

OPTIMAL CIRCLE FITTING VIA BRANCH AND BOUND

Emanuel E. Zelniker, Benjamin C. Appleton and I. Vaughan L. Clarkson

School of Information Technology and Electrical Engineering
The University of Queensland, Queensland, 4072, Australia.

ABSTRACT

In this paper, we examine the problem of fitting a circle to a set of noisy measurements of points from the circle's circumference, assuming independent, identically distributed GAUSSIAN measurement errors. We propose an algorithm based on Branch and Bound to obtain the Maximum Likelihood Estimate and show that this algorithm obtains the optimal estimate. We examine the rate of convergence and determine the computational complexity of the proposed algorithm. We also provide timings and compare to existing techniques for circle fitting proposed in the literature. Finally, we demonstrate that our algorithm is statistically efficient by comparing our results to the CRAMÉR-RAO lower bound.

1. INTRODUCTION

The accurate fitting of a circle to noisy measurements of points on its circumference is a much-studied problem in the scientific literature. In his paper, CHAN [1] proposes a 'circular functional relationship' which assumes that the measurement errors are instances of independent and identically distributed (i.i.d.) random variables and that the points lie at fixed but unknown angles around the circumference. This model requires the estimation of the unknown angles of each circumferential point, in addition to the center and radius of the circle. CHAN proposes an approximate method to find the MLE when the errors have a GAUSSIAN distribution. This method is identical to the least-squares method of [2]. He also examines the consistency of the estimator.

A disadvantage of the MLE for circles is that it can be difficult to obtain numerically. As a result, existing algorithms for computing the MLE only produce locally optimal estimates. It is well known that for high noise there are often several local minima [3–6]. BERMAN & CULPIN [3] have carried out a detailed statistical analysis of both the MLE and the DELOGNE-KÅSA estimator (DKE) which uses a least-squares approach. Specifically, they investigated the asymptotic consistency and variance of the estimates.

Due to the numerical difficulties of the MLE, there are several techniques for fitting which are widely used by practitioners. The NEWTON-RAPHSON method can often fail in the case of fitting circles [3,4], diverging to infinity or entering a limit cycle depending on the arrangement of the points and the initialisation. When it converges there is no guarantee that the local optimum which has been found is the global optimum. There are other iterative estimators which guarantee a reduction in the objective function with

each iteration [4–6]. This is an improvement on the NEWTON-RAPHSON method, however, this does not guarantee convergence in the estimates.

The principle of Branch and Bound was first proposed in a discrete setting by LAND & DOIG in [7] and the Branch and Bound search method is described in a discrete setting by WINSTON [8]. Branch and Bound has previously been proposed for locating approximately circular contours [9] and more general contours in images [10]. In [11], ZHANG & KORF present an analysis of the average computational complexity of Branch and Bound in a typical search application.

In this paper, we propose a new algorithm which uses the principle of Branch and Bound to estimate circle parameters from a set of noisy measurements of points on a circle's circumference. This algorithm repeatedly partitions a circle parameter space into subspaces. After each such refinement of the partitioning, lower and upper bounds on the objective function of the MLE are computed for each subspace. Many subspaces may then be discarded from consideration, leading to an efficient search algorithm which bounds the MLE within an arbitrarily small region in the circle parameter space. We show theoretically and empirically that this algorithm obtains the estimate of globally maximum likelihood.

2. THEORY

2.1. CHAN's Circular Functional Model

CHAN's circular functional model [1] for Cartesian coordinates \mathbf{p}_i , $i = 1, \dots, N$ can be expressed as $\mathbf{p}_i = \mathbf{c} + r\mathbf{u}(\theta_i) + \xi_i$, where $\mathbf{c} = (c_1, c_2)^T$ is the center of the circle, r is its radius, the $\mathbf{u}(\theta_i) = (\cos \theta_i, \sin \theta_i)^T$ are unit vectors and the ξ_i are instances of random vectors 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 covariance $\sigma^2\mathbf{I}$.

2.2. Maximum-Likelihood Estimation

If we let $\Omega = (\mathbf{c}, r, \{\theta_i\})^T$, the conditional likelihood for Ω is

$$L(\Omega | \mathbf{p}_i) = \frac{1}{(2\pi\sigma^2)^N} \prod_{i=1}^N \exp\left(-\frac{\|\mathbf{p}_i - (\mathbf{c} + r\mathbf{u}(\theta_i))\|^2}{2\sigma^2}\right). \quad (1)$$

By taking the logarithm of (1) and ignoring the constant offset and scaling, both of which depend on N and σ only, it is possible to simplify the objective function so that

$$\begin{aligned} F_{\text{ML}}(\mathbf{c}) &= \sum_{i=1}^N \|\mathbf{p}_i - \mathbf{c}\|^2 - \frac{1}{N} \left\{ \sum_{i=1}^N \|\mathbf{p}_i - \mathbf{c}\| \right\}^2 \\ &= N \text{VAR}[\|\mathbf{p}_i - \mathbf{c}\|], \end{aligned} \quad (2)$$

Emanuel Zelniker and Benjamin Appleton are additionally supported by a scholarship from the Commonwealth Scientific & Industrial Research Organisation.

where $\text{VAR}[\|\mathbf{p}_i - \mathbf{c}\|]$ is the empirical variance of $\|\mathbf{p}_i - \mathbf{c}\|$. Thus, if we define $d^i(\mathbf{x}) = \|\mathbf{p}_i - \mathbf{x}\|$ where $\mathbf{x} \in \mathbb{R}^2$ is the set of candidate circle centres, then the MLE of the circle parameters is

$$(\hat{\mathbf{c}}_{\text{ML}}, \hat{r}_{\text{ML}}) = \arg \min_{(\mathbf{x}, \rho)} \sum_{i=1}^N [d^i(\mathbf{x}) - \rho]^2, \quad (3)$$

$$\hat{\mathbf{c}}_{\text{ML}} = N \arg \min_{\mathbf{x}} \text{VAR}[d^i(\mathbf{x})], \quad (4)$$

where $\rho \in \mathbb{R}^+$ is a candidate circle radius.

2.3. Sets and Partitions

In this paper we consider functions applied to points, sets, and set partitions. It is assumed that we are searching a finite-volume subset Γ of a real space \mathbb{R}^N of finite dimension. We consider only compact sets, which are closed and bounded in the spaces that we consider.

Definition 1. A set partition P of a compact set Γ is a finite collection of compact subsets $P = \{\gamma_i\}$ with union Γ and pairwise intersection of LEBESGUE measure zero, i.e. $\forall i \neq j, \mathcal{L}(\gamma_i \cap \gamma_j) = 0$, and $\bigcup P = \Gamma$.

For clarity, functions whose argument is a single point will be denoted $f(\cdot)$, functions whose input is a set will be denoted $f[\cdot]$, and functions whose input is a set partition will be denoted $f\langle \cdot \rangle$. For example, the function $\min[\cdot]$ maps a set to its minimal element. Set functions $f[\cdot]$ may be applied to set partitions (which are simply sets of sets) to give set-valued output. When round brackets are used with set argument, we interpret the result as set-valued, i.e., $f(S) = \{f(x) \mid x \in S\}$.

Definition 2. The diameter $D[\cdot]$ of a compact set γ is the lowest upper bound on the distance between two elements of γ , i.e. $D[\gamma] = \sup_{\mathbf{p}_1, \mathbf{p}_2 \in \gamma} \|\mathbf{p}_1 - \mathbf{p}_2\|$.

The diameter $D[\gamma]$ is finite for a compact set $\gamma \subseteq \mathbb{R}^M$. Observe that if $\gamma_1 \subseteq \gamma_2$, $D[\gamma_1] \leq D[\gamma_2]$.

The diameter $D\langle \cdot \rangle$ of a set partition P is defined as the maximum diameter over all subsets $\gamma \in P$, i.e., $D\langle P \rangle = \max_{\gamma \in P} D[\gamma]$.

Definition 3. Given two set partitions P_1 and P_2 of the set Γ , we define a partial ordering $P_1 \leq P_2 \iff \forall \gamma \in P_2, \exists \gamma' \in P_1$ s.t. $\gamma \subseteq \gamma'$.

This means that P_2 more finely partitions Γ than P_1 . Observe that, for two partitions $P_2 \geq P_1$ of Γ , $D\langle P_2 \rangle \leq D\langle P_1 \rangle$.

For the remainder of this paper, we only provide lemma and theorem statements. All proofs may be found in [12].

2.4. Branch and Bound

Consider a scalar function $F: \Gamma \rightarrow \mathbb{R}$ over the domain Γ whose globally extremal point(s) we wish to locate. In this section we assume that the point of minimal value is desired, with the generalisation to finding maximal points implicit throughout. It is assumed that F is a continuous function and that Γ is a compact set.

In maximum-likelihood estimation, Γ corresponds to the parameter space of a model while F computes the negative log-likelihood of a given parameter vector. The search method known as Branch and Bound may be applied to search for a minimum of the function F in the parameter space Γ by repeatedly partitioning

Γ into compact subsets and bounding the range of values within each set. Sets which may contain the minimum are then considered in greater detail by partitioning into smaller subsets, while sets which cannot contain the minimum may be discarded from consideration. In this paper we apply Branch and Bound to solve continuous problems which are of the form $\hat{\mathbf{c}} = \arg \min_{\mathbf{x} \in \Gamma} F(\mathbf{x})$.

2.4.1. Bounding a function over a region

Consider a region in the parameter space $\gamma \subseteq \Gamma$. Then a pair of set functions $F_{\min}[\gamma]$, $F_{\max}[\gamma]$ bound F over the region γ if and only if $F_{\min}[\gamma] \leq \min[F(\gamma)]$ and $\max[F(\gamma)] \leq F_{\max}[\gamma]$. Such bounds need not be tight. Indeed the bounding functions $F_{\min}[\gamma]$, $F_{\max}[\gamma]$ may be chosen for simple computation so as to avoid evaluating F at all points in γ .

Lemma 1. Given two compact sets $\gamma_1 \subseteq \gamma_2 \subseteq \Gamma$ we have that $\min[F(\gamma_1)] \geq \min[F(\gamma_2)]$ and $\max[F(\gamma_1)] \leq \max[F(\gamma_2)]$.

So, as we remove points from a set γ_2 to produce the subset γ_1 , the lower bound is monotonically non-decreasing while the upper bound is monotonically non-increasing.

The analogous property for bounding functions now follows.

Definition 4. F_{\min} , F_{\max} are monotonic if and only if, for compact sets $\gamma_1 \subseteq \gamma_2 \subseteq \Gamma$, $F_{\min}[\gamma_1] \geq F_{\min}[\gamma_2]$ and $F_{\max}[\gamma_1] \leq F_{\max}[\gamma_2]$.

Lemma 2. Consider a sequence of non-empty compact sets $\gamma_1 \supseteq \gamma_2 \supseteq \dots \supseteq \gamma_\infty$ with monotonically decreasing diameter $\delta_i = D[\gamma_i]$ converging to 0. Then define γ_∞ is a set of zero radius, i.e. a point. Then, for a continuous function F , $\lim_{i \rightarrow \infty} \min[F(\gamma_i)] = F(\gamma_\infty) = \lim_{i \rightarrow \infty} \max[F(\gamma_i)]$.

Definition 5. The bounding functions F_{\min} , F_{\max} are convergent if and only if $\lim_{i \rightarrow \infty} F_{\min}[\gamma_i] = \lim_{i \rightarrow \infty} F_{\max}[\gamma_i] = F(\gamma_\infty)$.

2.4.2. Bounding a set partition

We may obtain bounds for the global minimum of F on Γ from the bounding functions F_{\min} , F_{\max} applied to a set partition P of Γ

$$\min[F_{\min}[P]] \leq \min[F(\Gamma)] \leq \min[F_{\max}[P]]. \quad (5)$$

Consider a sequence of increasingly fine partitions $P_1 \leq P_2 \leq \dots$ of Γ with diameter converging to zero. Then for convergent bounding functions F_{\min} and F_{\max} , the bounds on the global minimum computed from (5) converge to the global minimum, i.e. $\lim_{i \rightarrow \infty} \min[F_{\min}[P_i]] = \min[F(\Gamma)] = \lim_{i \rightarrow \infty} \min[F_{\max}[P_i]]$.

Lemma 3. Sets $\gamma' \in P$ with lower bound $F_{\min}[\gamma']$ greater than the minimal upper bound $\min[F_{\max}[P]]$ cannot contain a global minimum, i.e., $F_{\min}[\gamma'] > \min[F_{\max}[P]] \rightarrow \hat{\mathbf{c}} \notin \gamma'$. In the search for global minima such sets may be discarded.

2.4.3. Branch and Bound Algorithm

Here we describe an application of the Branch and Bound principle to locate the global minimum of a function F . Figure 1 depicts a Branch and Bound search tree alongside the corresponding partitioning of the search space Γ . We begin by considering a set in the parameter space Γ which contains the global minimum. This set is recursively partitioned into smaller subsets of equal size. We may then evaluate the bounds F_{\min} and F_{\max} on the minimum value of F within each set. The bounds on the minimum value of F

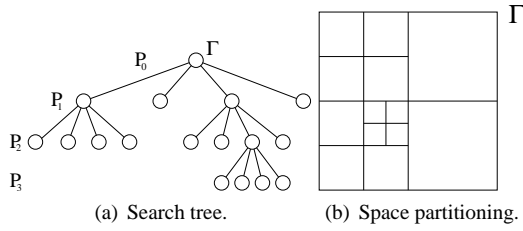


Fig. 1. Branch and bound applied to function optimisation.

guide the branching process, restricting the regions of Γ which need to be considered in more detail. In particular, the application of Lemma 3 allows us to remove from consideration a number of sets which cannot contain the global minimum. This process *prunes* the search tree, reducing the number of regions which must be searched. Finally, the algorithm halts when the position of the global minimum is known within a prescribed distance δ .

Algorithm 1: (Branch & Bound)

Locate the position of the minimum $\hat{\mathbf{c}} = \arg \min_{\mathbf{x} \in \Gamma} F(\mathbf{x})$ within a distance δ :

- Set $P = \{\Gamma\}$
- Until all sets $\gamma \in P$ may be contained within a diameter $D[\cup P] < \delta$

Branch: Refine P by partitioning each set in P

Bound: For all $\gamma \in P$, compute bounds $F_{\min}[\gamma]$ and $F_{\max}[\gamma]$

Prune: Discard all sets $\gamma \in P$ with $F_{\min}[\gamma] > \min[F_{\max}[P]]$

Theorem 1. For a uniformly continuous function F with a single point $\hat{\mathbf{c}}$ of globally minimal value and for convergent bounding functions F_{\min} and F_{\max} , Algorithm 1 terminates after a finite number of operations.

Theorem 2. For a function F with a single point $\hat{\mathbf{c}}$ of globally minimal value, at termination Algorithm 1 locates this point within a distance δ .

The rate of convergence of the algorithm depends upon the refinement scheme for the set partition P as well as the properties of the bounding functions F_{\min} and F_{\max} . Note that the algorithm described in this paper will not terminate when, due to a pathological configuration of the noisy circle points, there are multiple global minima of exactly equal value. As the set of such point configurations has measure zero it is not considered in this paper.

3. ALGORITHM

3.1. Log-Likelihood Bounds

In this section we consider the objective function in (2) This function is defined over the parameter space of possible circle centers, $\Gamma = \mathbb{R}^2$. In this section we consider set partitionings of Γ into rectangles γ with extent $[c_{\min}^1, c_{\max}^1] \times [c_{\min}^2, c_{\max}^2]$.

Given a rectangle $\gamma \subseteq \Gamma$ we wish to define bounding functions F_{\min} and F_{\max} which are monotonic, convergent and efficiently computable. We do so by bounding each term of the summations in (2), which in turn requires bounding the distances $d^i(\mathbf{x})$.

3.1.1. Bounding d^i

The likelihood function in (2) depends on $d^i(\mathbf{x})$, which are uniformly continuous convex functions. We treat $d^i(\mathbf{x})$ as a scalar field and bound it by 1st order polynomials, *i.e.*, we may define tight bounds on $d^i(\mathbf{x})$ over the rectangle γ of the form $d_{\min}^i(\mathbf{x}) \leq d^i(\mathbf{x}) \leq d_{\max}^i(\mathbf{x})$, where

$$d_{\min}^i(\mathbf{x}) = \nabla d^i(\mathbf{c}_\gamma) \cdot (\mathbf{x} - \mathbf{p}_i), \quad (6)$$

$$d_{\max}^i(\mathbf{x}) = \nabla d^i(\mathbf{c}_\gamma) \cdot (\mathbf{x} - \mathbf{p}_i) + B_{\max}^i,$$

$\nabla d^i(\mathbf{c}_\gamma)$ is the unit vector $\nabla d^i(\mathbf{c}_\gamma) = (\mathbf{p}_i - \mathbf{c}_\gamma) / \|\mathbf{p}_i - \mathbf{c}_\gamma\|$, B_{\max}^i is the intercept of the upper bound and \mathbf{c}_γ is the center of the rectangle γ . Note that $\nabla d^i(\mathbf{c}_\gamma)$ is not defined at the circle point \mathbf{p}_i . As a result, the bounding functions $d_{\min}^i(\mathbf{x})$ and $d_{\max}^i(\mathbf{x})$ in the rectangle containing the circle point \mathbf{p}_i are treated separately at the end of this section. Now, it can be shown that B_{\max}^i must be the value of $f(\mathbf{x})$ at one of the corners of the rectangle γ , *i.e.* $B_{\max}^i = \max_{\mathbf{x} \in \text{corners}(\gamma)} \{d^i(\mathbf{x}) - \nabla d^i(\mathbf{c}_\gamma) \cdot (\mathbf{x} - \mathbf{p}_i)\}$. These first order bounds are tight and hence both monotonic and convergent. They are also very simple to compute.

Rectangles γ which contain a circle point \mathbf{p}_i present difficulties in the computation of the bounds $d_{\min}^i(\mathbf{x})$ and $d_{\max}^i(\mathbf{x})$ above, as the distance function $d^i(\mathbf{x})$ is non-differentiable at \mathbf{p}_i . Consequently, these rectangles are bound by the constant functions $d_{\min}^i(\mathbf{x}) = \min_{\mathbf{x}' \in \gamma} \{d^i(\mathbf{x}')\}$ and $d_{\max}^i(\mathbf{x}) = \max_{\mathbf{x}' \in \gamma} \{d^i(\mathbf{x}')\}$. These bounds are also tight and hence are both monotonic and convergent. However as they are zero order functions they converge more slowly than the first order bounds described above. As a result we use the first order bounds for rectangles which do not contain any circle points, defaulting to the zero order bounds only when necessary.

3.1.2. Bounding F

Combining these we may derive bounds F_{\min} and F_{\max} on the uniformly continuous log-likelihood function F :

$$F_{\min}(\mathbf{x}) = \sum_{i=1}^N d_{\min}^i(\mathbf{x})^2 - \frac{1}{N} \left\{ \sum_{i=1}^N d_{\max}^i(\mathbf{x}) \right\}^2, \quad (7)$$

$$F_{\max}(\mathbf{x}) = \sum_{i=1}^N d_{\max}^i(\mathbf{x})^2 - \frac{1}{N} \left\{ \sum_{i=1}^N d_{\min}^i(\mathbf{x}) \right\}^2.$$

It is worth noting that the bounds $d_{\min}^i(\mathbf{x})$ and $d_{\max}^i(\mathbf{x})$ in the previous section were chosen to be uniformly non-negative, leading to simple expressions for $F_{\min}(\mathbf{x})$ and $F_{\max}(\mathbf{x})$. Note that here we are taking differences of 2nd order polynomials, and therefore obtain quadratic bounds on the likelihood function $F(\mathbf{x})$ over $\mathbf{x} \in \gamma$.

Finally, we may minimise the quadratics F_{\min} and F_{\max} over the associated rectangle γ to determine the bounds on the minimum of F within this rectangle, *i.e.* $F_{\min}[\gamma] = \min[F_{\min}[\gamma]]$, and $F_{\max}[\gamma] = \min[F_{\max}[\gamma]]$.

3.2. Complexity Analysis

A simple partition refinement scheme for Algorithm 1 is as follows. Set Γ to be a sufficiently large rectangle encapsulating the the center of the circle. At each refinement step, split each rectangle in the partition into four smaller rectangles of equal size, thereby ensuring that the rectangle diameter halves with each iteration. Hence, the number of refinement steps is proportional to the

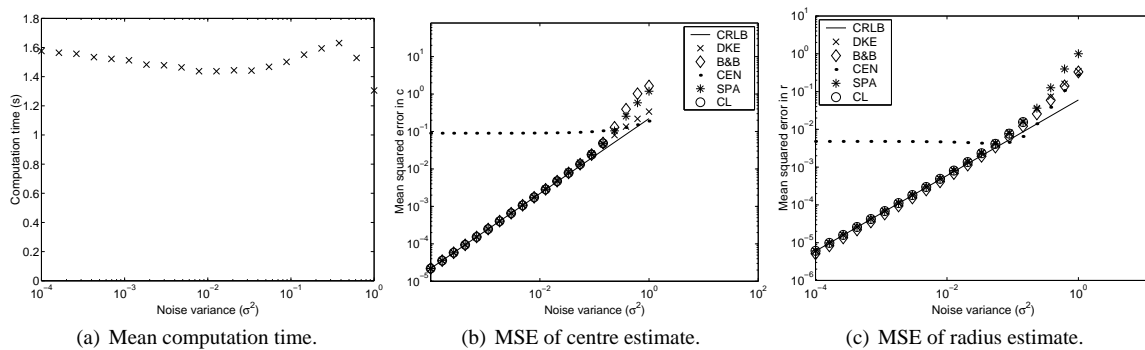


Fig. 2. Simulation results.

negative logarithm of the desired precision δ in the location of the global minimum.

In general Branch and Bound methods may have poor worst-case complexity. However as ZHANG & KORF [11] note they often have low average complexity. Here we briefly mention the major results of the arithmetic complexity of the proposed algorithm.

Theorem 3. *The diameter of a partition has order $O(D^2)$.*

Corollary 1. *Each iteration of Algorithm 1 takes a constant amount of time.*

Corollary 2. *As each iteration of Algorithm 1 improves the precision of the circle centre estimate \hat{c} by one bit in each coordinate, the running time of the algorithm is proportional to the logarithm of the desired precision of the circle centre estimate.*

4. SIMULATIONS

For the experiments in this paper an initial rectangle was selected which was centered on the points and had an area 100 times that of the bounding rectangle of the points.

The branch and bound (B&B) algorithm was simulated using a Monte-Carlo analysis. Twenty points with no noise ($N = 20$) were generated with a uniform distribution around the circumference of a unit circle. For each value of σ , the algorithm was evaluated over 10,000 trials. In each trial, noise was added to the true points to obtain estimates for the center of the circle \hat{c} and radius \hat{r} according to CHAN's circular functional model. This was used to generate mean square error (MSE) values. The amount of noise, σ was varied from 10^{-2} to 1 in equal geometric increments. The same was done for the DKE (DKE) [13], the centroid method (CEN) by taking the mean of the x and y coordinates, the SPÄTH algorithm (SPA) [5] and the CHERNOV & LESORT algorithm (CL) [4]. Also, for the algorithm proposed in this paper, the number of likelihood function evaluations was recorded in order to demonstrate its independence to σ . See Figure 2(a).

The MSE values in centre, \hat{c} , and radius, \hat{r} , are plotted against their corresponding CRAMÉR-RAO lower bound (CRLB, see [12]) for the same level of noise σ in Figure 2(b) and 2(c) on a logarithmic scale. As the noise level, σ , approaches zero, all methods except CEN approach statistical efficiency. The centroid method levels off due to bias. The CHERNOV & LESORT method diverged for the highest 4 values of noise and the results cannot be plotted.

5. REFERENCES

- [1] Nai Ng Chan, "On Circular Functional Relationships," *J. Roy. Statist. Soc. Ser. B Stat. Methodol.*, vol. 27, pp. 45–56, Apr. 1965.
- [2] Stephen M. Robinson, "Fitting Spheres by the Method of Least Squares," in *Comm. ACM*, 1961, p. 491.
- [3] Mark Berman and David 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.
- [4] N. Chernov and C. Lesort, "Least squares fitting of circles and lines," to appear in *Journal of Mathematical Imaging and Vision*.
- [5] H. Späth, "Least-Square Fitting by Circles," *Computing*, vol. 57, pp. 179–185, 1996.
- [6] U. M. Landau, "Estimation of a Circular Arc Center and its Radius," *Comput. Vision Graph. Image Process.*, vol. 38, pp. 317–326, June 1986.
- [7] A. Land and A. Doig, "An automatic method of solving discrete programming problems," *Econometrica*, vol. 28, pp. 497–520, 1960.
- [8] P. H. Winston, *Artificial Intelligence*, Addison-Wesley Publishing Company, Inc, 1984.
- [9] Ben Appleton and Changming Sun, "Circular Shortest Paths by Branch and Bound," *Pattern Recognition*, vol. 36, no. 11, pp. 2513–2520, Nov. 2003.
- [10] Ben Appleton and Hugues Talbot, "Globally optimal geodesic active contours," *Journal of Mathematical Imaging and Vision*, 2002, Accepted.
- [11] W. Zhang and R. Korf, "An average-case analysis of branch-and-bound with applications: Summary of results," in *Proc. 10th National Conf. on Artificial Intelligence, AAAI-92*, San Jose, CA, July 1992, pp. 545–550.
- [12] E. E. Zelniker, B. C. Appleton, and I. Vaughan L. Clarkson, "Optimal Circle Fitting via Branch and Bound," accepted to ICASSP2005.
- [13] E. E. Zelniker and I. Vaughan L. Clarkson, "A Statistical Analysis of the DELOGNE-KÁSA Method for Fitting Circles," submitted to *Digital Signal Processing*, unpublished.