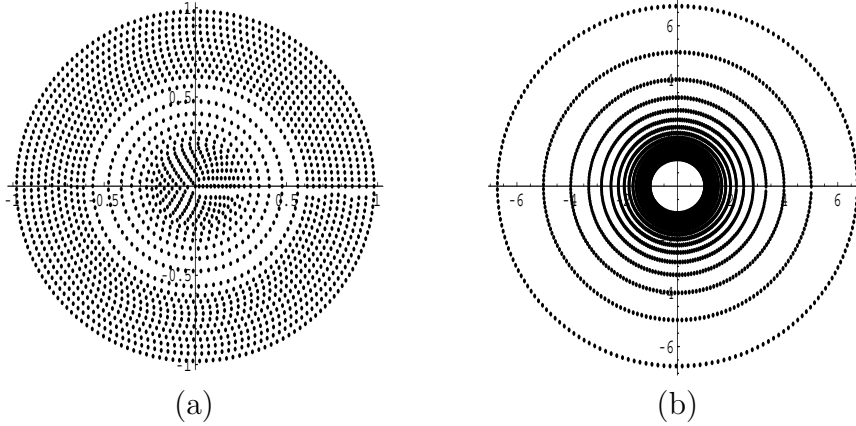


(a)                                    (b)

Figure 6: (a) points of $G$ in the image $D$; (b) points of $G$ outside $D$.

# 7  Experiment

The feasibility of the algorithm for detecting the focus of expansion was investigated using the two images in Fig. 6, taken from the "Map" data set available at

   http://www.middlebury.edu/stereo.

For further information about these images see [15,16].

The two images have the same size. Feature points were detected using the Sobel edge operator [8]. Let $N_1$ be a large integer, for example, $N_1 = 200$. In each image the locations of the top $N_1$ Sobel gradient magnitudes were recorded. Let $a_1(1), \ldots a_{N_1}(1)$ and $a_1(2), \ldots a_{N_1}(2)$ be the two lists of points. The simplified notation $a_i$ is used for $a_i(1)$ or $a_i(2)$ where appropriate. Each point $a_i$ is the centre of a $7 \times 7$ block $w_i$ of grey levels. The points $a_i(1)$ and $a_j(2)$ are matched if

$$\|w_i(1) - w_j(2)\| = \min_{1 \leq s,t \leq N_1} \|w_s(1) - w_t(2)\|.$$

The matching points $a_i(1)$, $a_j(2)$ are removed from the two lists and the process is repeated until the two lists are empty. Let $c$ be the centre of either image and let the shortest side of an image have length $2l$. Define translated and scaled image points $q_i$ by $q_i = (a_i - c)/l$. All pairs $q_i(1) \leftrightarrow q_j(2)$ such that $\|q_i(1)\| \geq 1$ or $\|q_j(2)\| \geq 1$ are discarded, leaving $N$ image correspondences.
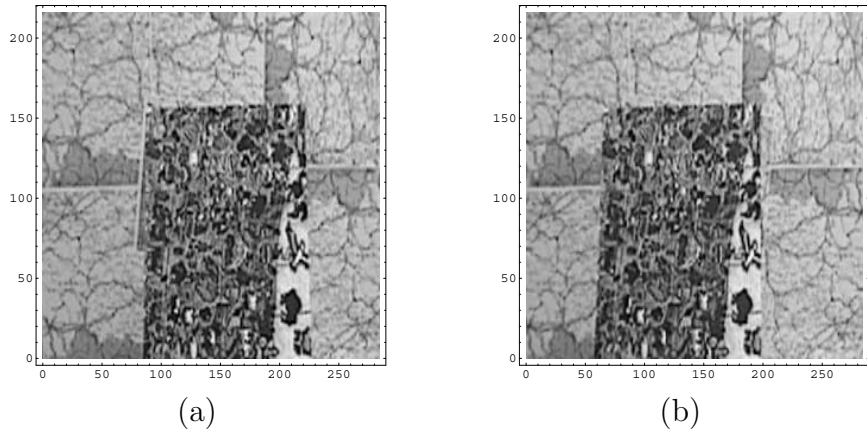
14

Figure 7: Images from the "Map" data set at http://www.middlebury.edu/stereo.
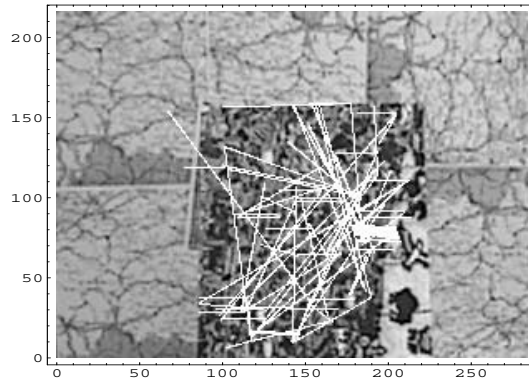


Figure 8: Image correspondences superposed on Fig. 7(a).

In the experiment, $N_1 = 200$ and $N = 95$. The image correspondences are shown in Fig. 7. Each correspondence $a_i(1) \leftrightarrow a_j(2)$ defines the end points of one of the white line segments, which are superposed on the image in Fig. 7(a).

The value of $\sigma$ was chosen to be $\sigma = 10^{-2}$ to ensure that the image dimensions are of order $2\sigma^{-1}$. The set $G$ was of size $|G| = 5201$. In comparison, the estimate (23) of the size of $G$ is 3316. The lower bound $M(N)$ on the number of inliers required for successful detection of a focus of expansion was estimated from (32) and (34), but with the term $\pi^{-1}V(K, T)$ in (32) replaced by $|G|$. The result was $M(N) = 25.37$.

The element $c_m$ of $G$ with the most inliers had polar coordinates $(r, \theta) = (5.01, 0.15\,\mathrm{rad})$. This focus is close to the horizontal axis and far from the origin, as expected given the displacement shown in Fig. 7. It had 43 inliers. There were in total 606 elements of $G$ with 26 or more inliers. Of these, 153 were contained in $D$. The maximum number of inliers for those elements of $G$ contained in $D$ was 30.

# 8  Conclusion

An accurate approximation to the Fisher-Rao metric for the focus of expansion has been obtained for the first time. The value of the approximating metric can be calculated

15

numerically for any given position of the focus of expansion, and the metric can be approximated by Taylor series in the polar coordinate $r$ when the focus of expansion is near to the origin or when it is far from the origin. The parameter space $T$ for the focus of expansion is sampled at a finite set $G$ of points, chosen such that each point of $T$ is near to at least one point of $G$. The set $G$ is the basis of an algorithm for estimating the focus of expansion, given a set of correspondences between two images taken by a translating camera.

Formulae were given for the probability of false detection of a focus of expansion and the probability of false rejection of a focus of expansion, under the assumption that image correspondences not associated with a focus of expansion, i.e. outliers, are distributed uniformly. Experimental results suggested that the algorithm successfully locates the focus of expansion even when the set of image correspondences contains a large number of outliers.

The analysis of the probability of false detection and the probability of false rejection can be improved in at least three ways. Firstly, it is necessary to estimate the change in the probability false detection on taking into account the fact that outlying points might support the detection of the true focus of expansion. Secondly, many algorithms for finding corresponding points in two images do not search the whole space $D^2$. Once a point $q_1$ is located in the first image, the search for a matching point is confined to a relatively small region of the second image. This constraint on the search space affects both the probability of a false detection and the probability of finding a correct correspondence $q_1 \leftrightarrow q_2$. Thirdly, it might be possible to improve the bound on the probability of a false detection by making more realistic assumptions about the distribution of outliers.

# References

1. M. Abramowitz and I.A. Stegun (eds.), Handbook of Mathematical Functions with formulas, graphs, and mathematical tables. (Dover, 1965).

2. J.K. Aggarwal and N. Nandhakumar, On the computation of motion from sequences of images - a review Proceedings of the IEEE. 76(1988)917-935.

3. S.-I. Amari, Differential-Geometrical Methods in Statistics, LNCS, vol. 28. (Springer, 1985).

4. T.M. Cover and J.A. Thomas, Elements of Information Theory. (John Wiley and Sons, 1991).

5. O. Faugeras, Three Dimensional Computer Vision. (MIT Press, 1993).

6. D.A. Forsyth and J. Ponce, Computer Vision, a Modern Approach. (Prentice Hall, 2003).

7. S. Gallot, D. Hulin and J. LaFontaine, Riemannian Geometry, 2nd edition, Universitext. (Springer, 1990).

8. R.C. Gonzalez and R.E. Woods, Digital Image Processing. 2nd Edition. (Prentice Hall, 2002).

9. S.J. Maybank, Fisher information and model selection for projective transformations of the line. Proc. Royal Society London, Series A, 459(2003)1-21.

10. S.J. Maybank, Detection of image structures using the Fisher information and the Rao metric. IEEE Transactions on Pattern Analysis and Machine Intelligence. 26(2004)1579-1589.

11. S.J. Maybank, The Fisher-Rao metric for projective transformations of the line. International Journal of Computer Vision. 63(2005)191-206.

12. S.J. Maybank, Application of the Fisher-Rao metric to structure detection. Journal of Mathematical Imaging and Vision. 25(2006)49-62.

13. S.J. Maybank, Application of the Fisher-Rao metric to ellipse detection. Int. J. of Computer Vision. 72(2007)287-307.

14. S.J. Maybank, The Fisher-Rao metric for lines in a convex image. Accepted by Int. J. Pattern Recognition and Artificial Intell. (2007).

15. D. Scharstein and R. Szeliski, A taxonomy and evaluation of dense two-frame stereo correspondence algorithms. Int. Journal of Computer Vision. 47(2002)7-42.

16. R. Szeliski and R. Zabih, An experimental comparison of stereo algorithms. Int. Workshop on Vision Algorithms, Kerkyra, Greece. (1999)1-19.

17. W. Triggs, Matching constraints and the joint image. Proc. 5th Int. Conf. on Computer Vision. (1995)338-343.

18. S. Wolfram, The Mathematica Book, 4th Edition. (Cambridge University Press, 1999).