Fitting Ellipses and Predicting Confidence Envelopes using a Bias Corrected Kalman Filter

John Porrill

AI Vision Research Unit
University of Sheffield
Sheffield S10 2TN
England

We describe the use of the Kalman filter to find optimal fits to short sections of ellipse data and to predict confidence envelopes in order to facilitate search for further ellipse data. The extended Kalman filter in its usual form is shown not to reduce the well known bias to high curvature involved in least squares ellipse fitting. This problem is overcome by developing a linear bias correction for the extended Kalman filter. The estimate covariance is used to evaluate confidence envelopes for the fitted ellipse. Performance is shown on both real and synthetic data.

1. INTRODUCTION

We are interested in fitting ellipses to 2D edge data as an extension to the range of 2D primitives available in the TINA vision system (see [1]) and as a preliminary stage in the accurate recovery of 3D circles from stereo edge data. Our starting assumption is that a non-straight edge segment has been identified, and we wish to find, relatively efficiently, the best fitting ellipse, to test whether the fit is reasonable, and to predict a confidence region to direct the search for continuations of the ellipse.

This last requirement is possibly the most important since most methods perform so well given sufficient data that it is easy to underestimate the instability of the problem for short curve sections. The following order of magnitude argument illustrates this instability dramatically. The overall shape of an ellipse depends on the second derivative of curvature at a point, that is, on fourth derivatives of position. Hence given N points of accuracy σ on a small length L of curve, position measurements are accurate to order $\sigma N^{-1/2}$, and 4'th derivatives of position are thus accurate to order $\sigma N^{-1/2} L^{-4}$. If the density of points is $\rho = N/L$ then the overall accuracy of the ellipse is of order $\sigma \rho^{-1/2} L^{-9/2}$. Given fixed accuracy in position, in order to achieve constant accuracy of the fitted ellipse for short sections the density of points must vary inversely as the ninth power of the length of the segment!

There are many ellipse fitting methods available in the literature, but most can be be ruled out for our purposes on the following grounds: requiring data from a large proportion of the ellipse, being non-optimal, being biased to high curvature solutions, being based on non-linear optimisation and hence too slow to be practical, or being incapable of providing reliable confidence estimates. For descriptions of other methods, and for ellipse detection methods, see for example [2,3,4,5].

The extended Kalman filter is finding increasing use in vision applications (for examples see [6]) and is the basis of the geometrical statistics module in the TINA system (described in [7]) It is efficient, estimates its own error, and is applicable to contaminated sequential data. It is thus a natural candidate for our purpose. However experiments with the standard form of the extended Kalman filter showed surprisingly that it is just as biased towards high curvature fits as naive least squares. We have developed a bias corrected Kalman filter based on a linearisation of the maximum likelihood principle rather than the measurement equation which greatly improves performance in this regard.

In §2 we present the problem of ellipse fitting in a suitable form for application of the Kalman filter the properties of which are described briefly in §3. In §4 the filter is applied to ellipse fitting and in §5 we use the covariance estimate from the Kalman filter to evaluate the confidence envelope for the predicted ellipse. In §6 we show examples of ellipse fitting and prediction for synthetic and real data.

2. DESCRIBING CONICS

We describe the general conic by the biquadratic equation

$$ax^2 + 2bxy + cy^2 + 2dx + 2ey + f = 0.$$

Since the trace $a + c$ can never be zero for an ellipse the arbitrary scale factor can be removed by the normalisation

$$a + c = 1$$

(with this normalisation rectangular hyperbolae cannot be represented, the more obvious scaling $f = 1$ would exclude ellipses which pass through the origin). All ellipses can then be described by the 5-vector

$$\mathbf{x} = (a, b, d, e, f)^t$$

which we wish to estimate given observations

$$\mathbf{y}_i^* = (x_i, y_i)^t$$

related to their true values by

$$\mathbf{y}_i^* = \mathbf{y}_i + \mathbf{v}_i$$

where the noise $\mathbf{v}_i$ has expected value zero and covariance

$$\text{Cov}[\mathbf{v}_i] = R_i = \begin{bmatrix} \Sigma^2 & 0 \\ 0 & \Sigma^2 \end{bmatrix}$$

For each point the equation of the conic becomes a non-linear constraint

$$F(\mathbf{x}, \mathbf{y}_i) = 0.$$

which we shall impose using the extended Kalman filter.

3. THE KALMAN FILTER

The linear, extended and bias corrected Kalman filters for measurement constraints such as that above are described briefly below (for more details on the Kalman filter see [8], for more details on correcting linearisation bias see [9]).

We use the fact that the minimum of the quadratic form

$$Q(\mathbf{x}) = (\mathbf{x} - \mathbf{x}_0)^t\, S_0^{-1}\, (\mathbf{x} - \mathbf{x}_0) + \sum_{i=1}^{p} \frac{(z_i - \mathbf{h}_i^t\, (\mathbf{x} - \mathbf{x}_i^*))^2}{\sigma_i^2}$$

can be calculated by iterating the Kalman filter

$$\mathbf{x}_i = \mathbf{x}_{i-1} - (z_i - \mathbf{h}_i^t\,(\mathbf{x}_{i-1} - \mathbf{x}_i^*))\, \frac{S_{i-1}\mathbf{h}}{\sigma_i^2 + \mathbf{h}_i^t S_{i-1}\mathbf{h}_i}$$

$$S_i = S_{i-1} - \frac{(S_{i-1}\mathbf{h}_i)\,(S_{i-1}\mathbf{h}_i)^t}{\sigma_i^2 + \mathbf{h}_i^t S_{i-1}\mathbf{h}_i}.$$

p times starting with the initial values $\mathbf{x}_0$, S_0.

This corresponds to finding the maximum likelihood estimate of a vector $\mathbf{x}$ given prior knowledge that $\mathbf{x}$ has expected value $\mathbf{x}_0$ and covariance S_0 and given the values z_i of a series of measurements

$$z_i = \mathbf{h}_i^t\,(\mathbf{x}_i - \mathbf{x}_i^*) + u_i \quad E[u_i] = 0 \quad \text{Var}[u_i] = \sigma_i^2.$$

The normalised residual

$$\varepsilon_i = \frac{(z_i - \mathbf{h}_i^t\,(\mathbf{x}_{i-1} - \mathbf{x}_i^*))^2}{\sigma_i^2 + \mathbf{h}_i^t S_{i-1}\mathbf{h}_i}$$

is a χ^2 variable on one degree of freedom.

This can be generalised to the case of a series of p measurements $\mathbf{y}_i^*$ of vectors $\mathbf{y}_i$ related to $\mathbf{x}$ by a non-linear scalar constraint

$$F(\mathbf{x}, \mathbf{y}_i) = 0.$$

where the measurement noise is Gaussian with

$$E[\mathbf{y}_i - \mathbf{y}_i^*] = 0 \quad \text{Cov}[\mathbf{y}_i - \mathbf{y}_i^*] = R_i$$

if we linearise the constraint about estimates $\mathbf{x}_i^*$ of $\mathbf{x}$ and the measurements $\mathbf{y}_i^*$

$$F + (\nabla_{\mathbf{x}}F)\,(\mathbf{x} - \mathbf{x}_i^*) + (\nabla_{\mathbf{y}}F)\,(\mathbf{y}_i - \mathbf{y}_i^*) = 0.$$

Making the substitutions (all function values are taken at $\mathbf{x}_i^*, \mathbf{y}_i^*$)

$$z_i = -F$$

$$\mathbf{h}_i = \nabla_{\mathbf{x}}F$$

puts this in the form of the linear measurement equation above with noise term

$$u_i = (\nabla_{\mathbf{y}}F)(\mathbf{y}_i - \mathbf{y}_i^*)$$

having zero expectation and variance

$$\sigma_i^2 = \text{Var}[u_i] = (\nabla_{\mathbf{y}}F)\, R_i\, (\nabla_{\mathbf{y}}F)^t$$

The Kalman filter for the linearised measurement constraint is called the extended Kalman filter.

An alternative derivation of the linearised filter begins by suitably approximating the maximum likelihood principle. The maximum likelihood estimate for $\mathbf{x}$ (with prior knowledge) minimises

$$Q = (\mathbf{x} - \mathbf{x}_0)^t\, S_0^{-1}\, (\mathbf{x} - \mathbf{x}_0) + \sum_{i=1}^{p} (\mathbf{y}_i - \mathbf{y}_i^*)^t\, R_i^{-1}\, (\mathbf{y}_i - \mathbf{y}_i^*)$$

over all $\mathbf{x}$, $\mathbf{y}_i$ subject to the measurement constraints. We will find the minimum in the Levenberg-Marquadt approximation that the d_i are small, and so can be approximated to linear order.

It is clear that the minima over $\mathbf{y}_i$ can be calculated independently for each term

$$d_i^2 = (\mathbf{y}_i - \mathbf{y}_i^*)^t\, R_i^{-1}\, (\mathbf{y}_i - \mathbf{y}_i^*)$$

subject to the single constraint

$$F(\mathbf{x}, \mathbf{y}_i) = 0$$

At the minimum d_i is the perpendicular distance from the point $\mathbf{y}_i^*$ to the surface $F = 0$ in $\mathbf{y}$ space. For $\mathbf{y}_i^*$ close to the surface d_i can be calculated to linearised order as

$$d_i = F(\mathbf{x}, \mathbf{y}_i^*)/\sigma_i(\mathbf{x})$$

where we have defined

$$\sigma_i(\mathbf{x}) = \left[(\nabla_{\mathbf{y}}F(\mathbf{x}, \mathbf{y}_i^*))\, R_i\, (\nabla_{\mathbf{y}}F(\mathbf{x}, \mathbf{y}_i^*))^t \right]^{1/2}$$

If we expand both numerator and denominator to linear order the estimate of $Q(\mathbf{x})$ is

$$Q(\mathbf{x}) = (\mathbf{x} - \mathbf{x}_0)^t\, S_0^{-1}\, (\mathbf{x} - \mathbf{x}_0) + \sum_{i=1}^{p} \frac{(z_i - \mathbf{h}_i^t\,(\mathbf{x} - \mathbf{x}_i^*))^2}{(\sigma_i + \nabla_{\mathbf{x}}\sigma_i(\mathbf{x}_i)(\mathbf{x} - \mathbf{x}_i^*))^2}$$

With the same substitutions as above. The minimum principle for the extended Kalman filter is thus the result of expanding the numerator of d_i to linear order in $\mathbf{x} - \mathbf{x}_i^*$ but ignoring the linear correction in the denominator. It fails to take into account the change in weight of a given measurement as the constraint parameters change. The result in our ellipse fitting application is that the fitted ellipse tries to locate areas of high curvature on the fitted segment artificially reducing Q, as will be seen in the examples.

The solution is to retain the linear order terms in the denominator, luckily this can be done without changing the form of Q, in fact to linear order
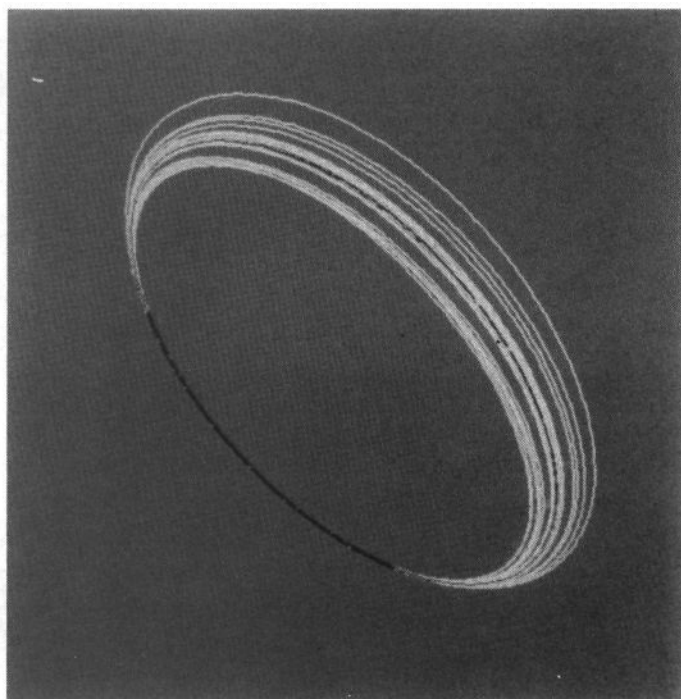
$$Q(\mathbf{x}) = (\mathbf{x} - \mathbf{x}_0)^t\, S_0^{-1}\, (\mathbf{x} - \mathbf{x}_0) + \sum_{i=1}^{p} \frac{(z_i - \bar{\mathbf{h}}_i^t\,(\mathbf{x} - \mathbf{x}_i^*))^2}{\sigma_i^2}$$

where

Figure 1 : Correction of bias to high curvature
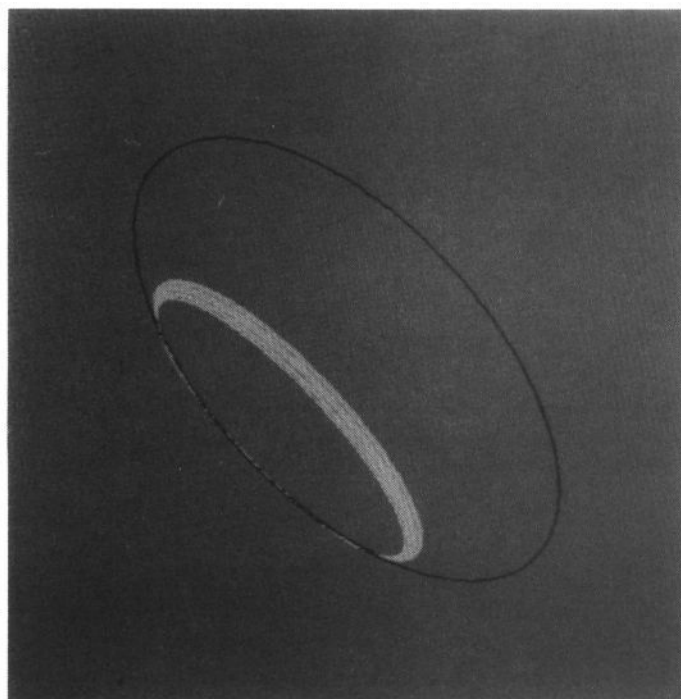
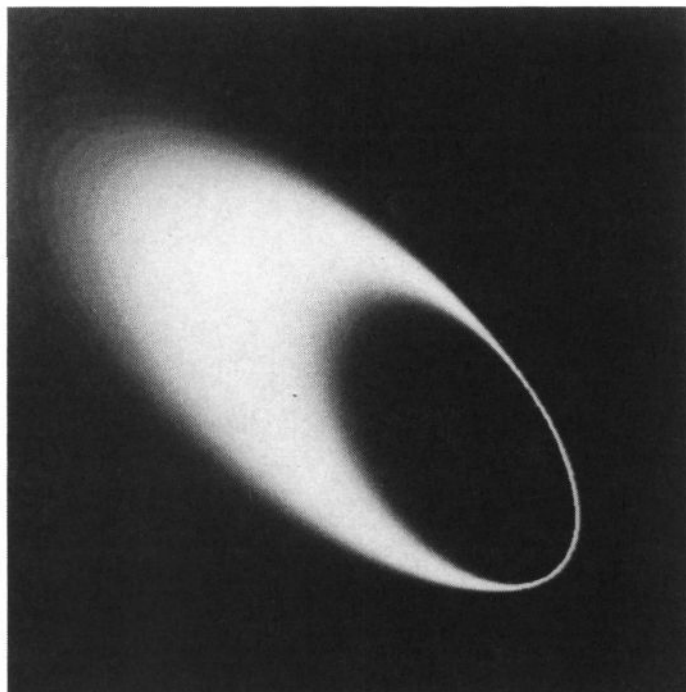(a) Extended Kalman Filter

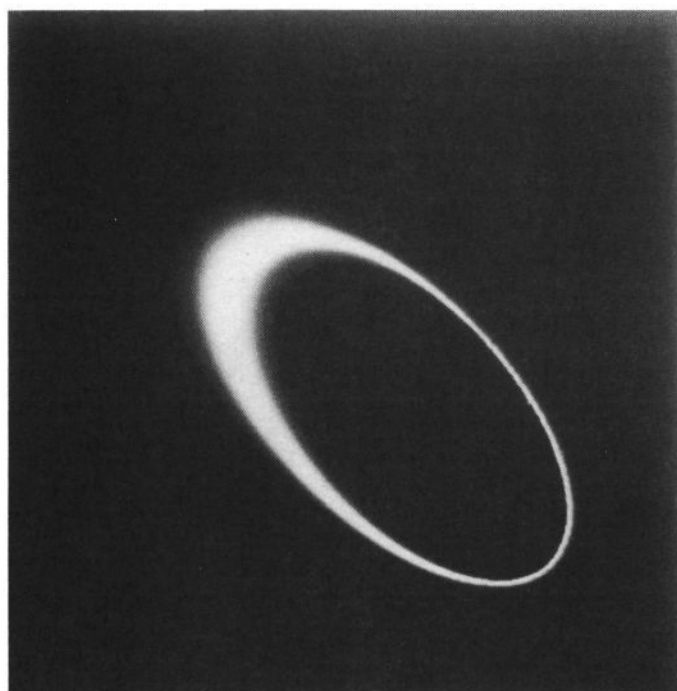(b) Bias Corrected Kalman Filter

(c) Predicted Confidence Surface

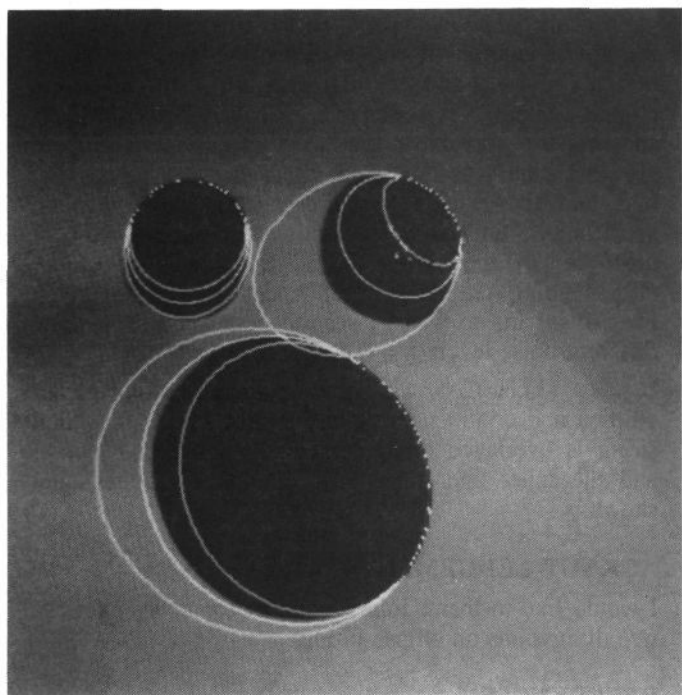# Figure 2 : Instability of prediction from short sections
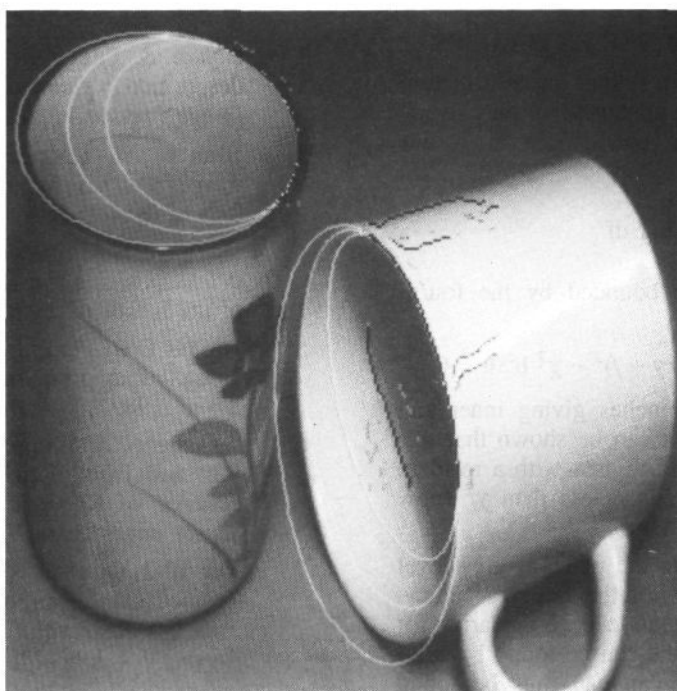

(a) Σ = 0.5

(b) Σ = 0.1

(c) Σ = 0.02

Figure 3 : Examples using real data

(a) Disc & Lens Caps

(b) Widgets

(c) Mugs

$$\tilde{\mathbf{h}}_i = \mathbf{h}_i + \frac{z_i}{\sigma_i} \nabla_{\mathbf{x}}\sigma_i(\mathbf{x}_i)$$

the correction is thus made by adding a term to the plant vector of the Kalman filter, this terms depends on an $\mathbf{x}$-derivative of σ and thus on the cross second derivative $\nabla_{\mathbf{y}}\nabla_{\mathbf{x}}F$ of the measurement constraint. We will call this the bias corrected Kalman filter.

4. APPLYING THE KALMAN FILTER

An initial estimate of the parameters $\mathbf{x}_0$ of the best fit conic to a point string is obtained by calculating the conic passing exactly through five well spaced points of the string using the four-line method (if this initial conic is not an ellipse convergence is more likely if a three point circle fit is chosen). This initial estimate is rendered infinitely uncertain by setting the initial covariance S_0 to be diagonal and large.

The extended Kalman filter described above is then applied sequentially to impose the measurement constraint at each point of the string with the substitutions

$$z = -(ax^2 + 2bxy + (1-a)y^2 + 2dx + 2ey + f = 0)$$

$$\mathbf{h} = \nabla_{\mathbf{x}}F = (x^2 - y^2 \quad 2xy \quad 2x \quad 2y \quad 1)^t$$

$$\sigma = 4\Sigma^2((ax+by+d)^2 + (bx+cy+e)^2)$$

The fitting process is iterated by using the result of this process as new initial estimate until changes in ellipse parameters are less than their expected error.

Note that a robust implementation of the filter such as the Bierman U-D factorised form [10] is essential to avoid numerical problems.

5. CALCULATING CONFIDENCE ENVELOPES

Once the best fit ellipse has been found, the information about the accuracy of the fit is contained in the covariance matrix S. The test that a general point (x, y) lies on the ellipse is given by

$$\chi^2 = \frac{z^2}{\mathbf{h}'S\mathbf{h}^t} \text{ on 1 df}$$

so that confidence regions are bounded by the fourth order curve

$$(ax^2 + 2bxy + cy^2 + 2dx + 2ey + f)^2 - \chi^2 \, \mathbf{h}'S\mathbf{h} = 0$$

(for an ellipse this has two branches giving inner and outer confidence boundaries). It can be shown that this curve is also the envelope of all ellipses with a residual deviation from the estimated ellipse less than χ^2 (now interpreted on 5 degrees of freedom).

Plotting the confidence values in the plane as grey level values gives a visual realisation of the fitting accuracy. The intersection of the 95% confidence boundaries for the fitted ellipse with a given line can be found by solving a quartic. The boundaries can can thus be found efficiently by plotting their intersection with lines through the ellipse centre.

6. EXAMPLES

All the examples show a notional 256×256 image.

Figures (1a,b,c) were constructed as an extreme example of the bias associated with using the extended Kalman filter. The data consists of 320 points generated on a short section of an exact ellipse to which Gaussian errors of standard deviation 0.2 pixels have been added. Such data sets were generated repeatedly, and the fitting method applied, a sample of 20 fits was taken to show the performance of each method. To allow fair comparison all methods were given the true ellipse as starting approximation. It shows the true ellipse the family of fitted ellipses, and the total data set overlayed on the ellipse. Figure (1a) uses the extended Kalman filter, showing the very clear bias to high curvature. Figure (1b) uses the bias corrected filter, there is no obvious bias and the sample variation agrees well with the predicted confidence surface shown in Figure (1c). Figures (2a,b,c) show very clearly the instability of the ellipse fitting problem. Confidence surfaces are given for 5-point fits to a short section of an ellipse at three different noise levels $\Sigma = 0.5, 0.1, 0.02$ pixels.

Figures (3a,b,c) shows some examples of predicted ellipses and confidence regions for real data. The fitted string is overlayed on the fitted ellipse (middle curve) and the inner and outer 95% confidence bounds are shown.

REFERENCES

1   **Pollard S B, Pridmore T P, Porrill J, Mayhew J E W & Frisby J P,** Geometrical modelling from multiple stereo views, *International Journal of Robotics Research,* to appear in 1989.

2   **Nalwa & Pauchon,** Edgel aggregation and edge description, *Computer Vision, Graphics, and Image Processing,* 40, 1987.

3   **Blunck M,** *Automatic analysis of diffraction fringe patterns in electron microscopy,* Proceedings of the fourth Alvey vision conference, 1988.

4   **Yuen H K, Illingworth J & Kittler J,** *Ellipse detection using the Hough transform,* Proceedings of the fourth Alvey vision conference, 1988.

5   **Davies E R,** Finding ellipses using the generalised Hough transform, *Pattern Recognition Letters,* 9, 87-96, 1989.

6   **Ayache N & Faugeras O D,** Building, registering, and fusing noisy visual maps, *International Journal of Robotics Research, special issue on sensor data fusion,* Vol 7, 45-66, 1988.

7   **Porrill J,** Optimal combination and constraints for geometrical sensor data, *International Journal of Robotics Research, special issue on sensor data fusion,* Vol 7, 66-77, 1988.

8   **Jazwinsci A H,** *Stochastic processes and filtering theory,* New York, Academic Press, 1970.

9   **Porrill J,** Reducing linearisation bias for the extended Kalman filter, *AIVRU Memo,* 1989.

10 **Bierman G J,** *Factorization methods for discrete sequential estimation,* New York, Academic Press, 1977.