Robust Estimation of Shape Parameters

G.A. Jones, J. Princen, J. Illingworth and J. Kittler

Dept. of Electronic and Electrical Engineering,
University of Surrey,
Guildford GU2 5XH, United Kingdom

We investigate the use of Robust Estimation in an application requiring the accurate location of the centres of circular objects in an image. A common approach used throughout computer vision for extracting shape information from a data set is to fit a feature model using the Least Squares method. The well known sensitivity of this method to outliers is traditionally accommodated by outlier rejection methods. These usually consist of heuristic applications of model templates or data trimming. Robust Estimation offers a theoretical framework for assessing such rejection schemes, and more importantly, provides an approach to parameter estimation in contaminated data distributions capable of greater accuracy.

Introduction

The detection and localisation of parametric shapes in edge maps is an important problem in computer vision and image analysis. It is difficult because edge maps contain a profusion of multiple structures and many contaminating features (see Figure 2). A common approach in such situations is to apply the Hough Transform (HT) method [1]. The HT detects structures by finding maxima in a parameter space defining particular instances of the structures. It copes well with multiple instances and contaminating feature points because the value of HT at a particular parameter is the number of points lying in a finite width template which surrounds the parametric curve. Hence the local peaks in the HT are not influenced by feature points which are far from the curve. This can be contrasted with a conventional least-squares method which attempts to minimise an error metric in which the values at a particular parameter value are influenced by all feature points. In typical estimation problems, straightforward least-squares is not particularly useful.

The performance of least-squares methods can be improved by changing the error metric so that points which lie far from the curve do not influence the final value - that is reject outliers. In this form the least-squares can be thought of as a method, like the HT, in which the value of the metric is a weighted sum of points falling inside a template which surrounds the curve. The template function though is quadratic rather than binary like the HT. In fact there are a whole class of methods which have been developed within the field of Robust Estimation, known as Generalised Maximum Likelihood (GML) methods, which can be thought of in this way. Methods which use finite, template like functions are known as redescending methods. As the template function is continuous, there are two important differences between redescending GML methods and the HT. Firstly, their properties such as accuracy and sensitivity to contaminations are well understood [2]. Secondly, continuity means that iterative methods can be employed to accurately identify extrema. This last property is particularly useful in problems with a large number of parameters. The above merits of Robust Estimation have not been generally exploited in computer vision applications [3].

In this paper we investigate the problem of circle centre estimation [4] using robust, redescending estimators. The problem is motivated by the task of accurately locating partial circular objects in low signal to noise ratio images. A HT is used to locate initial starting points for an iterative method which minimises an error metric defined by a redescending kernel. The iterative method solves a set of equations, which are obtained by taking the derivative of the error metric with respect to each of the unknown parameters, using a weighted least-squares formulation. One of the difficulties in using a fixed width template is that different SNR's will yield different edge point deviations around the ideal circular curve and as a result the performance of the method will differ. To avoid this problem we use a robust method to find an estimate of the standard deviation of edge points and use this estimate to iteratively define an appropriate kernel width. The inclusion of this scale estimation improves performance significantly in low SNR images.

Circle Finding

The application considered in this paper is the accurate, sub-pixel determination of the centre of partially occluded circular objects, see Figure 2. In a practical

image processing system, this problem is complicated by many factors associated with non-ideal imaging conditions. However, in this paper the problem is simplified by the use of synthetic data in order to gain a lower limit on the measurement accuracy attainable under "well understood" imaging conditions. The emphasis of the current work is to characterise the performance of different estimation procedures.

Image segmentation is often achieved by edge detection. However this is error prone and leads to the identification of non-boundary edge points throughout the image as well as imprecisely locating genuine boundary edge points. Figure 2 illustrates edge detection applied to a synthetic image of a single circular object. The difference between object and background gray levels was 100 and the image was corrupted by the addition of Gaussian distributed noise to each pixel.

In estimating the centre of a circular shape by applying a least squares fitting procedure, there is an implicit assumption that all the data considered must be Gaussianly distributed about the circular locus. However it can be seen in Figure 2 that not only is the circle contaminated by outliers across the image but there may also be other more structured distributions present i.e. the flat side of the object. Before applying a least squares procedure to such data these unwanted distributions must be removed. If initial estimates of the circle parameters are available then templates can be used to reject unwanted edge points. Wallace has implemented this technique using the Hough transform [6] to provide coarse initial estimates of circle parameters. The Hough Transform is a robust method for parameter estimation in noise but its accuracy depends on the quantization of the voting array and the accuracy of the edge information. Wallace detects candidate circles by finding peaks in a 3D Hough Transform space characterised by the parameters (a, b, r), where (a, b) is the circle centre and r is the circle radius. Points associated with each circle are isolated by defining a finite width binary template around the estimated parameters. Points that fall within the template are used for Least Squares fitting i.e. the template performs outlier rejection. However, for the method to work effectively the template width must be chosen carefully. Too small a template may cause valid points to be missed while too large a template will introduce outliers. The Least Squares fitting procedure is adversely affected by both of these effects.

Robust Estimation

As in the case of the Least Squared Error method, Robust Estimation methods require the minimization of an error criterion. This function K is composed of terms

from each of the n data points. It is given by

$$K = \sum_{n=1}^{N} \rho(e_n) \tag{1}$$

where $\rho(e_n)$ is a function defining the effect of errors. The error term is a function of the nth data point $\mathbf{x}_n = (x_n, y_n)$ and $\mathbf{a} = (a_0, a_1, .., a_j, ..)$ the parameters of the model function that is being fitted i.e.

$$e_n = f(x_n, y_n, \mathbf{a}) \tag{2}$$

In our circle estimation application the parameter vector is $\mathbf{a} = (a, b, r)$. The minimum of equation 1 will occur at the position $\mathbf{a}$ which is the best model fit to the data. This minimum is found by setting the partial derivatives of K with respect to each model parameter a_j to zero

$$\frac{\partial K}{\partial a_j} = \sum_{n=1}^{N} \phi(e_n) \frac{\partial e_n}{\partial a_j} = 0 \tag{3}$$

where $\phi(e_n) = \partial \rho(e_n)/\partial e_n$ is known as the *Influence Function* (IF). Solutions of this set of equations give possible values for $\mathbf{a}$.

In the least squares (or *mean*) formulation $\rho(e_n) = e_n^2$ and, consequently the influence function is given by $\phi(e_n) = \partial \rho(e_n)/\partial e_n = 2e_n$. This function, illustrated in Figure 3, assigns high weights to outliers and accounts for the sensitivity of the Least Square's method to noise. There is much statistical literature relating to the use of alternative IF's which are less sensitive or more robust in the presence of outliers. Examples, which are illustrated in Figure 3, include the *skipped mean*, the *skipped median* and the *Tukey Bi-weight* IF's.

It should be noted that a single application of the skipped mean to the data set is the same as the Least Squares fitting method with the use of a binary template. The cut-off parameter τ plays the same role as the choice of template width. However, assuming a Gaussian distribution with small levels of contamination, choice of the cut off point can be based on results from the robust Estimation literature. The least sensitive i.e. most robust estimator, of scale, s, or spread of a contaminated Gaussian distribution is the Median Absolute Deviation (MAD) estimator [2], viz

$$s = 1.4826 median(|e_n|) \; ; \; \forall n \tag{4}$$

The factor of 1.4826 is the ratio of the standard deviation to the median of absolute deviations from the mean of an uncontaminated Gaussian distribution. Having estimated the scale of the data distribution all measurements and their errors can be suitably adjusted.

$$K = \sum_{n} \rho(e_n/s) \tag{5}$$

A consequence of this is that the value of τ need only be expressed as a multiple of the scale. In this way the

fraction of data points within the fit can be set with reference to a Gaussian distribution i.e. to accept 98% of all good points set the cut-off or template width to $\tau = 3$.

The Circle Finding Problem

In this section we formulate the circle finding problem in terms of Robust Estimation. We consider a two stage method in which the first stage is a Hough based estimate of parameters. The second stage involves an iterative refinement of this estimate by solving equation 3. The error function in equation 2 for a point (x_n, y_n) is defined as

$$e_n = (x_n - a)^2 + (y_n - b)^2 - r^2 \qquad (6)$$

This can be shown to be approximately equivalent to $e_n = 2r\delta r_n$ where δr_n is the perpendicular distance between the data point (x_n, y_n) and the circle locus defined by the estimate (a, b, r). In order to simplify the solution of equation 3, equation 2 is rewritten as

$$e_n = 2ax_n + 2by_n + 2c - 2z_n \qquad (7)$$

where $c = (r^2 - a^2 - b^2)/2$ and $z_n = (x_n^2 + y_n^2)/2$. The solution vector $(\hat{a}, \hat{b}, \hat{c})$ is found by solving the following equations, derived from equation 3

$$\frac{\partial K}{\partial a} = \sum_{n=1}^{N} \frac{\partial e_n}{\partial a} \phi(e_n/s) = 0 \qquad (8)$$

$$\frac{\partial K}{\partial b} = \sum_{n=1}^{N} \frac{\partial e_n}{\partial b} \phi(e_n/s) = 0 \qquad (9)$$

$$\frac{\partial K}{\partial c} = \sum_{n=1}^{N} \frac{\partial e_n}{\partial c} \phi(e_n/s) = 0 \qquad (10)$$

where

$$\frac{\partial e_n}{\partial a} = 2x_n, \quad \frac{\partial e_n}{\partial b} = 2y_n, \quad \frac{\partial e_n}{\partial c} = 2 \qquad (11)$$

The above equations can be solved using the iterative Weighted Least Squares formulation. At the ith iteration, an influence weight for each point is defined as $w_n = \phi(e_{n,i-1}/s_i)/(e_{n,i-1}/s_i)$. $e_{n,i-1}$ is the error between the nth point and the parameter estimate at i^{th} iteration. Equations 8, 9 and 10 can then be expressed as

$$\sum_{n=1}^{N} x_n w_n (a_i x_n + b_i y_n + c_i - z_n) = 0$$

$$\sum_{n=1}^{N} y_n w_n (a_i x_n + b_i y_n + c_i - z_n) = 0 \qquad (12)$$

$$\sum_{n=1}^{N} w_n (a_i x_n + b_i y_n + c_i - z_n) = 0$$

The intermediate parameter vector $\mathbf{p}_i = (a_i, b_i, c_i)$ at each iteration can be computed by rearranging equation 12 into the matrix form below and finding the inverse.

$$A_i \mathbf{p}_i = \mathbf{b}_i \qquad (13)$$

$$A_i = \begin{bmatrix} \sum_1^N w_n x_n^2 & \sum_1^N w_n x_n y_n & \sum_1^N w_n x_n \\ \sum_1^N w_n x_n y_n & \sum_1^N w_n y_n^2 & \sum_1^N w_n y_n \\ \sum_1^N w_n x_n & \sum_1^N w_n y_n & \sum_1^N w_n \end{bmatrix}$$

$$\mathbf{b}_i = \begin{bmatrix} \sum_{n=1}^N w_n x_n z_n \\ \sum_{n=1}^N w_n y_n z_n \\ \sum_{n=1}^N w_n z_n \end{bmatrix}$$

Finally, the circle parameters $(\hat{a}, \hat{b}, \hat{r})$ can then be recovered using the conversion $\hat{r}^2 = 2\hat{c} + \hat{a}^2 + \hat{b}^2$.

The Iteration Scheme

An iterative scheme is used to minimise the error metric of equation 5. The estimate of a prior iteration is used to recompute the weights of each data point. The initial estimate $\mathbf{p}_0$ is found via equation 7 from the circle parameters (a, b, r) recovered by the first Hough Transform stage. Thus, the basic iterative cycle of the second stage can be described by the following steps.

1. **Compute Scale** : Compute the scale s_i from equation 4.

2. **Compute Weights** : For each data point (x_n, y_n) the weight w_n is found from $w_n = \phi(e_{n,i-1}/s_i)/(e_{n,i-1}/s_i)$.

3. **Compute Matrices** : Construct the matrices A_i and $\mathbf{b}_i$.

4. **Find Solution** : Find $\mathbf{p}_i$ by inverting matrix A_i.

5. **Compute Error Metric** : Compute equation 5.

6. **Check for Termination** : Terminate algorithm or repeat above steps.

As the iteration scheme progresses, K should converge to its minimum value K_{min}. Termination follows when K has been *judged* to have converged to K_{min}. K may only decrease and termination occurs when the decrease in K becomes small *i.e.* when $K_i > \alpha K_{i-1}$ where $\alpha = 0.999$.

At each iteration the scale of the data points is computed. The correct estimation of scale is important as it determines the width of the influence functions and

therefore outlier rejection. Unfortunately the robust MAD scale estimator is still sensitive to large numbers of outliers. It is therefore necessary to base scale estimation on a 'trimmed' sample of data points. This can be reliably achieved by using a trimmed median Influence Function with a fairly large cut-off so that all true data points are contained within it, but most outlying data points are excluded. Since $e_n \approx 2r_i\delta r_n$, the cut-off is chosen as $\tau = 2r_i\tau_r$ where τ_r is set at some value between 8 and 15 pixels. For the data presented here the result is insensitive to the choice of τ_r.

Experimental Results

In this section results are presented for the performance of the iterative robust estimation method and the more heuristic non-iterative Least Squares method of Wallace, which uses a fixed size template. Synthetic data provides a quantitative comparison of the best achievable accuracy as a function of additive Gaussian gray level noise. The algorithms are then applied to a real data image to illustrate the benefit of using the robust estimation procedure.

All experiments with the robust method used the smoothly redescending Tukey Bi-weight Influence Function. For the method of Wallace a fixed template width of ± 3.5 pixels was adopted. The edge map of the synthetic data used is shown in Figure 2. The original image contained a circle of radius 90 pixels with 20% of its boundary replaced by a linear segment. The image size is 256×256 pixels. Figure 4 compares the location accuracy of the first stage Hough Transform, the Least Squares algorithm and the Robust Estimation estimator as a function of the standard deviation of Gaussian noise added to each pixel. As the noise increases so does the positional inaccuracy of each pixel as well as the number of spurious points detected by the edge segmentation. Therefore, it is expected that the accuracy of circle parameters should decrease. Each data point is the result of 300 trials during which the orientation of the deformed circle and sub-pixel centre position were varied. The vertical axis plots the error between estimated and actual centre positions. Curves are drawn such that 98% of all the measured values deviate from the true centre by less than the plotted values. It can be seen that the Hough Transform cannot reliably achieve sub-pixel accuracy. The Least Squares method performs well for low noise levels but breaks down for moderate levels of noise. The Robust Estimation method extends the range of tolerable noise for which sub-pixel accuracy is attainable to twice that of the Least Square method. This improvement is gained at the computational expense of iterating the estimation scheme. Table 1 shows the average number of iterations for each noise level. However the rate determining step of the total algorithm is the first stage Hough Transform which takes about 100 times the time of the Least Squares method.

Some of the advantages of the Robust Estimation method over the Least Square method are illustrated with respect to real data in Figure 5. This shows the edge map of a complicated scene containing two man-made objects. The problem is to accurately estimate the parameters of the large circular arc which is part of the outer boundary of the object on the right. The algorithms are manually given an initial estimate, and the results are shown by overlaying on the edge map the circle corresponding to the finally estimated parameters. Figure 5(a) illustrates the result of giving an accurate initial estimate to the Least Squares with a template of width 16 pixels. The result is corrupted by other data structures in this wide template. This effect is reduced using a smaller template. Figure 5(b) presents the result produced by the Robust method for the same initial estimate. After several iterations it converges to the correct arc parameters.

Figures 5(c) and (d) shows the effect of introducing inaccuracy into the initial estimate of circle parameters. In this example the initial estimate is out by two pixels in each of the three parameters. The Least Squares (c) is even more distorted by the surrounding structure than previously. Reducing the template width can result in complete failure to find any sensible solution. For the same inaccuracy, Figure 5(d) shows that the Robust Estimation scheme (which calculates its own scale estimate) manages to lock onto the correct parameters.

Number of Iterations					
		Noise Levels			
σ	0.0	12.5	25.0	37.5	50.0
No. Iterations	7	10	15	16	18

Figure 1: Number of Iterations to Convergence

Conclusions

We have outlined the relationship between a general class of robust estimation techniques and the heuristic outlier rejection methods which have been applied to the standard Least Squares. Robust statistics provides a framework for analysing the properties of these methods. An algorithm which applies a robust estimation technique to the problem of circle parameter estimation has been developed. The method utilizes scale estimation to iteratively select an appropriate template width, as a function of the dispersion of edge points from the circular boundary. The range of signal to noise ratio's with which the algorithm can cope is considerably greater than that of the standard least-squares with outlier rejection. The performance of the robust estimator degrades gracefully as the noise level increases. Experimental results using

both synthetic and real data illustrate the robustness of the method.

Robust estimation methods coupled with some fairly robust, though inaccurate initial stage, such as the HT, provide a useful approach to parameter estimation problems. These methods have wide application to problems in computer vision and image analysis.

References

[1] **Yuen H.K., Princen J., Illingworth J., Kittler J.**, "Comparative Study of Hough Transform methods for Circle Finding", Image and Vision Computing, Vol. 8, No. 1, pp 71-77, February 1989

[2] **Hampel F., Ronchetti E., Rousseeuw P., Stahel W.**, "Robust Statistics: The Approach Based on Influence Functions" John Wiley & Sons 1986

[3] **Besl P.J., Birch J.B., Watson L.T.**, "Robust Window Operators", Proc. IEEE 2nd ICCV, Florida December 1988, pp 591-600

[4] **Jones G.A., Yuen H.K., Princen J., Illingworth J., Kittler J.**, "Accurate Estimation of Circle Centres", VSSP-TR9/90, Dept. Elect. Eng., University of Surrey 1990

[5] **Canny J.**, "A Computational Approach to Edge Detection", IEEE Trans. Pattern Analysis and Machine Intelligence, Vol. PAMI-8 pp 679, November 1986

[6] **Wallace A.M.**, "Greyscale Image Processing for Industrial Applications", Image and Vision Computing, Vol. 1, No. 4 November 1983
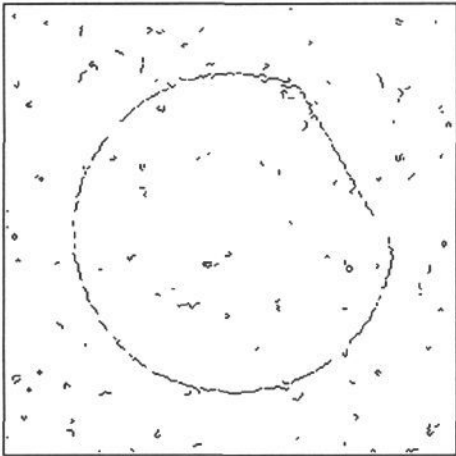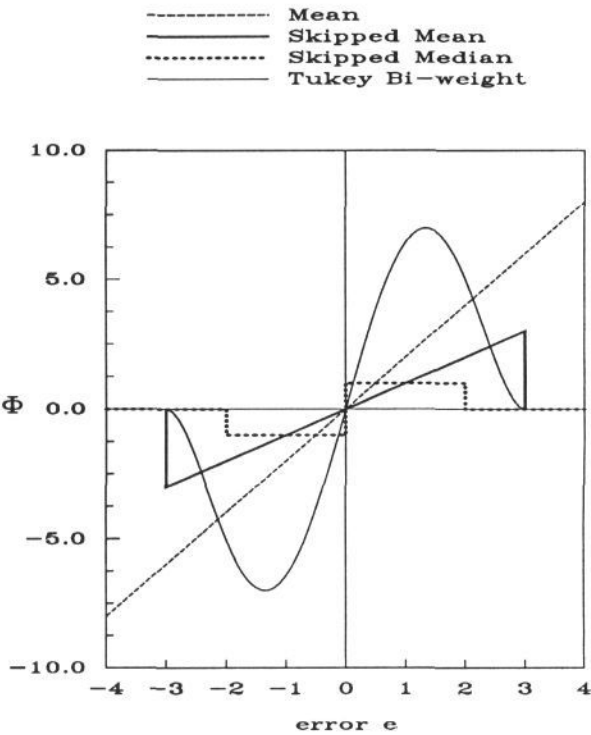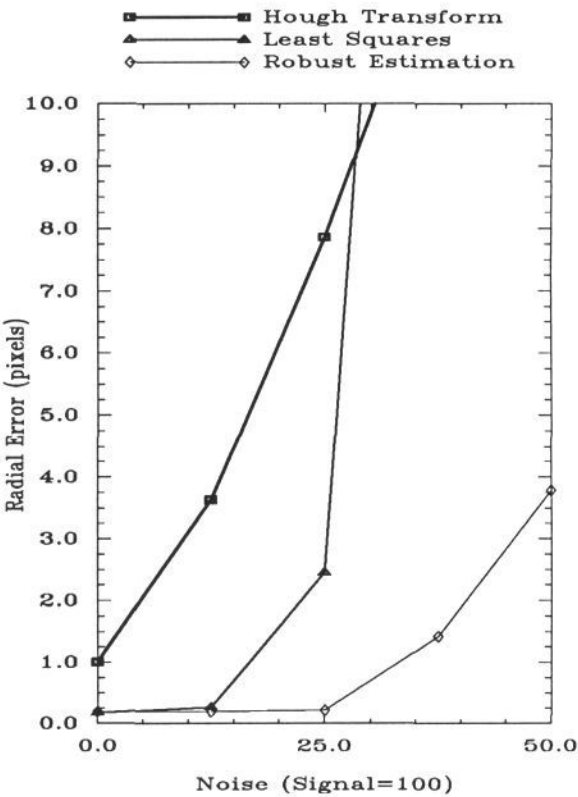
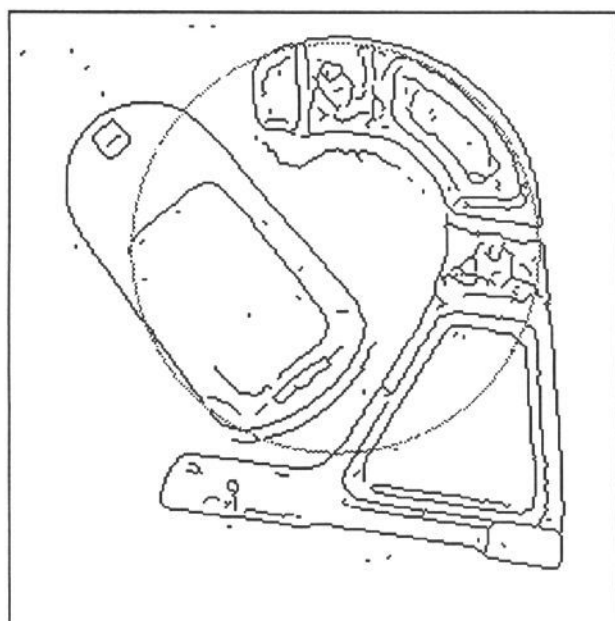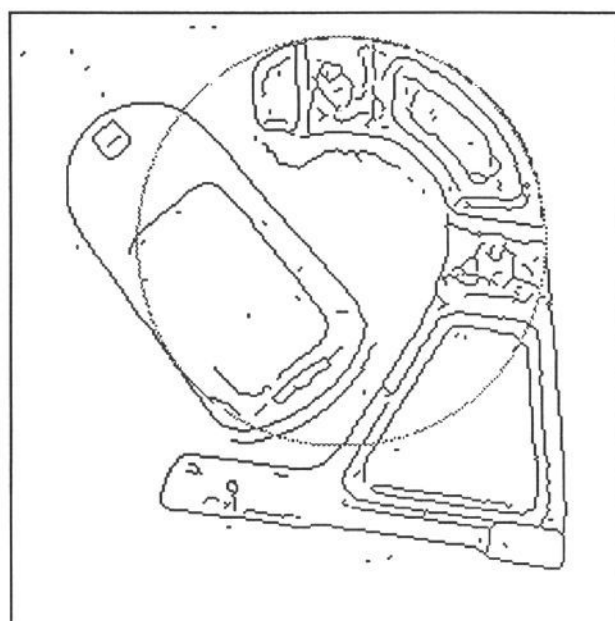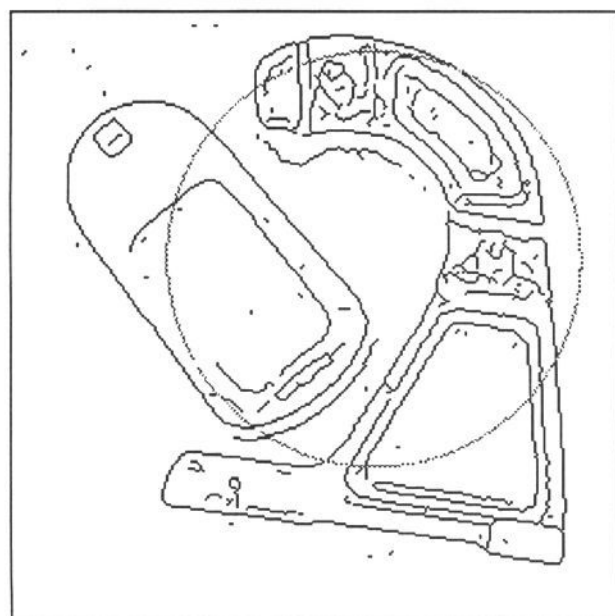Figure 3: Influence Functions

Figure 2: Circular Object

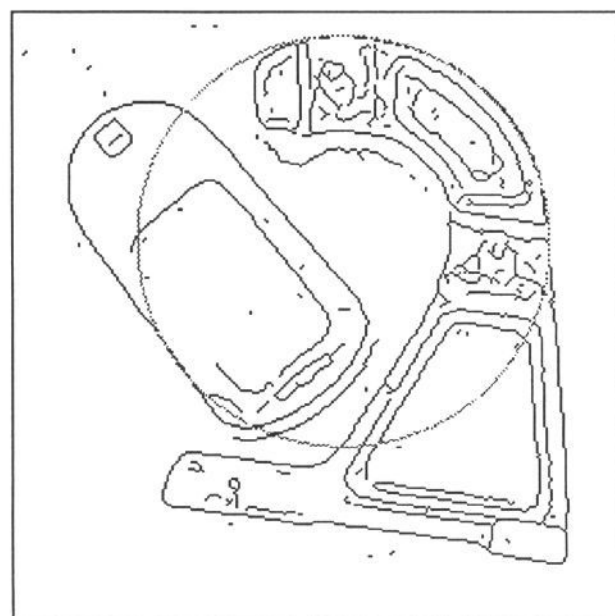Figure 4: Centre Location error as a function of noise

(a) Least Squares : Template = 16
Using accurate initial estimate

(b) Robust Estimator
Using accurate initial estimate

(c) Least Squares : Template = 16
Using inaccurate initial estimate

(d) Robust Estimator
Using inaccurate initial estimate

Figure 5: Comparison of Algorithms