Journal of Mathematical Imaging and Vision manuscript No.
(will be inserted by the editor)

A Fisher-Rao Metric for Curves Using the Information in Edges

Stephen J. Maybank

Provisional: 1 August 2014

Abstract Two curves which are close together in an image are indistinguishable given a measurement, in that there is no compelling reason to associate the measurement with one curve rather than the other. This observation is made quantitative using the parametric version of the Fisher-Rao metric. A probability density function for a measurement conditional on a curve is constructed. The distance between two curves is then defined to be the Fisher-Rao distance between the two conditional pdfs. A tractable approximation to the Fisher-Rao metric is obtained for the case in which the measurements are compound in that they consist of a point $\mathbf{x}$ and an angle α which specifies the direction of an edge at $\mathbf{x}$. If the curves are circles or straight lines, then the approximating metric is generalized to take account of inlying and outlying measurements. An estimate is made of the number of measurements required for the accurate location of a circle in the presence of outliers. A Bayesian algorithm for circle detection is defined. The prior density for the algorithm is obtained from the Fisher-Rao metric. The algorithm is tested on images from the CASIA Iris Interval database.

Keywords Bayesian curve detection, CASIA Iris Database, circle detection, Hough transform, Riemannian metric, step edges

1 Introduction

Curve detection is an important task in image processing and computer vision. In many applications of curve detection the curves belong to a family parameterised by the points of a manifold, such that each point in the manifold specifies a unique curve in the image. A family of this type contains an infinite number of distinct curves. However, due to the finite resolution of the image, two distinct curves which are sufficiently close to one another are in practice indistinguishable, in that there is no compelling reason to choose one curve in preference to the other. The notion of "sufficiently close" is made quantitative by defining a particular Riemannian metric on the parameter space for the curves. If the distance between two curves, as measured by this Riemannian metric, is less than a given threshold, then the two curves are regarded as indistinguishable, in the sense that a measurement inlying to one curve can equally well be considered as an inlier to the other curve.

The Riemannian metric on the parameter space is known as the Fisher-Rao metric. It is defined using the probability density function (pdf) for a measurement conditional on a curve in the image. The distance between two curves is defined to be the distance between the two corresponding conditional pdfs, as specified by the Fisher-Rao metric (Amari 1985; Maybank 2004; Rao 1945). The square of this distance has a probabilistic interpretation: it is twice the average of the log likelihood ratio for the two curves. The number of distinguishable curves, given a measurement, is proportional to the volume of the parameter space under the Fisher-Rao metric. The Fisher-Rao metric defines a probability density function on the parameter space, in which the probability of a subset of the parameter space is proportional to the volume of the subset under the metric. This pdf is a suitable prior in Bayesian parameter estimation (Balasubramanian 1997; Jaynes 2003; Jeffreys 1998).

The contents of this paper are summarised as follows.

i) A parameterised family of image curves is chosen and two Fisher-Rao metrics are defined on the parameter space. In the case of the first metric it is assumed that the measurements consist of image points $\mathbf{x}$. In the case of the second metric it is assumed that the measurements con-

S.J. Maybank
Department of Computer Science and Information Systems, Birkbeck College, Malet Street, London, WC1E 7HX, UK. E-mail: sjmaybank@dcs.bbk.ac.uk

sist of pairs $(\mathbf{x}, \alpha)$, such that $\mathbf{x}$ is an image point and α is the direction of a step edge centred at $\mathbf{x}$. These pairs are referred to as compound measurements.

ii) The probability density function for a measurement $(\mathbf{x}, \alpha)$ is constructed by assuming in effect that $\mathbf{x}$ is uniformly distributed along the curve and Gaussian distributed normal to the curve. The component α of the measurement is assumed to have a von Mises distribution centred at the direction of the tangent line to the curve at the point of the curve nearest to $\mathbf{x}$. The two Fisher-Rao metrics in (i) are approximated under the assumption that the standard deviation σ of the Gaussian component of $\mathbf{x}$ is small and the von Mises distribution approximates to a Gaussian distribution with a small standard deviation τ. The Gaussian densities are maximum entropy choices given σ, τ (Jaynes 2003). The choice of a uniform density for $\mathbf{x}$ along the curve is appropriate if there is no prior knowledge of the position of $\mathbf{x}$ along the curve. If the two approximating metrics differ significantly, then the component α of the measurement contains useful information which is not already contained in $\mathbf{x}$.

iii) Fisher-Rao metrics are defined for the case in which the measurement $\mathbf{x}$ or $(\mathbf{x}, \alpha)$, as appropriate, is an inlier with probability δ or an outlier with probability $1 - \delta$. It is shown that these metrics have tractable numerical approximations for the family of lines in an image and for the family of circles in an image.

iv) The metrics for circles obtained in (iii) are used to estimate the number of measurements required for the accurate detection of a circle in the presence of outliers. The inclusion of the component α in the measurement reduces significantly the required number of measurements if the probability δ of an inlier is small.

v) A Bayesian algorithm for curve detection is defined. The Fisher-Rao metric in (iii) based on compound measurements $(\mathbf{x}, \alpha)$ is used to define a prior probability density on the parameter space. The advantage of this algorithm is that it is not necessary to specify in advance which measurements are inliers and which are outliers. A version of this algorithm, adapted to circle detection, was tested on the CASIA Iris Interval database, with good results.

The new results are

- the definition of Fisher-Rao metrics firstly for compound measurements and secondly for compound measurements in the presence of outliers;
- an approximation to the Fisher-Rao metric for compound measurements when the noise level is low;
- numerically tractable approximations to the Fisher-Rao metrics for circles and straight lines in the presence of outlying measurements;
- a method for estimating the number of measurements necessary for the accurate detection of a circle in the presence of outlying measurements;
- a Bayesian algorithm for curve detection, in which it is not necessary to know a priori which measurements are inliers and which are outliers.

Related work is described in Sect. 2. A model for step edges and a probability density function for edge directions are defined in Sect. 3. The approximations to the Fisher-Rao metrics are obtained in Sect. 4. The approximating metric in the presence of outliers is obtained in Sect. 5.1 for circles. The number of measurements required for the accurate detection of a circle is estimated in Sect. 5.2. The approximating metric in the presence of outliers is obtained in Sect. 5.3 for straight lines. The Bayesian algorithm for curve detection is described in Sect. 6.1, and experimental results are reported in Sect. 6.2. A comparison with two Hough based algorithms for circle detection is made in Sect. 6.3. Some concluding remarks are made in Sect. 7.

2 Related Work

Three areas of related work are described, namely the Fisher-Rao metric (Sect. 2.1), metrics on parameter spaces for curves (Sect. 2.2) and methods for curve detection (Sect. 2.3).

2.1 Fisher-Rao metric

The Fisher-Rao metric is described by Amari (1985) and Rao (1945). Amari notes that it is the only known metric for conditional pdfs which is invariant under reparameterisations of the data and of the conditioning parameters. Cover and Thomas (1991) describe the one dimensional version of the Fisher-Rao metric which is referred to as the Fisher information. Fisher-Rao metrics on the space of lines in the plane and the space of ellipses in the plane are described in Maybank (2004) and Maybank (2007) respectively. In both cases the measurements are points in the image. It is shown that the Fisher-Rao metrics for lines and ellipses can be approximated by simpler closed form metrics provided the errors in the measurements are small compared with the scale of the curves. An approximation to the Fisher-Rao metric for projections of lines into a paracatadioptric image is obtained by Maybank et al. (2012). A general approximation to the Fisher-Rao metric is obtained by Kanatani (1996).

Srivastava et al. (2007) describe several applications of the Fisher-Rao metric to computer vision, including the comparison of pdfs for the values obtained by filtering images, the classification of planar shapes and the measurement of distances between time warping functions used for activity

analysis. The Fisher-Rao metric is preferred over competing metrics because the values of distances and volumes obtained using the metric do not depend on the choice of parameterization of the space on which the metric is defined. Peter and Rangarajan (2009) define a Fisher-Rao metric for planar shapes that are specified by a finite number of landmark points, however in their experiments they replace the Fisher-Rao metric by a new metric which has a closed form definition. The deformations of shapes are investigated using the new metric. Ceolin and Hancock (2010) use a Fisher-Rao metric to measure distances between 3D scans of face images. The resulting edge weighted graph is simplified by projecting it into a low dimensional space. The projected graph is used to classify facial expressions. Ceolin and Hancock (2012) use a similar method to classify faces according to gender.

2.2 Metrics on parameter spaces for curves

The space of all curves obtained by mapping the unit circle homeomorphically to the plane is infinite dimensional. Many metrics for this space have been suggested, with applications to shape description, tracking and interpolating between two given shapes. Sundaramoorthi et al. (2011) describe a metric in which the effects of centroid translations, scale changes and deformations are orthogonal to each other. They apply the metric to the tracking of the contours in images of a beating heart. Mio et al. (2006) define a metric on the space of all curves in the plane using concepts from elasticity, and then use geodesics in the space of curves to interpolate between two given planar shapes. A mathematical overview of Riemannian metrics on spaces of curves and the applications of these methods to shape description is given by Michor and Mumford (2007).

Srivastava et al. (2011) define a number of metrics on the space of curves in $\mathbb{R}^n$. Shape spaces are defined for general curves and for closed curves by identifying any two curves that differ by a reparameterisation, a rigid motion or a scale change. Algorithms are described for computing geodesics between points in the shape spaces. Applications to the classification of curves are described.

Any metric on the space of all curves induces a metric on any finite dimensional submanifold of the space of all curves. However, there are relatively few computer vision papers that explicitly describe Riemannian metrics on finite dimensional parameter manifolds for curves. Tatu et al. (2010) describe a Riemannian metric which is induced on a finite dimensional family of closed B-spline curves in which the number of nodes is fixed and the nodes are equally spaced along each curve. These B-spline curves are used for modeling shapes in images.

2.3 Curve detection

The Hough transform is the basis for many curve detection methods in which the parameter space is partitioned into disjoint subsets. Each subset is assigned a value which is equal to the number of measurements exactly compatible with one or more curves specified by the points in the subset. The variable holding this value is referred to as an accumulator. Curve detection is based on the subset with the highest accumulator value. Introductory material on the Hough transform can be found in the books by Forsyth and Ponce (2011), Gonzalez and Woods (2008) and Szeliski (2011). A survey of the literature on the Hough transform is given by Leavers (1993). Kimme et al. (1975) describe a Hough transform based algorithm for circle detection in which the local edge orientations are used to simplify the updating of the Hough accumulators associated with the different subsets of the parameter space. Olson (1999) describes a constrained Hough transform for curve detection. The transform is defined using subsets of the measurements, rather than single measurements. Aguado et al. (1996) describe a two stage Hough transform for detecting circles. In the first stage, candidate centres of circles are identified. In the second stage the corresponding radii are found. The two stage method is then extended to detect ellipses. Ballard (1981) uses table look up to define a Hough transform method for detecting arbitrary shapes. Woodford et al. (2014) use a minimum entropy criterion to select Hough accumulators which receive contributions from many measurements. Code for a Hough transform based algorithm for detecting circles can be found in OpenCV (2014). The MATLAB function imfindcircles also uses a Hough transform to detect circles. See 'More About' on the MathWorks page for imfindcircles[1].

In Bayesian algorithms for curve detection each subset of the parameter space is assigned a value equal to the probability that the subset contains a parameter vector for a curve that is a good fit to the measurements. The advantage of Bayesian algorithms is that they make the best possible use of the available data (Jaynes 2003). Bonci et al. (2005) describe a Hough transform method for curve detection which is Bayesian provided all the measurements are known to be inliers. Bonci et al's method is similar to one due to Ji and Haralick (1999), except that Bonci et al. multiply probabilities in order to obtain the values of the Hough accumulators, while Ji and Haralick add probabilities. The advantage of adding probabilities is that it is not necessary to exclude measurements which are outliers. Each outlying measurement far from the curve makes only a very small contribution to the sum. The disadvantage of adding probabilities in this way is that the method is no longer Bayesian. Toronto et al. (2007) describe a Bayesian foundation for the Hough transform, however their pdf for a measurement con-

[1] http://uk.mathworks.com/help/images/ref/imfindcircles.html

ditional on a given curve is in effect constant on the space of measurements, once the Lebesgue measure is taken into account. Werman and Keren (2001) describe a general method for curve detection which is Bayesian, provided all the measurements are known to be inliers. Their method can be applied even if the errors in the measurements of an image point are relatively large.

3 Step Edges

There are many different types of edges in natural images, for example step edges, ramps and roof edges (Gonzalez and Woods 2008). In this paper the edges are chosen to be step edges. This choice is appropriate for applications in which the curve to be detected is the boundary of an object seen against a contrasting background. The model for step edges is described in Sect. 3.1 and a probability density function for the measurements of the direction of a step edge is described in Sect, 3.2.

3.1 Step edges

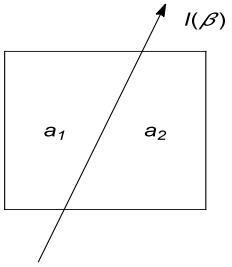

Fig. 1 A pixel cut by a line $l(\beta)$ into two parts, with areas a_1 and a_2 respectively.

Let $\mathbb{Z}^2$ be the integer lattice in the plane $\mathbb{R}^2$. Let I be an image such that each pixel in I corresponds to a unit square centred at a point of $\mathbb{Z}^2$. Let β be an angle such that $0 \leq \beta < 2\pi$ and let $l(\beta)$ be the oriented line through the origin $(0,0)^\top$ with direction

$$\mathbf{v} = (\cos(\beta), \sin(\beta))^\top. \tag{1}$$

Each pixel in I is given a value equal to the area of that part of the pixel to the left of the line $l(\beta)$, as shown in Fig. 1.

Let m be odd positive integer and let $\mathbf{R}(\beta)$ be the $m \times m$ matrix of pixel values obtained from the $m \times m$ image region centred at $(0,0)^\top$ in I. Let $\mathbf{M}$ be an $m \times m$ matrix of pixel values, let μ be the average value of the elements of $\mathbf{M}$, let $\mathbf{e}$ be the vector of dimension m^2 in which every entry is equal to 1 and let $\mathbf{g}(\mathbf{M})$ be the m^2 dimensional vector obtained by flattening $\mathbf{M}$. If $\mathbf{g}(\mathbf{M}) \neq \mu\mathbf{e}$, then the function $\zeta(\mathbf{M})$ is defined by

$$\zeta(\mathbf{M}) = \|\mathbf{g}(\mathbf{M}) - \mu\mathbf{e}\|^{-1}(\mathbf{g}(\mathbf{M}) - \mu\mathbf{e}),$$

where $\|.\|$ is the Euclidean norm.

An image region with an $m \times m$ matrix $\mathbf{M}$ of pixel values is said to contain an edge if $\zeta(\mathbf{M})$ is near to $\zeta(\mathbf{R}(\beta))$ for some value of β. In detail, let α be defined by

$$\alpha = \text{argmax}_{0 \leq \beta < 2\pi}\, \beta \mapsto \zeta(\mathbf{R}(\beta))^\top \zeta(\mathbf{M}). \tag{2}$$

Let b, ψ be non-negative thresholds. The image region from which $\mathbf{M}$ is obtained contains an edge with measured direction α if the standard deviation of the pixel values in $\mathbf{M}$ is greater than or equal to b and

$$\cos^{-1}(\zeta(\mathbf{R}(\alpha))^\top \zeta(\mathbf{M})) \leq \psi.$$

These criteria are used for step edge detection in the experiments reported in Sect. 6.2 and Sect. 6.3.

3.2 PDF for the direction of a step edge

Let H_β be the hypothesis that a given image region contains a step edge subject to noise and with an underlying true direction β. Let α be the measured direction. The conditional probability density function $p(\alpha|H_\beta)$ is assumed to be a von Mises density of the form

$$p(\alpha|H_\beta) = (2\pi I_0(\kappa))^{-1} \exp(\kappa\cos(\alpha - \beta)), 0 \leq \alpha < 2\pi, \tag{3}$$

where κ is a strictly positive constant and I_0 is a modified Bessel function of the first kind (Mardia and Jupp 2000). If κ is large then the von Mises density is closely approximated by a Gaussian density with expected value β and variance κ^{-1}. The advantage of using the von Mises density is that it automatically takes into account the fact that α and β are angles defined on $[0, 2\pi)$.

4 Fisher-Rao Metrics

Let $\mathbf{J}$ be the Fisher-Rao metric for curves based on the information in compound measurements of the form $(\mathbf{x}, \alpha)$ where $\mathbf{x}$ is a pixel in the image and α is a measurement of the direction of a step edge in an $m \times m$ image region centred at $\mathbf{x}$, as described in Sect. 3.1. It is shown that $\mathbf{J} = \mathbf{J}' + \mathbf{J}''$, where $\mathbf{J}'$ is the Fisher-Rao metric for curves based on the information in the measurements $\mathbf{x}$, and $\mathbf{J}''$ is obtained from the measurements α. Approximations $\mathbf{K}'$, $\mathbf{K}''$ to $\mathbf{J}'$, $\mathbf{J}''$ are obtained, such that the metric $\mathbf{K}$ defined by $\mathbf{K} = \mathbf{K}' + \mathbf{K}''$ is an approximation to $\mathbf{J}$.

A conditional pdf for the measurement $\mathbf{x}$ is described in Sect. 4.1. The Fisher-Rao metrics $\mathbf{J}$, $\mathbf{J}'$ and the approximation $\mathbf{K}'$ to $\mathbf{J}'$ are obtained in Sect. 4.2. The approximation $\mathbf{K}''$ to $\mathbf{J}''$ is obtained in Sect. 4.3.

4.1 Probability density for the measurement points

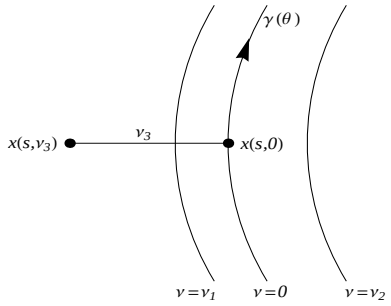

Fig. 2 Coordinates (s,ν) for points $\mathbf{x}(s,\nu)$ near to or on the trace $\gamma(\theta)$ of the curve θ. The arrow on $\gamma(\theta)$ shows the direction in which the arc length s increases. The curves defined by $\nu=\nu_1$ and $\nu=\nu_2$, with $\nu_1<0<\nu_2$, are shown. A point $\mathbf{x}(s,\nu)$ is on $\gamma(\theta)$ if and only if $\nu=0$.

Let T be a parameter manifold for a family of curves in an image I, let $\|.\|$ be the Euclidean metric in I and let $\gamma(\theta)$ be the trace of the curve specified by the point in T with coordinates θ in a suitable parameterisation of T. By definition, $\gamma(\theta)$ is the subset of points in I that are on the curve specified by θ. It is assumed that $\gamma(\theta)$ has a tangent line at each of its points. Let s be an arc length parameter on $\gamma(\theta)$. The arc length parameter specifies an orientation of $\gamma(\theta)$ in which the positive direction is the direction in which the arc length increases. Coordinates (s,ν) are chosen in a neighbourhood of $\gamma(\theta)$ such that ν is the signed distance from the image point $\mathbf{x}(s,\nu)$ to $\gamma(\theta)$,

$$\nu(\mathbf{x},\theta)=\pm\min_{\mathbf{y}\in\gamma(\theta)}\|\mathbf{x}(s,\nu)-\mathbf{y}\|, \tag{4}$$

where the minus sign is chosen if $\mathbf{x}(s,\nu)$ is on the left hand side of $\gamma(\theta)$ and the plus sign is chosen if $\mathbf{x}(s,\nu)$ is on the right hand side of $\gamma(\theta)$, as illustrated in Fig. 2. The point on $\gamma(\theta)$ nearest to $\mathbf{x}(s,\nu)$ is $\mathbf{x}(s,0)$.

Let $L(\theta)$ be the length of $\gamma(\theta)$ and let σ be the standard deviation of the error in the displacement of the measurement $\mathbf{x}$ normal to $\gamma(\theta)$. The pdf $p(\mathbf{x}|\theta)$ is defined as in Maybank (2004) by

$$\begin{aligned}&p(\mathbf{x}|\theta)d\mathbf{x} \;=\; p(\mathbf{x}(s,\nu)|\theta)\,ds d\nu \;=\\ &L(\theta)^{-1}(2\pi\sigma^2)^{-1/2}\exp\left(-\nu^2/(2\sigma^2)\right)ds d\nu,\;\; \mathbf{x}\in I.\end{aligned} \tag{5}$$

Intuitively, $p(\mathbf{x}|\theta)$ is uniform along the trace $\gamma(\theta)$ and Gaussian with standard deviation σ perpendicular to $\gamma(\theta)$.

4.2 The Fisher-Rao metric

Let $d(T)$ be the dimension of the parameter manifold T for the family of curves in the image I. The Fisher-Rao metric $\mathbf{J}$ on T, for the compound measurements $(\mathbf{x},\alpha)$, is

$$\begin{aligned}J_{ij}(\theta)=-\int_I\int_0^{2\pi}\left(\frac{\partial^2}{\partial\theta_i\partial\theta_j}\ln(p(\mathbf{x},\alpha|\theta))\right)\times\\ p(\mathbf{x},\alpha|\theta)\,d\alpha d\mathbf{x},\;\; 1\le i,j\le d(T).\end{aligned} \tag{6}$$

The conditional probability $p(\mathbf{x},\alpha|\theta)$ is written in the form

$$\begin{aligned}p(\mathbf{x},\alpha|\theta)=p(\alpha|\mathbf{x},\theta)p(\mathbf{x}|\theta),\\ \mathbf{x}\in I,\alpha\in[0,2\pi),\theta\in T.\end{aligned} \tag{7}$$

It follows from (6) and (7) that $\mathbf{J}(\theta)$ is the sum of terms $\mathbf{J}'(\theta)$ and $\mathbf{J}''(\theta)$, that are defined by

$$\begin{aligned}&J'_{ij}(\theta)\;=\\ &-\int_I\int_0^{2\pi}\left(\frac{\partial^2}{\partial\theta_i\partial\theta_j}\ln(p(\mathbf{x}|\theta))\right)p(\alpha|\mathbf{x},\theta)p(\mathbf{x}|\theta)\,d\alpha d\mathbf{x},\\ &=-\int_I\left(\frac{\partial^2}{\partial\theta_i\partial\theta_j}\ln(p(\mathbf{x}|\theta))\right)p(\mathbf{x}|\theta)d\mathbf{x},1\le i,j\le d(T),\end{aligned} \tag{8}$$

$$\begin{aligned}&J''_{ij}(\theta)=\\ &-\int_I\int_0^{2\pi}\left(\frac{\partial^2}{\partial\theta_i\partial\theta_j}\ln(p(\alpha|\mathbf{x},\theta))\right)p(\alpha|\mathbf{x},\theta)p(\mathbf{x}|\theta)\,d\alpha d\mathbf{x},\\ &1\le i,j\le d(T).\end{aligned} \tag{9}$$

It is apparent from (8) that $\mathbf{J}'$ is the Fisher-Rao metric on T based on the information contained in the measurements $\mathbf{x}$. It is shown in Maybank (2004) that if the pdf $p(\mathbf{x}|\theta)$ is given by (5) and if σ is small, then $\mathbf{J}'$ is closely approximated by the Riemannian metric $\mathbf{K}'$ defined by

$$\begin{aligned}K'_{ij}(\theta)=\frac{1}{2\sigma^2L(\theta)}\int_0^{L(\theta)}\left(\frac{\partial^2}{\partial\theta_i\partial\theta_j}\nu(\mathbf{x},\theta)^2\right)_{\mathbf{x}=\mathbf{x}(s,0)}ds,\\ 1\le i,j\le d(T),\end{aligned} \tag{10}$$

where s is an arc length parameter on $\gamma(\theta)$, chosen such that $0\le s\le L(\theta)$ and $\nu(\mathbf{x},\theta)$ is given by (4). See also Kanatani (1996, Section 14.4). It is noted that $\mathbf{K}'(\theta)$ has the equivalent form

$$\begin{aligned}&K'_{ij}(\theta)=\frac{1}{\sigma^2L(\theta)}\times\\ &\int_0^{L(\theta)}\left(\left(\frac{\partial}{\partial\theta_i}\nu(\mathbf{x},\theta)\right)\left(\frac{\partial}{\partial\theta_j}\nu(\mathbf{x},\theta)\right)\right)_{\mathbf{x}=\mathbf{x}(s,0)}ds,\\ &1\le i,j\le d(T).\end{aligned} \tag{11}$$

4.3 Contribution of the directions to the Fisher-Rao metric

Let $\tilde{\mathbf{J}}''(\mathbf{x},\theta)$ be the matrix defined by

$$\begin{aligned}\tilde{J}''_{ij}(\mathbf{x},\theta)=-\int_0^{2\pi}\left(\frac{\partial^2}{\partial\theta_i\partial\theta_j}\ln p(\alpha|\mathbf{x},\theta)\right)p(\alpha|\mathbf{x},\theta)\,d\alpha,\\ 1\le i,j\le d(T).\end{aligned} \tag{12}$$

It is assumed that $p(\alpha|\mathbf{x},\theta)\equiv p(\alpha|H_\beta)$, where $p(\alpha|H_\beta)$ is given by (3). The angle β specifies the direction of the tangent to $\gamma(\theta)$ at the nearest point of $\gamma(\theta)$ to $\mathbf{x}$, as illustrated

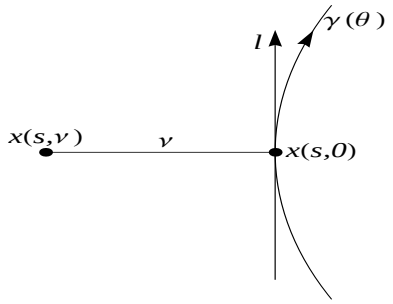

Fig. 3 The line l is tangent to $\gamma(\theta)$ at the point $\mathbf{x}(s,0)$ of $\gamma(\theta)$ nearest to $\mathbf{x}(s,v)$. The direction of l is $\beta(\mathbf{x}(s,0),\theta)$.

in Fig. 3. On substituting $p(\alpha|H_\beta)$ for $p(\alpha|\mathbf{x},\theta)$ in (12), it follows that

$$J''_{ij}(\mathbf{x},\theta) = U(\beta)\frac{\partial\beta}{\partial\theta_i}\frac{\partial\beta}{\partial\theta_j},\ \ 1\le i,j\le d(T), \tag{13}$$

where

$$\begin{aligned} U(\beta) &= \int_0^{2\pi} p(\alpha|H_\beta)^{-1}\left(\frac{\partial}{\partial\beta}p(\alpha|H_\beta)\right)^2 d\alpha. \\ &= \kappa I_1(\kappa)/I_0(\kappa). \end{aligned} \tag{14}$$

The terms $I_0(\kappa)$, $I_1(\kappa)$ are modified Bessel functions of the first kind. The argument β is omitted from $U(\beta)$ in the rest of this section, because $U(\beta)$ is independent of β.

It follows from (9), (12) and (13) that

$$J''_{ij}(\theta) = U\int_I \frac{\partial\beta}{\partial\theta_i}\frac{\partial\beta}{\partial\theta_j}p(\mathbf{x}|\theta)\,d\mathbf{x},\ \ 1\le i,j\le d(T).$$

If the image noise level σ in the definition (5) of $p(x|\theta)$ is small then $\mathbf{J}''(\theta)$ is closely approximated by the matrix

$$L(\theta)^{-1}U\int_0^{L(\theta)}\frac{\partial\beta}{\partial\theta_i}\frac{\partial\beta}{\partial\theta_j}\,ds,\ \ 1\le i,j\le d(T), \tag{15}$$

where s is an arc length parameter on the trace $\gamma(\theta)$ and β is a function of s. The expression (15) is a product of two terms. The first term, U, depends only on the probabilistic model for edges. The second term is purely geometric, in that it depends only on the parameterised family of curves. If κ is large then the von Mises pdf (3) is closely approximated by a Gaussian pdf with standard deviation $\tau=\kappa^{-1/2}$. In this case U is closely approximated by τ^{-2}.

To summarise, the Fisher-Rao metric $\mathbf{J}$ is closely approximated by the metric $\mathbf{K}$ defined by

$$K_{ij}(\theta) = K'_{ij}(\theta) + K''_{ij}(\theta),$$

$$= \frac{1}{L(\theta)}\int_0^{L(\theta)}\left(\frac{1}{\sigma^2}\frac{\partial v}{\partial\theta_i}\frac{\partial v}{\partial\theta_j}+\frac{1}{\tau^2}\frac{\partial\beta}{\partial\theta_i}\frac{\partial\beta}{\partial\theta_j}\right)_{\mathbf{x}=\mathbf{x}(s,0)} ds, \quad 1\le i,j\le d(T), \tag{16}$$

where $\mathbf{K}''(\theta)$ is obtained on replacing U by τ^{-2} in (15).

5 Inliers and Outliers

In this section and in later sections it is convenient to assume that the image I has the form of the unit disk D, centred at the origin. The symmetry of D ensures that certain integrations can be carried out exactly.

Let δ be the prior probability that a measurement $(\mathbf{x},\alpha)$ is an inlier to $p(\mathbf{x},\alpha|\theta)$. As in earlier sections, $\mathbf{x}$ is a point in the image and α is the direction of a step edge centred at $\mathbf{x}$. It is assumed that the outliers to $p(\mathbf{x},\alpha|\theta)$ are uniformly distributed on $D\times[0,2\pi)$. The pdf $q(\mathbf{x},\alpha|\theta)$ for $(\mathbf{x},\alpha)$ taking the outliers into account is given by

$$q(\mathbf{x},\alpha|\theta) = \delta p(\mathbf{x},\alpha|\theta) + (1-\delta)/(2\pi^2), \quad (\mathbf{x},\alpha)\in D\times[0,2\pi),\ \ \theta\in T. \tag{17}$$

Let $\mathbf{J}(q,\theta)$ be the Fisher-Rao metric for the family of pdfs q. In this section tractable approximations $\mathbf{K}(q,\theta)$ to $\mathbf{J}(q,\theta)$ are obtained for the family of circles in D and the family of lines in D. The metric $\mathbf{K}(q,\theta)$ is the basis for a circle detection algorithm described in Sect. 6.

The metric $\mathbf{K}(q,\theta)$ for circles is obtained in Sect. 5.1. In Sect. 5.2 this metric is used to assess the effect of outliers on the accuracy of circle detection. The metric $\mathbf{K}(q,\theta)$ for lines is obtained in Sect. 5.3.

5.1 Circles

Let $\theta=(\xi,c_1,c_2)^\top$ be the parameter vector for a circle in the unit disk D. The circle has radius ξ and centre $\mathbf{c}=(c_1,c_2)^\top$. The entire circle is required to be in D, thus

$$\xi + \|(c_1,c_2)^\top\| \le 1.$$

Let q be the pdf (17) for the family of circles in D. It follows from the definition of the Fisher-Rao metric that

$$\begin{aligned} J_{ij}(q,\theta) &= \int_0^{2\pi}\int_D q^{-1}\frac{\partial q}{\partial\theta_i}\frac{\partial q}{\partial\theta_j}\,d\mathbf{x}d\alpha, \\ &= \delta^2\int_0^{2\pi}\int_D q^{-1}p^2\left(\frac{\partial}{\partial\theta_i}\ln p\right)\left(\frac{\partial}{\partial\theta_j}\ln p\right)d\mathbf{x}d\alpha, \\ &\qquad 1\le i,j\le 3. \end{aligned} \tag{18}$$

The arguments $\mathbf{x}$, α, θ are omitted from $p(\mathbf{x},\alpha|\theta)$ and $q(\mathbf{x},\alpha|\theta)$ in (18) to reduce clutter.

It follows from (3), (5) and (7) that the pdf $p(\mathbf{x},\alpha|\theta)$ is given by

$$\begin{aligned} p(\mathbf{x},\alpha|\theta) \ =\ & \left((2\pi)^{3/2}\sigma I_0(\kappa)L(\theta)\right)^{-1}\times \\ & \exp\left(-\frac{1}{2\sigma^2}(\|\mathbf{x}-\mathbf{c}\|-\xi)^2+\kappa\cos(\alpha-\beta)\right),\ \ \theta\in T, \end{aligned} \tag{19}$$

where β is the direction of the tangent at the nearest point of the curve θ to $\mathbf{x}$. Let $\mathbf{x}=(x_1,x_2)^\top$. The angle β is given by

$$\tan(\beta) = -(x_1-c_1)/(x_2-c_2). \tag{20}$$

The coordinate transformation between $\mathbf{x}$ and (s,ν) is defined by

$$\mathbf{x} = \mathbf{c} + (\xi+\nu)(\cos(\xi^{-1}s), \sin(\xi^{-1}s)).$$

It follows from (19) that to leading order,

$$\ln p(\mathbf{x},\alpha|\theta) = -(2\sigma^2)^{-1}(\|\mathbf{x}-\mathbf{c}\|-\xi)^2 + \kappa\cos(\alpha-\beta),$$

thus to leading order

$$\begin{aligned}
\partial \ln p/\partial\xi &= \sigma^{-2}(\|\mathbf{x}-\mathbf{c}\|-\xi),\\
\partial \ln p/\partial c_1 &= \sigma^{-2}\|\mathbf{x}-\mathbf{c}\|^{-1}(\|\mathbf{x}-\mathbf{c}\|-\xi)(x_1-c_1) +\\
&\quad \kappa\sin(\alpha-\beta)\partial\beta/\partial c_1,\\
\partial \ln p/\partial c_2 &= \sigma^{-2}\|\mathbf{x}-\mathbf{c}\|^{-1}(\|\mathbf{x}-\mathbf{c}\|-\xi)(x_2-c_2) +\\
&\quad \kappa\sin(\alpha-\beta)\partial\beta/\partial c_2. \qquad (21)
\end{aligned}$$

It follows from (20) that the derivatives of β are given by

$$\begin{aligned}
\partial\beta/\partial\xi &= 0,\\
\partial\beta/\partial c_1 &= \|\mathbf{x}-\mathbf{c}\|^{-2}(x_2-c_2),\\
\partial\beta/\partial c_2 &= -\|\mathbf{x}-\mathbf{c}\|^{-2}(x_1-c_1). \qquad (22)
\end{aligned}$$

Let the function $f(\nu,\alpha)$ be defined by

$$f(\nu,\alpha) = q(\mathbf{x}(s,\nu),\alpha|\theta)^{-1}p(\mathbf{x}(s,\nu),\alpha|\theta)^2, \quad \theta\in T. \qquad (23)$$

It is noted that the right-hand side of (23) is independent of s, and that $f(\nu,\alpha)$ is an even function of ν. It follows from (18), (21), (22) and (23) that $J_{11}(q,\theta)$ is approximated by

$$\begin{aligned}
K_{11}(q,\theta) &= \sigma^{-4}\delta^2\int_0^{2\pi}\int_{\mathbb{R}}\int_0^{2\pi\xi} f(\nu,\alpha-\beta)\nu^2 ds d\nu d\alpha,\\
&= 2\pi\sigma^{-4}\delta^2\xi\int_0^{2\pi}\int_{\mathbb{R}} f(\nu,\alpha)\nu^2\, d\nu d\alpha. \qquad (24)
\end{aligned}$$

Similarly, $J_{22}(q,\theta)$ is approximated by

$$\begin{aligned}
K_{22}(q,\theta) &= \delta^2\int_0^{2\pi}\int_{\mathbb{R}}\int_0^{2\pi\xi} f(\nu,\alpha-\beta)\times\\
&\left(\sigma^{-2}\nu\cos(\xi^{-1}s) + \kappa\xi^{-1}\sin(\alpha-\beta)\sin(\xi^{-1}s)\right)^2 ds d\nu d\alpha,\\
&= \pi\delta^2\xi\int_0^{2\pi}\int_{\mathbb{R}} f(\nu,\alpha)\left(\sigma^{-4}\nu^2 + \kappa^2\xi^{-2}\sin^2(\alpha)\right) d\nu d\alpha. \qquad (25)
\end{aligned}$$

It follows from symmetry that $K_{33}(q,\theta) = K_{22}(q,\theta)$. Similar calculations show that $\mathbf{K}(q,\theta)$ is a diagonal matrix, which is function of ξ but independent of $\mathbf{c}$. It is straightforward to estimate the integrals in (24) and (25) numerically.

On setting $\delta = 1$ and approximating the von Mises density with a Gaussian density with standard deviation τ, the following expression for $\mathbf{K}(\theta)$ is obtained,

$$\mathbf{K}(\theta) = \frac{1}{2\sigma^2}\begin{pmatrix}2&0&0\\0&1&0\\0&0&1\end{pmatrix} + \frac{1}{2\tau^2\xi^2}\begin{pmatrix}0&0&0\\0&1&0\\0&0&1\end{pmatrix}. \qquad (26)$$

The expression (26) for $\mathbf{K}(\theta)$ can be obtained directly from (16), using (19).

5.2 Effect of outliers on circle detection

An estimate is made of the number of measurements required to locate a circle accurately in the presence of outliers. Let θ and $\theta+\Delta\theta$ be two nearby points in the parameter space T for circles. Let X be a set of n compound measurements sampled from the circle θ. The Kullback-Leibler divergence of $q(X|\theta+\Delta\theta)$ from $q(X|\theta)$ is equal to the expected value of the log likelihood ratio

$$\ln(q(X|\theta)/q(X|\theta+\Delta\theta))$$

where the expectation is taken over X. The Kullback-Leibler divergence of $q(X|\theta+\Delta\theta)$ from $q(X|\theta)$ is given to leading order by

$$\frac{1}{2}n\Delta\theta^\top\mathbf{J}(q,\theta)\Delta\theta + O(\|\Delta\theta\|^3).$$

It is assumed that θ can be reliably distinguished from $\theta+\Delta\theta$ if the expected value of the log likelihood ratio is at least $\ln(10)$. It follows that the minimum value of the number n of measurements required on average to distinguish θ from $\theta+\Delta\theta$ can be estimated using the equation

$$\ln(10) = \frac{1}{2}n\Delta\theta^\top\mathbf{K}(q,\theta)\Delta\theta. \qquad (27)$$

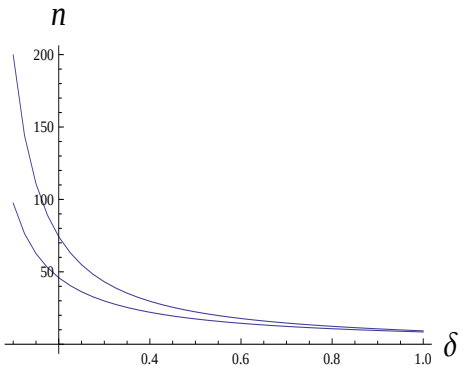

Fig. 4 Lower curve: number of compound measurements $(\mathbf{x},\alpha)$ required to detect a circle, as a function of δ. Upper curve: number of measurements $\mathbf{x}$ as a function of δ.

A numerical example of (27) is investigated. The parameters σ, τ are assigned the values $\sigma = 0.05$, $\tau = 0.1$. Let $\theta = (0.5,0,0)^\top$ and let the increment $\Delta\theta$ is chosen such that

$$\Delta\theta = (0,\Delta c_1,\Delta c_2)^\top \text{ and } \Delta c_1^2 + \Delta c_2^2 = \sigma^2. \qquad (28)$$

It follows from (27), (28) and the properties of the matrix $\mathbf{K}(q,\theta)$ that

$$n = 2\ln(10)/(\sigma^2 K_{22}(q,\theta)).$$

A graph of n as a function of the prior probability δ that a measurement is an inlier is shown as the lower curve in Fig.

4. The upper curve in Fig. 4 is the graph of n as a function of δ for measurements consisting of points $\mathbf{x}$ in the image. At high values of δ the two curves are similar. At low values of δ the two curves rise rapidly but the rate of increase of the curve for measurements $\mathbf{x}$ is much greater than the rate of increase of the curve for measurements $(\mathbf{x},\alpha)$.

5.3 Lines

The space T of lines which intersect the unit disk D is parameterised by the polar coordinates $(\xi,\gamma)^\top \equiv \theta$ of the nearest point on a line to the origin. The coordinates ξ, γ take values in the ranges

$$0 \le \xi \le 1 \text{ and } 0 \le \gamma < 2\pi.$$

Let

$$\mathbf{u} = (\cos(\gamma),\sin(\gamma))^\top.$$

It follows from (3), (5) and (7) that the pdf $p(\mathbf{x},\alpha|\theta)$ is given by

$$p(\mathbf{x},\alpha|\theta) = ((2\pi)^{3/2}\sigma I_0(\kappa)L(\theta))^{-1} \times \exp\left(-\frac{1}{2\sigma^2}(\mathbf{x}.\mathbf{u}-\xi)^2 + \kappa\cos(\alpha-\beta)\right). \quad (29)$$

The angle β is given by $\beta = \gamma + \pi/2$. The coordinate transformation between $\mathbf{x}$ and (s,ν) is defined by

$$\mathbf{x} = (\xi+\nu)\mathbf{u} + s(-\sin(\gamma),\cos(\gamma))^\top.$$

It follows from (29) that to leading order

$$\begin{aligned}\partial \ln p/\partial\xi &= \sigma^{-2}(\mathbf{x}.\mathbf{u}-\xi),\\ \partial \ln p/\partial\gamma &= \sigma^{-2}(\mathbf{x}.\mathbf{u}-\xi)(x_1\sin(\gamma)-x_2\cos(\gamma)) + \\ &\quad \kappa\sin(\alpha-\beta). \end{aligned} \quad (30)$$

Let $f(\nu,\alpha)$ be defined by

$$f(\nu,\alpha) = q(\mathbf{x}(s,\nu),\alpha|\theta)^{-1}p(\mathbf{x}(s,\nu),\alpha|\theta)^2, \quad \theta \in T. \quad (31)$$

It is noted that the right-hand side of (31) is independent of s and that $f(\nu,\alpha)$ is an even function of ν. It follows from (30) and (31) that $\mathbf{J}(q,\theta)$ is approximated by the matrix $\mathbf{K}(q,\theta)$ defined by

$$\begin{aligned} K_{11}(q,\theta) &= 2\sigma^{-4}\delta^2(1-\xi^2)^{1/2}\int_0^{2\pi}\int_{\mathbb{R}} f(\nu,\alpha)\nu^2 d\nu d\alpha,\\ K_{22}(q,\theta) &= 2\delta^2\int_0^{2\pi}\int_{\mathbb{R}} f(\nu,\alpha)\times\\ &\quad \left((3\sigma^4)^{-1}(1-\xi^2)^{3/2}\nu^2 + (1-\xi^2)^{1/2}\kappa^2\sin^2(\alpha)\right)d\nu d\alpha,\\ K_{12}(q,\theta) &= K_{21}(q,\theta) = 0. \end{aligned} \quad (32)$$

It is straightforward to estimate the integrals in (32) numerically.

On setting $\delta = 1$ and approximating the von Mises density with a Gaussian density with standard deviation τ, the following expression for $\mathbf{K}(\theta)$ is obtained,

$$\mathbf{K}(\theta) = \sigma^{-2}\begin{pmatrix} 1 & 0 \\ 0 & 3^{-1}(1+3\tau^{-2}\sigma^2-\xi^2)\end{pmatrix}. \quad (33)$$

The expression (33) for $\mathbf{K}(\theta)$ can be obtained directly from (16). A consequence of (33) is that if $\tau^{-1}\sigma$ is small, then the contribution of $\mathbf{K}''(\theta)$ to $\mathbf{K}(\theta)$ is small. In particular, suppose that a disk shaped image with radius r is scaled to the unit disk, such that $\sigma = O(r^{-1})$. If r is large then $\tau^{-1}\sigma$ is small. It follows that there is an upper limit on the size of an image for which $\mathbf{K}''(\theta)$ makes a significant contribution to $\mathbf{K}(\theta)$.

6 Bayesian Algorithm for Curve Detection

A Bayesian algorithm for curve detection is described in Sect. 6.1. Experimental results obtained on applying the algorithm to circle detection are reported in Sect. 6.2. Experimental results obtained on applying a Hough transform based algorithm to the same task of circle detection are reported in Sect. 6.3.

6.1 Curve detection

Let the S_i for $1 \le i \le N$ be subsets of the parameter manifold T, and let X be a set of n compound measurements,

$$X = \{(\mathbf{x}_1,\alpha_1),\ldots,(\mathbf{x}_n,\alpha_n)\},$$

where the $\mathbf{x}_i$ are points in the image and α_i is the orientation of a step edge centred at $\mathbf{x}_i$. It is assumed that the elements of X are independent samples from a conditional pdf

$$q(\mathbf{x},\alpha|\theta) = \delta p(\mathbf{x},\alpha|\theta) + (1-\delta)/(2\pi^2), \quad \mathbf{x}\in D, \alpha\in[0,2\pi), \theta\in T,$$

where δ is the prior probability that the measurement $(\mathbf{x},\alpha)$ is an inlier to $p(\mathbf{x},\alpha|\theta)$. The outlying measurements are sampled from the uniform distribution on $D\times[0,2\pi)$.

Let H_i be the hypothesis that there is a single curve to be detected and that this curve is specified by a point in S_i. The conditional probability $P(H_i|X)$ is given by

$$P(H_i|X) = \int_{S_i} q(\theta|X)\,d\theta, \quad 1 \le i \le N. \quad (34)$$

Let $q(\theta)$ be the prior pdf for θ and let $q(X)$ be the pdf for the measurements X. An application of Bayes rule to (34) yields

$$P(H_i|X) = q(X)^{-1}\int_{S_i}\left(\prod_{j=1}^n q(x_j,\alpha_j|\theta)\right)q(\theta)\,d\theta, \quad 1 \le i \le N. \quad (35)$$

If the S_i are disjoint, then $q(X)$ can be found using the condition

$$\sum_{i=1}^{N} P(H_i|X) = 1.$$

Let $P(B_i|X)$ be the probability assigned to S_i under the hypothesis that the measurements in X are sampled from the uniform distribution on $D \times [0, 2\pi)$. These measurements contain no information about the presence of a curve, thus $P(B_i|X)$ is given by

$$P(B_i|X) = \int_{S_i} q(\theta) d\theta.$$

If the sets S_i are disjoint but do not cover T, then the $P(B_i|X)$ are scaled such that

$$\sum_{i=1}^{n} P(B_i|X) = 1.$$

A curve associated with S_i is detected if

$$P(H_i|X)/P(B_i|X) \gg 1. \tag{36}$$

The above algorithm can be simplified if T is a bounded open subset of $\mathbb{R}^d$ for some value of d and the sets S_i have a small diameter. To be specific, let $t > 0$ be a scale factor, let θ_i for $1 \leq i \leq N$ be points in T at the vertices of a scaled integer lattice $t\mathbb{Z}^d$ and let S_i be the closure of the set of points in T that are nearer to θ_i than to any other lattice point. Subsets S_i, S_j with $i \neq j$ may overlap but this does not cause any difficulty because the d-volume of any intersection $S_i \cap S_j$ with $j \neq i$ is zero. Sets S_i that are not wholly contained in T are discarded. Each remaining set S_i is a hypercube centred at θ_i and with side length t.

It is assumed that t is relatively small. The conditional probability $P(H_i|X)$ is estimated by

$$P(H_i|X) = \lambda_1 \left(\prod_{j=1}^{n} q(x_j, \alpha_j|\theta_i) \right) P(B_i|X), \quad 1 \leq i \leq N, \tag{37}$$

where λ_1 is a scale factor chosen such that

$$\sum_{i=1}^{N} P(H_i|X) = 1.$$

The conditional probability $P(B_i|X)$ is estimated by

$$P(B_i|X) = \lambda_2 t^d \sqrt{\det(\mathbf{K}(q, \theta_i))}, \quad 1 \leq i \leq N. \tag{38}$$

where λ_2 is a scale factor chosen such that

$$\sum_{i=1}^{N} P(B_i|X) = 1.$$

The algorithm for curve detection is summarised as follows.

Name: Algorithm 1
Inputs: i) a bounded open subset T of $\mathbb{R}^d$; ii) parameters σ, τ, δ, t, a; and iii) a list X of compound measurements $(\mathbf{x}, \alpha)$.
Output: A finite subset Θ of T.

1. obtain the set $T \cap t\mathbb{Z}^d \equiv \{\theta_1, \ldots, \theta_N\}$;
2. $\Theta \leftarrow \emptyset$;
3. for $1 \leq i \leq N$
   3.1 define $P(B_i|X)$ by (38); // σ, τ, δ required
   3.2 define $P(H_i|X)$ by (37); // σ, τ, δ required
   3.3 if $P(H_i|X)/P(B_i|X) \geq a$, then $\Theta \leftarrow \Theta \cup \{\theta_i\}$;
4. endFor;
5. return Θ

The efficiency of the curve detection is increased by applying Algorithm 1 iteratively. Suppose that the first iteration yields a set Θ of parameter vectors. In the next iteration the value of t is reduced and the sets S_i are chosen near to the parameter vectors in Θ. The iteration is continued until t is less than a specified threshold.

Algorithm 1 appears at first sight to be based on a Hough transform. However, there is an essential difference to Hough transform based algorithms. In (37) the product of the densities $q(\mathbf{x}_j, \alpha_j|\theta_i)$ is never zero. In Hough transform based methods, measurements that are not on any curve specified by an element of S_i make a zero contribution to the accumulator for S_i. It follows that the contributions to an accumulator from different measurements cannot be multiplied, because in practice the result would always be zero. This difficulty can be avoided by adding the contributions instead of multiplying them, but the result of this addition is not a probability. In this case, the curve detection method ceases to be Bayesian, and thus does not make full use of the available information (Jaynes 2003).

6.2 Iris detection using Algorithm 1

Algorithm 1, as described in Sect. 6.1, was adapted for iris detection using the images in the CASIA Iris Interval database (CASIA 2010). The aim is to detect the two circles that together form the boundary of the iris. The purpose of the adaptations is to reduce the number of measurements and the number of lattice points that need to be taken into account. The parameters for the edge detection described in Section 3 are: image size 280×320, $m = 11$, $b = 8$, $\psi = 0.75$. The remaining parameters are $\sigma = 0.014$, $\tau = 1/m$, $\delta = 0.2$, $\xi_{min} = 15\sigma$, $t = 10\sigma$ on the first iteration and $t = 10\sigma/7$ on the second iteration.

The set X of measurements and the set of lattice points θ_i, $1 \leq i \leq N$ in T are adapted as follows. Let $\mathbf{c}_I$ be the centre of the iris image and let $D(\mathbf{c}_I, r)$ be the disk centred at $\mathbf{c}_I$ and with the maximum radius r, given that $D(\mathbf{c}_I, r)$ is contained in the iris image. All measurements $(\mathbf{x}, \alpha)$ such that $\mathbf{x}$ is not in $D(\mathbf{c}_I, r)$ are discarded. The remaining measurements are translated and scaled,

$$(\mathbf{x}, \alpha) \mapsto (r^{-1}(\mathbf{x} - \mathbf{c}_I), \alpha)$$

to obtain measurements in the unit disk D centered at the origin. It is assumed that the two circles bounding the iris

are concentric and that the common centre is contained in a disk $D' \subset D$, centred at the origin and with radius 1/4. All lattice points θ_i for circles with centres outside D' are discarded. Let $(\mathbf{x}, \alpha)$ be a measurement. Any circle through $\mathbf{x}$ and with tangent direction α at $\mathbf{x}$ has its centre on a line though $\mathbf{x}$. The measurement is discarded if this line does not intersect D'.

After relabeling if necessary, let θ_i, $1 \leq i \leq N$, be the remaining lattice points and let X be the set of remaining measurements. Let i_0 be the index defined by

$$i_0 = \text{argmax}_{1 \leq i \leq N}\, i \mapsto P(H_i|X).$$

The circle θ_{i_0} is detected. The condition (36) is not required because in this application $P(H_{i_0}|X)$ is very near to 1. In the second iteration, $t = 10\sigma/7$ and T is sampled at $7 \times 7 \times 7$ points at the centres of a set of cubes centred on θ_{i_0}. The circle θ_{i_1} detected at this stage is the boundary between the iris and the pupil. The parameter space for the circle forming the second part of the boundary of the iris has dimension 1, because the centre of the second circle is set equal to the centre of θ_{i_1}. Only one iteration of the Bayesian algorithm is required, with $t = \sigma/2$.

The results of circle detection are shown in Fig. 5 for the first nine images in the CASIA Iris Interval database[2].

6.3 Iris detection using Hough transform based algorithms

Algorithm 1, as described in Sect. 6.1 and adapted for iris detection in Sect. 6.2, is in this Section reduced to a Hough transform based algorithm, which is referred to as Algorithm 2. The notation in Sect. 6.1 is used.

As in Algorithm 1, the set $T \cap t\mathbb{Z}^3$ is obtained. This set and the set of measurements are edited as described in Sect. 6.2. Each lattice point θ_i in $T \cap t\mathbb{Z}^3$ is at the centre of a cube S_i in $\mathbb{R}^3$ with side length t. The associated accumulator is A_i.

The points of T corresponding to circles exactly compatible with the compound measurement $(\mathbf{x}, \alpha)$ are contained in the lines

$$\xi \mapsto (\xi, x + \xi(-\sin(\alpha), \cos(\alpha)))^\top, \quad \xi_{min} \leq \xi \leq 1, \qquad (39)$$

$$\xi \mapsto (\xi, x - \xi(-\sin(\alpha), \cos(\alpha)))^\top, \quad \xi_{min} \leq \xi \leq 1. \qquad (40)$$

The value of A_i is equal to the total number of measurements $(\mathbf{x}, \alpha)$ for which either of the lines (39), (40) intersects S_i. The circle θ_i with the largest value of A_i is detected.

The results obtained on applying Algorithm 2 to the same 9 iris images to which Algorithm 1 is applied are shown in Fig. 6. In all 9 cases the inner boundary of the iris is detected. However, the less well defined outer boundary is detected in only 4 out of the 9 cases.

[2] http://biometrics.idealtest.org/dbDetailForUser.do?id=4

Circle detection on the same 9 images was also carried out using the Hough transform based MATLAB function imfindcircles[3]. The following parameters were used: radius range [30 150], Sensitivity 0.97 and ObjectPolarity set to dark. The results are shown in Fig. 7. In all 9 cases the inner boundary of the iris is detected but the outer boundary is not detected. There are many false detections. If the Sensitivity is reduced to 0.9 then in every case the inner boundary of the iris is detected, the outer boundary is not detected and there are no false alarms.

7 Conclusion

In this paper a tractable approximation is obtained to the Fisher-Rao metric for a parameterised family of curves in an image. The metric is based on compound measurements that consist of a pixel and the direction of the grey level gradient at the pixel. The additional information in the grey level gradients is measured by a non-negative matrix, which is added to the matrix for the Fisher-Rao metric based on measurements that consist only of pixels in the image.

The probability density function for a measurement given a curve is extended in the usual way to account for outliers. If the curves are circles or straight lines, then the associated Fisher-Rao metric has an accurate approximation which can be evaluated numerically. The approximating metric is used to estimate the minimum number of measurements required to detect a circle in the presence of outliers. In addition, the approximating metric is the basis of the first Bayesian algorithm for curve detection in the presence of outliers. The extension of the pdf to take account of outliers ensures that it is not necessary to separate the inliers from the outliers during the process of curve detection.

There are two major advantages of Bayesian methods for curve detection. The first advantage is that the best use is made of the available information. The second advantage is that the parameter space for the curves can be edited to remove subsets that are known a priori to be irrelevant. The Bayesian nature of the algorithm is maintained by renormalising the probabilities for the remaining parts of the parameter space.

References

1. Aguado, A.S., Montiel, M.E., Nixon, M.S.: On using directional information for parameter space decomposition in ellipse detection. Pattern Recognition, 29, 369-381 (1996)
2. Amari, S.-I.: Differential-Geometric Methods in Statistics. Lecture Notes in Statistics, vol. 28. Springer (1985)

[3] http://uk.mathworks.com/help/images/ref/imfindcircles.html

3. Balasubramanian, V.: Statistical inference, Occam's razor and statistical mechanics on the space of probability distributions. Neural Computation, 9, 349-368 (1997)
4. Ballard, D.H.: Generalizing the Hough transform to detect arbitrary shapes. Pattern Recognition, 13, 111-122 (1981)
5. Bonci, A., Leo, T., Longhi, S.: A Bayesian approach to the Hough transform for line detection. IEEE Transactions on Systems, Man, and Cybernetics, Part A: Systems and Humans, 35, 945-955 (2005)
6. CASIA Iris Image Database: http://biometrics.idealtest.org (2010)
7. Ceolin, S., Hancock, E.R.: Distinguishing facial expression using the Fisher-Rao metric. Proc. IEEE Conf. on Image Processing (ICIP), 1437-1440 (2010)
8. Ceolin, S.R., Hancock, E.R.: Computing gender difference using Fisher-Rao metric from facial surface normals. Proc. Conference on Graphics, Patterns and Images,(SIBGRAPI), 336-343 (2012)
9. Cover, T.M., Thomas, J.A.: Elements of Information Theory. John Wiley and Sons (1991)
10. Forsyth, D.A., Ponce, J.: Computer Vision: A Modern Approach, 2nd Edition. Prentice Hall (2011)
11. Gonzalez, R.C., Woods, R.E.: Digital Image Processing. 3rd edition, Pearson Education (2008)
12. Jaynes, E.T.: Probability Theory: the logic of science, CUP (2003)
13. Jeffreys, H.: Theory of Probability. Oxford Classics Series, OUP (1998)
14. Ji, Q., Haralick, R.M.: An optimal Bayesian Hough transform for line detection. Proceedings of the 1999 International Conference on Image Processing, 2, 691-695 (1999)
15. Kanatani, K.-I.: Statistical Computation for Geometrical Optimization, Elsevier (1996)
16. Kimme, C., Ballard, D., Sklansky, J.: Finding circles by an array of accumulators. Communications of the Association of Computing Machinery, 18, 120-122 (1975)
17. Leavers, V.F.: Which Hough transform? Computer Vision, Graphics, and Image Processing, 58, 250-264 (1993)
18. Mardia, K.V., Jupp, P.E.: Directional Statistics. Wiley (2000)
19. Maybank, S.J.: Detection of image structures using the Fisher information and the Rao metric. IEEE Transactions on Pattern Analysis and Machine Intelligence, 26, 1579-1589 (2004)
20. Maybank, S.J.: Application of the Fisher-Rao metric to ellipse detection. International Journal of Computer Vision, 72, 287-307 (2007)
21. Maybank, S.J., Ieng, S., Benosman, R.: A Fisher-Rao metric for paracatadioptric images of lines. International Journal of Computer Vision, 99, 147-165. DOI 10.1007/s11263-012-0523-x (2012)
22. Michor, P.W., Mumford, D.: An overview of the Riemannian metrics on spaces of curves using the Hamiltonian approach. Applied and Computational Harmonic Analysis, 23, 74-113 (2007)
23. Mio, W., Srivastava, S., Joshi, S.: On shape of plane elastic curves. International Journal of Computer Vision, 73, 307-324 (2006)
24. Olson, C.F.: Constrained Hough transform for curve detection. Computer Vision and Image Understanding, 73, 329-345 (1999)
25. OpenCV: http://docs.opencv.org/modules/imgproc/doc/feature_detection.html#houghcircles (2014). Page accessed 20.2.14.
26. Peter, A., Rangarajan, A.: Information geometry for landmark shape analysis: unifying shape representation and deformation. IEEE Transactions on Pattern Analysis and Machine Intelligence,31, 337-350 (2009)
27. Rao, C.: Information and the accuracy obtainable in the estimation of statistical parameters. Bulletin Calcutta Mathematical Society, 37, 81-91 (1945)
28. Srivastava, A., Jermyn, I., Joshi, S.: Riemannian analysis of probability density functions with applications in vision. IEEE Conference on Computer Vision and Pattern Recognition, CVPR2007, 1-8 (2007)
29. Srivastava, A., Klassen, E., Joshi, S. H., Jermyn, I.: Shape analysis of elastic curves in Euclidean spaces. IEEE Transactions on Pattern Analysis and Machine Intelligence, 33, 1415-1428 (2011)
30. Sundaramoorthi, G., Mennucci, A.C.G., Soatto, S., Yezzi, A.: A new geometric metric in the space of curves, and applications to tracking deforming objects by prediction and filtering. SIAM Journal on Imaging Sciences, 4, 109-145 (2011)
31. Szeliski, R.: Computer Vision: algorithms and applications. Springer-Verlag London Limited (2011)
32. Tatu, A., Lauze, F., Sommer, S., Nielsen, M.: On restricting planar curve evolution to finite dimensional implicit subspaces with non-euclidean metric. Journal of Mathematical Imaging and Vision, 38, 226-240 (2010)
33. Toronto, N., Morse, B.S., Ventura, D., Seppi, K.: (2007) The Hough transform's explicit Bayesian foundation. Proceedings of the 14th International Conference on Image Processing, IV, 377-380 (2007)
34. Werman, M., Keren, D.: A Bayesian method for fitting parametric and nonparametric models to noisy data. IEEE Transactions on Pattern Analysis and Machine Intelligence, 23, 528-534 (2001)
35. Woodford, O. J., Pham, M.-T., Maki, A., Porbet, F., Stenger B.: Demisting the Hough transform for 3D shape recognition and registration. International Journal of Computer Vision, 106, 332-341 (2014)

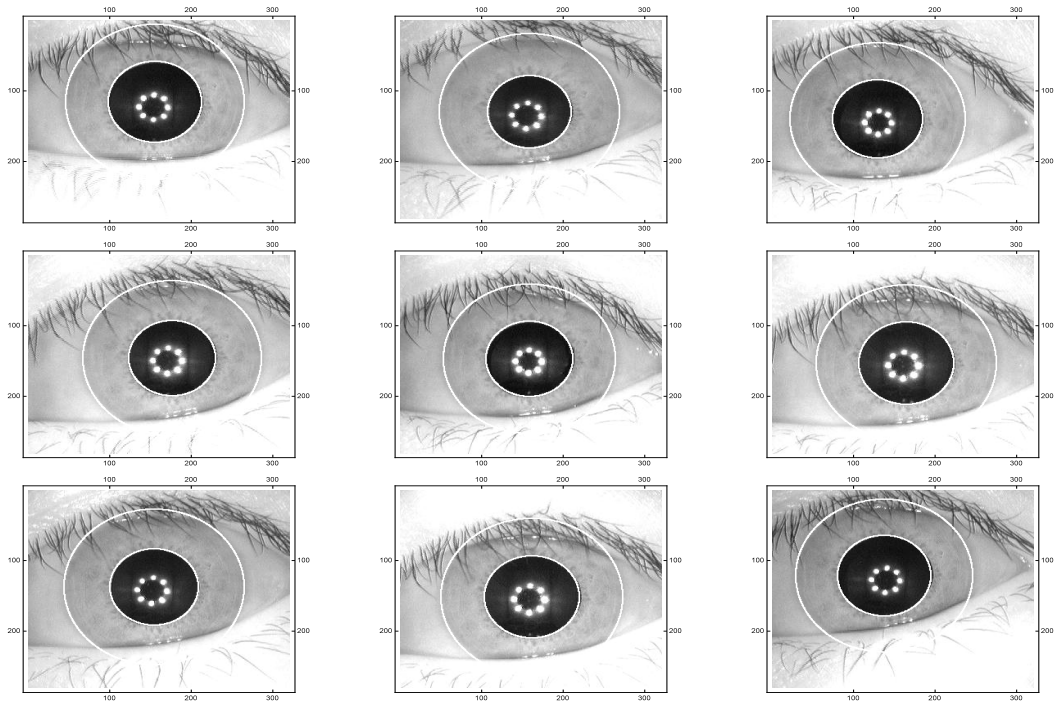

Fig. 5 Circles detected by Algorithm 1 in the first nine images of the CASIA Iris Interval database.

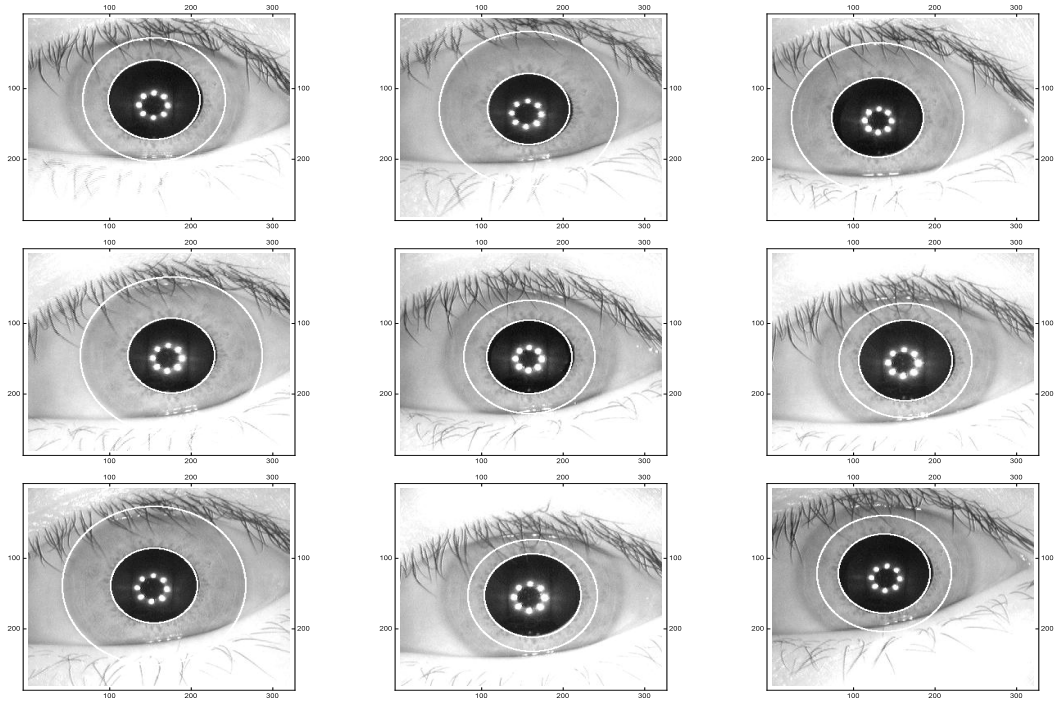

Fig. 6 Circles detected by Algorithm 2 in the first nine images of the CASIA Iris Interval database.

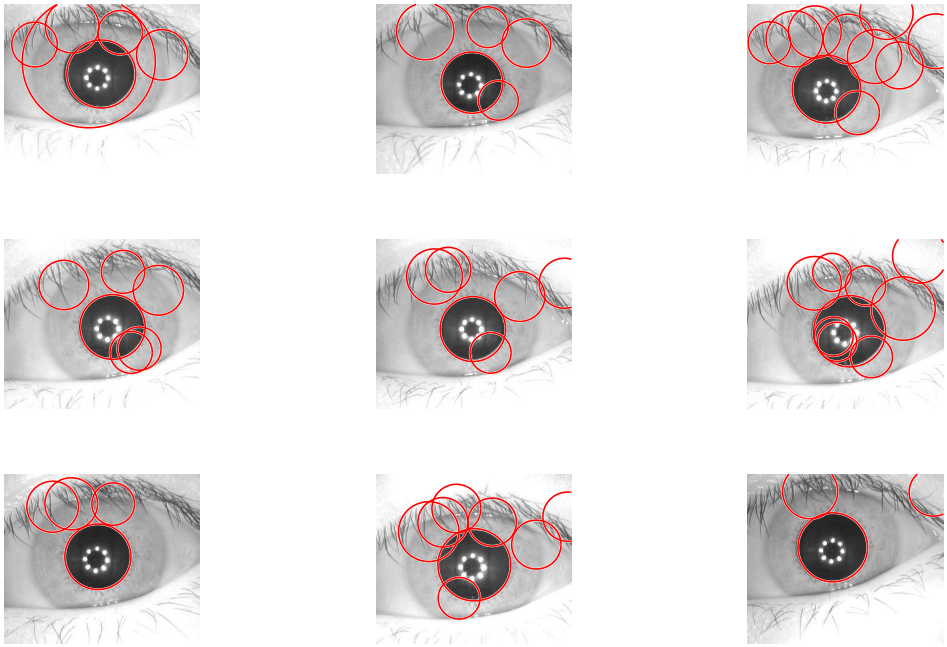

Fig. 7 Circles detected by the MATLAB function imfindcircles with Sensitivity 0.97.