APPROXIMATE MAXIMUM-LIKELIHOOD ESTIMATION OF CIRCLE PARAMETERS USING A PHASE-CODED KERNEL

Emanuel E. Zelniker[†] *and I. Vaughan L. Clarkson*

Intelligent Real-Time Imaging and Sensing Group
School of Information Technology & Electrical Engineering
University of Queensland, Queensland 4072, AUSTRALIA
phone: +61 7 3365 4510, fax: +61 7 3365 4999, email: {zelniker,v.clarkson}@itee.uq.edu.au
web: http://www.itee.uq.edu.au/∼ zelniker

ABSTRACT

The accurate fitting of a circle to noisy measurements of points on its circumference is an important and much-studied problem in statistics. ATHERTON & KERBYSON (*Image and Vision Computing* 17, 1999, 795-803) have proposed a complex convolutional circle parameter estimator. One of the estimators proposed is a 'Phase-Coded Annulus' to estimate for the centre and radius. ZELNIKER & CLARKSON (*Digital Image Computing: Tech. and Appl.* 2003, 509-518) have shown that it is possible to exactly describe the Maximum Likelihood Estimator (MLE) in terms of convolution under a certain model for ideal images formed from noisy circle points. In this paper, we investigate the relationship between the convolution of an ideal image with a Phase-Coded Kernel and the MLE. We compare our approximate MLE (AMLE) method to the DELOGNE-KÅSA Estimator which uses a least squares approach to solve for the circle parameters, the MLE as well as the theoretical CRAMÉR-RAO Lower Bound.

1. INTRODUCTION

The accurate fitting of a circle to noisy measurements of circumferential points has applications in many areas of research including archæology [1], geodesy [2], microwave engineering [3] and computer vision and metrology [4]. Several methods can be used to estimate the circle parameters, for example, the Circular Hough Transform (CHT) [5, 6], Maximum Likelihood Estimation (MLE) [7] and the DELOGNE-KÅSA Estimator (DKE) [3, 8].

The CHT is an iterative approach to circle parameter estimation in images. Since a circle is characterised by 3 parameters, *i.e.*, the two centre coordinates and the radius, a 3-dimensional Hough accumulator space is chosen where every point in this space represents a circle of a certain size in x-y space. Then, for every point in x-y space, we check for all possible circles which pass through this point and increment the corresponding Hough space coordinates (which represent these circles) once. Depending on how noisy the image is, there will be one Hough space coordinate which will be incremented N times, where N is the number of points on the circle's circumference. This Hough space coordinate represents the circle in the image. Other Hough space coordinates will be incremented zero times, once or possibly a few times. Therefore, the Hough space can be viewed as an intensity space and the coordinate with the highest intensity corresponds to the circle estimates.

[†] Emanuel Zelniker is additionally supported by a scholarship from the Commonwealth Scientific & Industrial Research Organisation.

Computation of the CHT is very slow and memory intensive and as a result, better formulations have been proposed which look for a range of circle sizes in one go. ATHERTON & KERBYSON [9] mention that the Hough transform can be implemented by convolving a Single Circle with an edge magnitude image (matched filtering). They build on this idea by defining an Orientation Annulus which detects a range of radii of circles, but also uses edge orientation information by taking the dot-product between the image edge orientation and an orientation field within the annulus. A complex Phase-Coded Annulus (PCA) is also described which detects a range of radii of circles by using phase to code for radius. The complex vector convolution of this annulus with an edge magnitude image results in a circle detection operation which uses both edge orientation information and size information.

The above-mentioned techniques are all designed to operate on digital images. ZELNIKER & CLARKSON [10] show that it is possible to exactly implement the MLE as well as the DKE in terms of convolution under a certain model for an ideal image, which they define as an unbounded image with continuous-values in intensity and in spatial coordinates. Their technique can also be adapted to digital images to produce necessarily coarser estimates.

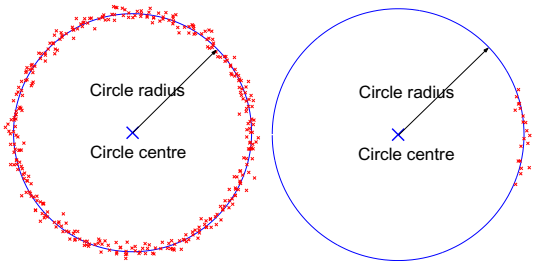
In this paper, we show that an approximate MLE (AMLE) can be developed for these same ideal images. This AMLE can also be calculated by convolution and the kernel turns out to be a version of the PCA.

The paper reads as follows. In Section 2, we provide a brief overview of CHAN's Circular Functional Model. An ideal image of noisy circumferential points is defined. The CRAMÉR-RAO Lower Bound (CRLB) for the circle-centre estimates and radius estimate is discussed. We revise the objective function of the MLE and show how the objective function can be implemented via convolution. In Section 3, we define the objective function of the AMLE, we show the relationship between the objective function of the AMLE and the objective function of the MLE and discuss applications to real images. In Section 4, we present simulation results to compare the AMLE, MLE and DKE to the CRLB.

2. BACKGROUND

2.1 Chan's Circular Functional Model

In this Section, we briefly present CHAN's circular functional model [11]. In this model, we assume that the positions of N points on the circumference of a circle are measured. The measurement process introduces random errors so that the Cartesian coordinates (x_i, y_i), $i = 1, \ldots, N$ can be expressed



(a) Circle (b) Arc

Figure 1: Noisy measurements of points on the circumference of a circle and an arc.

as

$$x_i = a + r\cos\theta_i + \xi_i,$$
$$y_i = b + r\sin\theta_i + \eta_i.$$

Here, $\mathbf{c} = (a,b)$ is the centre of the circle, r is its radius, the θ_i are the angles around the circumference on which the points lie and the ξ_i and η_i are instances of random variables representing the measurement error. They are assumed to be zero-mean and i.i.d. In addition, we will specify that they are Gaussian with variance σ^2.

Figure 1 shows some data with N points for the circumference of a circular arc, $\mathbf{p}_1 = (x_1,y_1),\ldots,\mathbf{p}_N = (x_N,y_N)$, displaced from the circumference by noise.

2.2 Ideal Images of Noisy Circumferential Points

In [10], an ideal image of the noisy circular points is defined to be an image which is unbounded and is continuous-valued in intensity and in spatial coordinates. Under these conditions, it can be assumed that the points are represented by 2-dimensional delta functions, that is, it is possible to consider the image as a function $f(a,b)$, where

$$f(a,b) = \sum_{i=1}^{N} \delta(a - x_i, b - y_i). \tag{1}$$

An ideal image in some cases can be approximated by an edge magnitude image.

2.3 Cramér-Rao Lower Bound (CRLB)

The CRLB (see [12]) for CHAN's circular functional model with Gaussian random variables was derived by CHAN & THOMAS [13], but see also [14] for a more straightforward derivation. For an unbiased estimator of circle centre (a,b), radius r and angles $\theta_1,\ldots,\theta_N$, we find that the variances for the estimates of $a,b,r,\theta_1,\ldots,\theta_N$ must not be less than the diagonal elements of the inverse of the FISHER Information Matrix, $\mathbf{J}$, where

$$\mathbf{J} = \begin{pmatrix} \mathbf{J}_{11} & \mathbf{J}_{12} \\ \mathbf{J}_{21} & \mathbf{J}_{22} \end{pmatrix} \tag{2}$$

and the 4 sub-matrices of $\mathbf{J}$ are as follows

$$\mathbf{J}_{11} = \frac{1}{\sigma^2}\sum_{i=1}^{N}\begin{pmatrix} 1 & 0 & \cos\theta_i \\ 0 & 1 & \sin\theta_i \\ \cos\theta_i & \sin\theta_i & 1 \end{pmatrix}, \tag{3}$$
$$\mathbf{J}_{12} = \frac{1}{\sigma^2}\begin{pmatrix} -r\sin\theta_1 & \cdots & -r\sin\theta_N \\ r\cos\theta_1 & \cdots & r\cos\theta_N \\ 0 & \cdots & 0 \end{pmatrix},$$
$$\mathbf{J}_{21} = \mathbf{J}_{12}^T,$$
$$\mathbf{J}_{22} = \frac{r^2}{\sigma^2}\mathbf{I}_{N\times N}.$$

From (3), it can be seen that $\mathbf{J}$ is an $(N+3)\times(N+3)$ symmetric matrix. Since we are interested in the CRLBs of a, b and r, via the block matrix inversion lemma, these will lie along the diagonal of the upper 3×3 sub-matrix of $\mathbf{J}^{-1}$, which is the inverse of

$$\mathbf{J}_{11} - \mathbf{J}_{12}\mathbf{J}_{22}^{-1}\mathbf{J}_{21} = \frac{1}{\sigma^2}\sum_{i=1}^{N}\begin{pmatrix} \cos^2\theta_i & \cos\theta_i\sin\theta_i & \cos\theta_i \\ \cos\theta_i\sin\theta_i & \sin^2\theta_i & \sin\theta_i \\ \cos\theta_i & \sin\theta_i & 1 \end{pmatrix}. \tag{4}$$

2.4 Maximum Likelihood Estimation (MLE)

The conditional probability density function for $\mathbf{p}_1,\ldots,\mathbf{p}_N$ is as follows

$$P(\mathbf{p}_1,\ldots,\mathbf{p}_N \mid a,b,r,\theta_1,\ldots,\theta_N) = \frac{1}{(2\pi\sigma^2)^N}\prod_{i=1}^{N}\exp\left(-\frac{\xi_i^2+\eta_i^2}{2\sigma^2}\right). \tag{5}$$

By definition, maximum likelihood estimates are given by those $\hat{a},\hat{b},\hat{r},\hat{\theta}_1,\ldots,\hat{\theta}_N$ that maximise (5).

In [10], the logarithm of (5) is considered to obtain an objective function, $f_{\text{obj}}(a,b,r,\theta_1,\ldots,\theta_N \mid \mathbf{p}_1,\ldots,\mathbf{p}_N)$, which is related to the log-likelihood by a scaling factor and constant offset, both of which depend only on σ and N. A partial derivative with respect to θ_i firstly and then with respect to r shows that the objective function is maximised when

$$\hat{\theta}_i = \arctan\left(\frac{y_i - b}{x_i - a}\right),$$
$$\hat{r} = \frac{1}{N}\sum_{i=1}^{N}\|\mathbf{p}_i - \mathbf{c}\|_2.$$

It is then possible to express the objective function as follows

$$f_{\text{obj}}(a,b \mid \mathbf{p}_1,\ldots,\mathbf{p}_N) = -\sum_{i=1}^{N}\|\mathbf{p}_i - \mathbf{c}\|_2^2 + \frac{1}{N}\left[\sum_{i=1}^{N}\|\mathbf{p}_i - \mathbf{c}\|_2\right]^2. \tag{6}$$

2.5 MLE via Convolution

One of the main results in [10] is to show how (6) can be exactly implemented/interpreted by convolving an ideal image (1) with a conic kernel $g(a,b) = \sqrt{a^2+b^2}$ as follows

$$-f_{\text{obj}}(a,b \mid \mathbf{p}_1,\ldots,\mathbf{p}_N) = f(a,b) * g^2(a,b) - \frac{(f(a,b)*g(a,b))^2}{f(a,b)*g^0(a,b)}. \tag{7}$$

The MLE is therefore shown to be equivalent to minimising the intensity of an ideal image obtained through convolution.

In Section 3, we show how the convolution of an ideal image (1) with a version of ATHERTON & KERBYSON's PCA is approximately the MLE (AMLE), *i.e.*, this convolution is approximately (6).

3. AN APPROXIMATE MLE (AMLE)

Consider the function

$$\Theta(a,b) = e^{j\sqrt{a^2+b^2}}. \tag{8}$$

This is a complex kernel that uses phase to code for radius and can be regarded as another member of a class of kernels which ATHERTON & KERBYSON call 'Phase-Coded Annulus'. When this kernel is convolved with an ideal image, a complex output results whose magnitude can be interpreted as a 2-dimensional intensity image where the coordinate of the maximum is the circle-centre estimates.

We now define a new objective function, $h_{\text{obj}}(a,b)$ by taking the squared magnitude of the convolution of (1) and (8), *i.e.*

$$\begin{aligned} &h_{\text{obj}}(a,b) \\ &= |f(a,b) * \Theta(a,b)|^2 \qquad (9) \\ &= \left(\sum_{i=1}^{N} \cos(r+\Delta r_i)\right)^2 + \left(\sum_{i=1}^{N} \sin(r+\Delta r_i)\right)^2, \qquad (10) \end{aligned}$$

where

$$\begin{aligned} r_i &= \sqrt{(a-x_i)^2+(b-y_i)^2} \\ &= \|\mathbf{p}_i - \mathbf{c}\|_2, \end{aligned}$$

and

$$\begin{aligned} \Delta r_i &= r_i - r \\ &= \|\mathbf{p}_i - \mathbf{c}\|_2 - r. \end{aligned}$$

The TAYLOR-series approximation of the *cosine* and *sine* terms in (10) neglecting third and higher order terms is

$$\begin{aligned} \cos(r+\Delta r_i) &= \cos r - \Delta r_i \sin r - \frac{(\Delta r_i)^2}{2}\cos r, \qquad (11) \\ \sin(r+\Delta r_i) &= \sin r + \Delta r_i \cos r - \frac{(\Delta r_i)^2}{2}\sin r. \end{aligned}$$

Substituting (11) into (10) and simplifying, it is possible to show that the objective function (9) becomes

$$h_{\text{obj}}(a,b) \approx N^2 - N\sum_{i=1}^{N} \|\mathbf{p}_i - \mathbf{c}\|_2^2 + \left[\sum_{i=1}^{N} \|\mathbf{p}_i - \mathbf{c}\|_2\right]^2. \tag{12}$$

The expression in (12) is the same as the expression in (6) which was derived in [10] but with a constant scaling and offset factors. Therefore, we have demonstrated an approximate relationship between the MLE (6) and the convolution of (1) with (8) to get (12). The AMLE is therefore equivalent to maximising the intensity of the magnitude of an ideal complex image obtained through convolution.

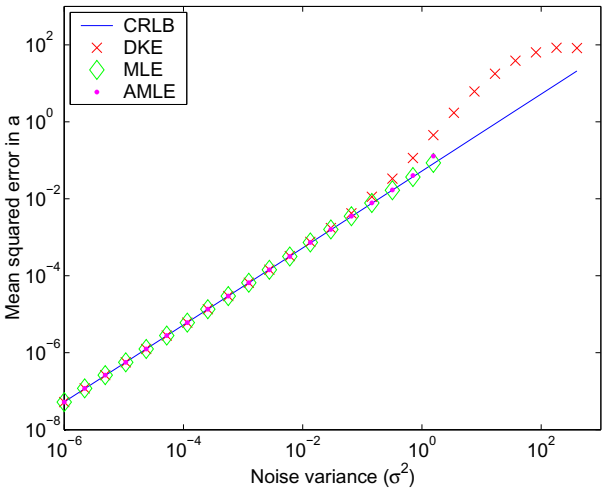


Figure 2: Simulation results for varying σ for an arc length of 180°.

We note that the kernel in (8) is a version of the PCA in [9]. ATHERTON & KERBYSON propose a complex kernel that uses phase to code for a range of radii, $R_{\min}$ to $R_{\max}$ as follows

$$O_{\text{PCA}}(a,b) = \begin{cases} e^{j\phi_{a,b}} & \text{if } R_{\min}^2 < a^2+b^2 < R_{\max}^2, \\ 0 & \text{otherwise,} \end{cases}$$

where

$$\phi_{a,b} = 2\pi\left(\frac{\sqrt{a^2+b^2}-R_{\min}}{R_{\max}-R_{\min}}\right).$$

We propose that this kernel should only be applied to real images via convolution in order to implement the AMLE when the size of the circle is known to be within certain bounds. The size of the circle is then represented by the phase of the complex convolution. Furthermore, because phase is used to code for radius (8), circle points are in phase while non-circle points are out of phase, meaning that the AMLE can be implemented on actual images which have many edge points that do not belong to the circle/s in them.

4. RESULTS FROM SIMULATIONS

Two simulations were performed, one with fixed σ and the other with fixed N. In the first simulation, The AMLE was simulated using a Monte-Carlo analysis. In each trial, 200 noisy points ($N = 200$) were generated in equal increments around half a circle's circumference. The radius was set to 16. Noise was then added to each (x_i, y_i) coordinate pair in the form of (ξ_i, η_i). The amount of noise, σ was varied from 10^{-3} to 20 in equal geometric increments. The AMLE was implemented via the Newton-Raphson (NR) algorithm, so a coarse grid search was performed in the neighbourhood of the true centre prior to NR in order to determine where to initiate the NR algorithm. The AMLE was run 10 000 times for each value of σ to obtain estimates for the centre of the circle $(\hat{a}, \hat{b})$ and radius $\hat{r}$, and so generate mean square error (MSE) values. The same was done for the DKE (see [15]) and the MLE which was also implemented via the Newton-

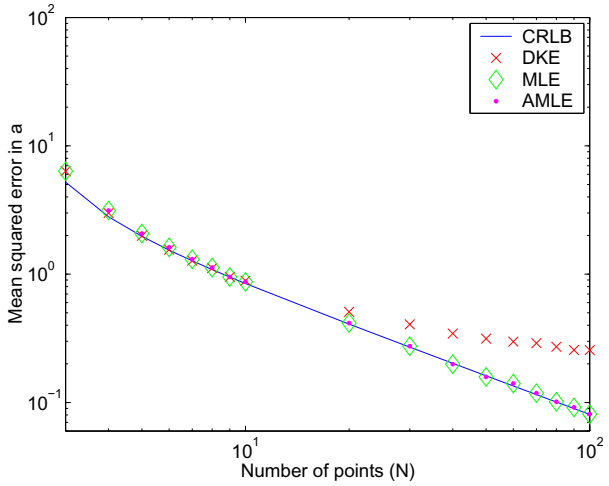


Figure 3: Simulation results for varying N for an arc length of 90°.

Raphson algorithm, so a coarse grid search was performed to determine where to initiate the NR algorithm for the MLE.

The MSE values in $\hat{a}$ for the AMLE are plotted against their corresponding theoretical CRLB for the same level of noise σ in Figure 2 on a logarithmic scale. It can be seen that as the noise level, σ, approaches zero, the estimator $\hat{a}$ approaches the theoretical CRLB. The same is observed for the MLE and DKE. For the MLE and AMLE, the MSE is very large for high values of σ. This is because the Newton-Raphson algorithm diverges in some cases, and so the MSE cannot be plotted. We would like to point out though that the MLE and AMLE conform to the CRLB for longer than the DKE as σ increases. We have chosen to illustrate the results for $\hat{a}$. Plots for $\hat{b}$ and $\hat{r}$ follow a similar pattern.

In the second set of simulations, σ was fixed and N was varied. The number of points were varied from $N = 3$ to $N = 100$ and they were generated in equal increments around a quarter of a circle's circumference. The radius was set to 16 and the noise (ξ_i, η_i) added to each (x_i, y_i) coordinate pair was set to 0.25 ($\sigma = 0.25$). The AMLE was run 10 000 times for each value of N to obtain estimates for the centre of the circle $(\hat{a}, \hat{b})$ and radius $\hat{r}$, and so generate MSE values.

The MSE values in $\hat{a}$ for the AMLE, MLE and DKE are plotted against the corresponding theoretical CRLB for each value of N in Figure 3 on a logarithmic scale. It can be seen that, unlike the DKE, the MLE and AMLE do not diverge from the CRLB as N increases. The DKE is not asymptotically efficient [7, 15] and is known to diverge from the CRLB as N approaches infinity. For the AMLE, the MSE is very large for $N = 3$ and so the MSE cannot be plotted.

5. CONCLUSIONS

We have shown that the convolution of an ideal image with a Phase-Coded Kernel is approximately the MLE (AMLE). Furthermore, because this kernel is complex and uses phase to code for radius, circle points are in phase while non-circle points are out of phase, meaning that the AMLE can be implemented on actual images which have many edge points that do not belong to the circle/s in them.

When applying to actual images, the size of the circle is usually known and so the complex PCA should use phase to code for a range of radii, $R_{\min}$ to $R_{\max}$ before convolving it with the image in order to implement the AMLE.

Our simulations showed that the AMLE, MLE and DKE quickly approached the CRLB as the noise variance approached zero, suggesting that all methods are statistically efficient.

REFERENCES

[1] A. Thom, "A Statistical Examination of the Megalithic Sites in Britain," *J. Roy. Statist. Soc. Ser. A General*, vol. 118, pp. 275–295, Mar. 1955.

[2] S. M. Robinson, "Fitting Spheres by the Method of Least Squares," in *Comm. ACM*, 1961, p. 491.

[3] I. Kåsa, "A circle fitting procedure and its error analysis," *IEEE Trans. Instrum. Meas.*, vol. 25, pp. 8–14, Mar. 1976.

[4] U. M. Landau, "Esimation of a Circular Arc Center and its Radius," *Comput. Vision Graph. Image Process.*, vol. 38, pp. 317–326, June 1986.

[5] R. O. Duda and P. E. Hart, "Use of the Hough Transform to Detect Lines and Curves in Pictures," *Comm. ACM*, vol. 15, pp. 11–15, 1972.

[6] E. Trucco and A. Verri, *Introductory Techniques for 3D Computer Vision*. New Jersey: Prentice Hall, 1998.

[7] M. Berman and D. Culpin, "The Statistical Behaviour of Some Least Squares Estimators of the Centre and Radius of a Circle," *J. Roy. Statist. Soc. Ser. B Stat. Methodol.*, vol. 48, pp. 183–196, Oct. 1986.

[8] P. Delogne, "Computer Optimization of Deschamps' method and error cancellation in reflectometry," in *Proc. IMEKO-Symp. Microwave Measurements*, Budapest, Hungary, 1972, pp. 117–123.

[9] T. J. Atherton and D. J. Kerbyson, "Size invariant circle detection," *Image and Vision Computing*, vol. 17, pp. 795–803, Aug. 1999.

[10] E. E. Zelniker and I. V. L. Clarkson, "Maximum-Likelihood Circle-Parameter Estimation via Convolution," in *Digital Image Computing: Tech. and Appl.*, Sydney, Australia, Dec. 2003, pp. 509–518.

[11] N. N. Chan, "On Circular Functional Relationships," *J. Roy. Statist. Soc. Ser. B Stat. Methodol.*, vol. 27, pp. 45–56, Apr. 1965.

[12] H. L. Van Trees, *Detection, Estimation and Modulation Theory, Part 1*. New York: Wiley, 1968.

[13] Y. T. Chan and S. M. Thomas, "Cramér-Rao Lower Bound for Estimation of a Circular Arc Center and its Radius," in *Graph. Models Image Process.*, Nov. 1995, pp. 527–532.

[14] K. Kanatani, "Cramér-Rao Lower Bounds for Curve Fitting," in *Graph. Models Image Process.*, 1998, pp. 93–99.

[15] E. E. Zelniker and I. V. L. Clarkson, "A Statistical Analysis of the DELOGNE-KÅSA Method for Fitting Circles," June 2003, submitted to *IEEE Trans. Signal Processing*.