Digitized circular arcs: characterization and parameter estimation

Worring, M.; Smeulders, A.W.M.

Published in:
IEEE Transactions on Pattern Analysis and Machine Intelligence

DOI:
10.1109/34.387505

Citation for published version (APA):

General rights
It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations
If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: http://uba.uva.nl/en/contact, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.
Abstract—The digitization of a circular arc causes an inherent loss of geometrical information. Arcs with slightly different local curvature or position may lead to exactly the same digital pattern. In this paper we give a characterization of all centers and radii of circular arcs yielding the same digitization pattern. The radius of the arcs varies over the set. However, only one curvature or radius estimate can be assigned to the digital pattern. We derive an optimal estimator and give expressions for the bound on the precision of estimation. This bound due to digitization is the deterministic equivalent of the Cramér/Rao bound known from parameter estimation theory.

Consider the estimation of the local curvature and local radius of a smooth object. Typically such parameters are estimated by moving a window along the digital boundary. Methods in literature show a poor precision in estimating curvature values, relative errors of over 40% are often found [34]. From the definition of curvature it follows that locally the curve can be considered a circular arc and hence the method presented in this paper can be applied to the pattern in the window giving estimates with optimal precision and a measure for the remaining error.

On the practical side we present examples of the residual error due to the discrete grid. The estimation of the radius or curvature of a circular arc at random position with an estimation window containing 10 points (coded with nine Freeman codes) has a relative deviation exceeding 2%. For a full disk the deviation is below 1% when the radius exceeds four grid units.

The presented method is particularly useful for problems where some prior knowledge on the distribution of radii is known and where there is a noise-free sampling.

Index Terms—Circular arcs, parameter estimation, curvature, precision, digitization.

I. INTRODUCTION

In this paper we consider the estimation of geometric features of smooth continuous objects when digitized on a regular grid.

We rely on the fundamental theorem in the differential geometry of curves [28] to find what features are of importance in the analysis of smooth boundary curves. The theorem states that one can reconstruct any smooth curve up to a rigid transformation if curvature is known as function of arc length, where curvature is the inverse of the radius of the circle locally coinciding with the curve. This makes curvature (or the radius of the circular arc for that matter) and arc length the two basic features to consider in the analysis of any smooth (line) figure.

For the estimation of the curvature of the digitized boundary, five essentially different methods are found in [1], [2], [9], [20], [29]. These five methods were compared in [33], [34] on the basis of their performance in measuring the curvature of circular arcs repeatedly placed on a random position with respect to the digitization grid. It turned out that none of the methods yields accurate and precise estimates of curvature.

A theoretical analysis of the methods revealed a number of clues for the poor performance. In [20] the curvature is estimated by applying a linear differentiating filter to the x- and y-coordinate separately. For circular arcs this separation, bypassing the 2D nature of the problem, introduces significant errors due to the truncation of the filter. The methods in [1], [2], [9] find curvature by applying a linear differentiating filter to the estimated orientation. The method in [9] is the only method which explicitly takes the anisotropy of the grid into account and in fact with proper scaling this is the method with best performance. In the other cases errors in the order of 40% are common.

Fitting a circular arc to the digital data does take the two dimensional shape of the digital boundary into account, and algorithms can be found in [4], [18], [29]. It turned out that for noise free digitized circular arcs the precision is poor.

Clearly, the precision of the estimation is limited as a consequence of both the stochastic noise introduced in the imaging system and by the digitization noise. There is, however, an important difference in the characteristics of the two types of noise. Stochastic noise is a random process, and the precision of measurement is limited by the Minimum Variance or Cramér/Rao bound [5]. In contrast, the digitization noise is not a random process. It is completely governed by the position and shape of the object and by the deterministic characteristics of the spatial quantization, dependent on the resolution and connectivity of the digitization grid. Hence, for digitization there is a deterministic bound on the precision equivalent to the Cramér/Rao bound. With high precision scanning systems as flat bed scanners or when images are recorded at limited resolution, the stochastic noise is in general much smaller than the digitization noise. For the purpose of accurate measurement this makes the effect of digitization most important. This will, therefore, be the focus of this paper.

Extensive studies on the effect of spatial quantization on straight boundaries [6], [14], [19] led to accurate estimators for measuring arc length and bounds on their precision [7], [8], [16], [17], [24]. These studies also led to very accurate registration of straight digitized boundaries in an image pair [3]. We proceed along the same path as [6], [14], [19] for the analysis of curved boundaries and seek expressions for the deterministic equivalent of the Cramér/Rao bound in local
shape estimation as well as a method reaching this bound.

To introduce the concept consider the curvature of a set of points in a window on the digital contour resulting from digitizing the contour of a smooth object. The interest is not in the curvature of the digital pattern, but in the curvature of the preimage of the contour. The digitization causes an inherent loss of geometrical information. Other arcs (with slightly different local curvature or position) may lead to exactly the same digital pattern. Thus, the precision of estimation is limited by the capability of the digital pattern to discriminate small variations in the curvature of the arc. Here we are concerned with the ultimate precision one can reach.

The first step in the analysis is to verify that the digital pattern indeed is the digitization of some circular preimage. Algorithms for this step are presented in [15], [21], [22]. One step further beyond the recognition is to render all centers of circular arcs which generate this pattern [10], not considering the radius of the arc. A complete characterization of all centers and radii is presented in [32]. This is done in [30] also, for a few very simple digital patterns only, using the theory of locales of an object [13]. The theory of locales is not applicable if only part of the boundary of the object is given.

We start off in Section II with an analysis of the digitization process and give a type classification of all possible point sets of limited size. From there the full characterization of digitized circles for a given pattern is studied based on the results in [32]. From there (Section III) we will find expressions for the bound on precision and present an optimal shape parameter estimator for noise free digitized circular arcs. A very preliminary presentation of the concept was presented in [27]. Finally, in Section IV practical bounds are calculated.

II. DIGITIZATION AND DOMAINS

A. Digitization

In this section we make precise how the digitization affects a curved object boundary.

For the digitization of an object X we can consider two different digitization models, Grid Intersect Quantization (GIQ) or Object Boundary Quantization (OBQ) [12]. GIQ is appropriate for the generation of arcs in computer graphics, for image processing OBQ is the more common model. Thus we consider OBQ only; the results derived here, however, are applicable for GIQ with minor modifications. We further restrict our attention to eight-connected grids, but this is also a choice not affecting the general idea in the paper; similar results are obtained for four- or six-connected grids.

For eight-connected objects pixels are on the boundary of the digital object if they are four-connected to a pixel in the background. The eight-connected contour of the object X is hence given by:

**DEFINITION 1. (Object Boundary Quantization)**

\[ D_{OBQ}(X) = \{ s \in \mathbb{Z}^2 | s \in X \land \exists s' \in N_4(s) : s' \in X \}, \]

where \( N_4(s) \) is the set of points four-connected to \( s \). We can order the points of \( D_{OBQ}(X) \) in counterclockwise fashion [25].

So we write \( D_{OBQ}(X) = \{ s_k \}_{k=0}^{m-1} \), where \( m \) is the number of contour points in \( X \) and \( s_0 \) an arbitrary starting point.

In the practice of curvature or radius estimation a window of size \( n' \) is moved along the points of \( D_{OBQ}(X) \), and from this set of points local estimates are derived (see Fig. 1(a) where \( n' = 6 \)). The set of \( n' \) points starting at point \( s_0 \) is defined by:

\[ S_p(X, k, n') = \{ s_j \in D_{OBQ}(X) \} \text{ for } j = k, \ldots, (k + n' - 1) \mod m, \]

where the modulo operator is used because we are dealing with closed, and therefore periodic, contours.

![Fig. 1.](image)

The points of the pattern \( S_p \) are the points just to the inside of the continuous boundary of \( X \) (Definition 1). These points have at least one four-connected neighbor not in \( S_p \). The set \( S_Q \) of such associated background pixels is given by:

\[ S_Q(S_p) = \{ s \in \mathbb{Z}^2 | s \in X \land \exists s' \in S_p : s' \in N_4(s) \}. \]

In the sequel we will omit the parameters of \( S_p \) and \( S_Q \). The ordered pair of point sets \((S_p, S_Q)\) together constitutes the digital pattern \( S \).

The points of \( S_p \) are described using Freemancodes [11]. For each point \( s_i \in S_p \), the Freemancode \( f_i \) indicates which of the eight possible neighbors of \( s_i \) is the next point in \( S_p \) (see Fig. 1). Note that the Freemancode is the coding of a vector. The vector (not its coding) is denoted by \( v(f_i) \), where \( f_i \in \{0, 7\} \). The sequence \( f_i \) with \( f_{i+1} \) is defined for \( t \leq n - 2 \). The forward difference is taken modulo 8 such that \( \Delta f_i = f_{i+1} - f_t \) defined for \( t \leq n - 2 \).

To generate all valid Freemancodes of given length \( n \), and hence all valid sets \( S_p \) with \( n + 1 \) points, which might occur in the digitization of some circular arc, we consider the contour tracing algorithm as defined in [25]. Starting from a contour point \( s_0 \in X \) for which \( s_0 + v(0) \notin X \) this algorithm considers neighbors of \( s_0 \) in counterclockwise order, starting from \( s_0 + v(0) \), until it encounters another element in \( X \). Then the algorithm continues with this new element in \( X \) found and the previous pixel encountered which consequently is not in \( X \).

From the definition of the tracing algorithm it follows that for any point \( s_i \in S_p \), with associated Freemancode \( f_i \), the point...
$s_i + v(f_i - 1)$ must be an element of $S_Q$. Furthermore, if $f_i$ is even, the point $s_i + v(f_i - 2)$ must also be in $S_Q$, as otherwise the tracing algorithm would have chosen this point as the next contour point instead of $s_i + v(f_i)$.

In the digitization of an arbitrary object a Freemancode could be followed by a Freemancode in exactly the opposite direction, i.e., $\Delta f = 4$. However, the resulting pattern, $(S_p, S_Q)$ can never be part of the digitization of a circular arc, which is easily verified using the results in Section II.B.

For all the remaining cases we can give one continuous arc (using window size $n = 2$) which has the resulting pattern $(S_p, S_Q)$ as its digitization. Consequently, the possible Freeman differences which can occur in the digitization of a circular arc are given by:

$$\Delta f_i \in \{-1, \ldots, 3\} \quad (f_i \in \{0, 2, 4, 6\})$$

By looking at all points not in the object $X$ which the algorithm considers for an element $s_i \in S_p$ and keeping only those points which are four-connected to a point in $S_p$ we find that the background points $q_i$ associated with $s_i$ are given by:

$$q_i = \{s_i, s_i + v(f_i - j) \} \quad (f_i \in \{0, 2, 4, 6\} \land \Delta f_i \neq 2)$$

and finally,

$$S_Q = \bigcap_i q_i.$$}

The domain of the patterns is exemplified in Fig. 2.

As there is a direct mapping from a Freemanchain $F$ to the corresponding digital pattern $(S_p, S_Q)$ we will not make a distinction between the two in the sequel and will use the appropriate format to describe the pattern.

**Definition 2. (Circular Disk)**

$$B(m,r) = \{c \in \mathbb{R}^2 \mid d(m,c) \leq r\}$$

$$B^*(m,r) = \{c \in \mathbb{R}^2 \mid d(m,c) > r\}$$

The 3D-parameter space $(m, r)$ of all centers and radii of arcs is denoted by $\mathbb{R}$. The digitization operator $DOB$ (Definition 1) maps elements from the continuous space $\Omega$ into a pattern in the discrete space $2^\mathbb{R}$. A continuum of various (but closely related) arcs $\omega \in \Omega$ are indistinguishable after digitization. Hence, there exists no inverse operator. Now, let $(S_p, S_Q)$ be the pattern corresponding to an estimation window, positioned somewhere along the digital boundary resulting from digitizing disk $B(\omega_0)$. The domain of $S$ is the set of all arcs, which after digitization generate the pattern $S$ somewhere in their digital boundary. To that end we define the domain operator $B$, being the pseudo-inverse of the digitization operator. It gives the equivalence class of $\omega_0$:

**Definition 3. (Domain Operator)**

$$B(S_p, S_Q) = \{\omega \in \Omega \mid S_P \subseteq D_OB(B(\omega))\}.$$
With (8) we have formulated the set of centers of all pre-image arcs which generate the digital pattern. Before we proceed with the characterization of the domain of centers and radii in Section II.D we distinguish a number of different domain types based on the set of centers only and use it to derive a bound on the radii of arcs which can be analyzed on the basis of a point set described by a Freemanchain of length $n$.

C. Characterization of Domain Type

In the preceding section we investigated a pattern $(S_p, S_q)$ and found disks including all of $S_p$ and excluding all of $S_q$ locally assuming a convex object. However, when estimating the curvature of an arbitrary smooth object, moving an estimation window along the digital boundary, the original curve can locally either be convex or concave. In the latter case (7) can still be applied, but with the roles of $S_p$ and $S_q$ reversed.

As in general it is not known beforehand which case applies both have to be considered. Depending on whether the domains for the two cases are empty or nonempty four different types are distinguished. For each nonempty domain we consider $r_{\text{max}}$, the radius of the largest disk, which might be finite or infinite. This leads to two more types of patterns (see Fig. 4).

Fig. 4. Examples of patterns corresponding to chains of $n = 4$ codes.

These relations are used in conjunction with (3) to efficiently find all patterns in the set of all Freemanchains of length $n$ (denoted by $\mathcal{F}^n$) which can be in the digitization of a circular arc. Without loss of generality we limit ourselves to patterns starting with Freemancode 0 or 1. Patterns which are self-intersecting are classified as noncircular as the tracing algorithm discussed in Section II will never generate such patterns. For the same reason patterns with partial overlap are classified noncircular, except when the last point of the pattern equals the first point of the pattern indicating a closed boundary. Counting the number of different types found yields the type classification in Table I.

<table>
<thead>
<tr>
<th>Table I</th>
<th>Counting the Type of all $2 \times 8^2$ Freemanchains of Length $n$ With 0 or 1 as First Code. The Classification of the Patterns is Given. Also Given is the Maximum Recognizable Radius for Varying $n$, Where for Concave Patterns the Obvious Extension of MRR to Such Patterns is Used.</th>
</tr>
</thead>
<tbody>
<tr>
<td>$n$</td>
<td>convex</td>
</tr>
<tr>
<td></td>
<td>strict</td>
</tr>
<tr>
<td>3</td>
<td>10</td>
</tr>
<tr>
<td>4</td>
<td>56</td>
</tr>
<tr>
<td>5</td>
<td>171</td>
</tr>
<tr>
<td>6</td>
<td>399</td>
</tr>
<tr>
<td>7</td>
<td>821</td>
</tr>
<tr>
<td>8</td>
<td>1,266</td>
</tr>
<tr>
<td>9</td>
<td>2,575</td>
</tr>
</tbody>
</table>

It follows from the table that only a small fraction of elements in $\mathcal{F}^n$ lead to valid circular patterns. Furthermore, most of these circular patterns are either strictly convex or strictly concave.

Now, consider a continuous disk placed on an arbitrary position with respect to the grid. When moving a window along the digital contour of the disk we expect to find strictly
convex patterns only. However, when the window size is chosen too small, the local pattern might be the infinite convex or even straight type. This is undesirable and poses a limit on the radius of continuous circular arcs we can analyze with a given windowsize. To be precise, for a window with corresponding Freeman chains of size \( n \) the maximum recognizable radius \( MRR(n) \) is defined as the radius of the largest disk which at an arbitrary position is assured to yield convex-type patterns only. To find \( MRR \) we have to consider all straight and infinite convex patterns corresponding to chains in \( F^* \) and make sure that the radius of the original disk is smaller than the smallest radius in any such domain i.e.:

**Definition 4. (Maximum Recognizable Radius)**

\[
MRR(n) = \min_{F \in F^*} \min_{\text{finite convex, straight}} \{\text{radius of } F\},
\]

where \( \text{radius of } F \) is the minimum radius of any disk in the domain of the pattern associated with \( F \).

In Table I the maximum recognizable radius is given for \( n = 3, \ldots, 9 \). The \( MRR \) values show that only small radii can be recognized if the number of codes is limited. For larger \( n \) the \( MRR(n) \) tends to a linear relationship with \( n \).

**D. Characterization of the Arc Domain**

In Section II.B the set of all centers of circular arcs which could lead to the given pattern \( S \) was shown to be the generalized Voronoi polygon. In this section we consider the full extent of the domain in \( \Omega \) and present an algorithm for computing a characterization of the domain. As the emphasis in this paper is on the shape estimation part of the problem proofs concerning the characterization will be omitted. Here we will give a summary only; interested readers are referred to [32]. We restrict ourselves to the convex case. However, results are exactly the same if the roles of \( S_P \) and \( S_Q \) are reversed throughout this section. In fact, \( S_P \) and \( S_Q \) can be two arbitrary disjunct point sets, they do not even have to be confined to the discrete grid.

Now, consider the set of centers for a fixed value of \( r \), denoted by \( M_r(S_P, S_Q) \), which consequently is a subset of the generalized Voronoi polygon. This set is given by:

**Definition 5. (Arc Centers)**

\[
M_r(S_P, S_Q) = \{m \in \mathbb{R}^2 \mid S_P \subset B(m, r) \land S_Q \subset B^*(m, r)\}.
\]

Thus, the set of centers is bounded by a polygon with convex and concave edges in its boundary. For an example see Fig. 3. Clearly we have:

\[
\bigcup_r M_r(S_P, S_Q) - V(S_P, S_Q)
\]

The curved edges in the boundary of \( M_r(S) \) form a sequence, where each curved edge is centered at some element in \( S \). They are the points of the set which constitute the local constraint on \( M_r \). Hence, we have an ordering of the elements in \( S \) which is denoted by \( SEQ_r(S) \).

**Definition 6. (r-sequence)**

\[
SEQ_r(S) = \text{ordered sequence of the elements of } S, \text{ corresponding to the ordered sequence of edges in } \partial M_r(S).
\]

For example, in Fig. 3 we have \( SEQ_r(S) = p_2, p_3, q_0, q_1, q_3, q_5 \). It should be noted that the \( r \)-sequence may actually contain a number of sequences each of which has a cyclic ordering. Further, note that not necessarily every element of \( S \) is part of the sequence. The remaining points of \( S \) do not play a limiting role on the arcs of radius \( r \).

Every three subsequent points \( (s', s, s'') \) in the sequence define one curved edge in the boundary of \( M_r(S) \). The point \( s \) defines the center of the edge and the points \( s' \) and \( s'' \) in conjunction with \( s \) determine the endpoints of the edge. Whether the edge is convex or concave depends on whether \( s \in S_P \) (convex) or \( s \in S_Q \) (concave).

For varying \( r \), the shape of \( \partial M_r(S) \) changes in a continuous fashion in the sense that the radius of the bounding edges increases. The composition of \( \partial M_r \), i.e., \( SEQ_r(S) \) changes only at discrete values of \( r \). At these values of \( r \) a new edge is added to the boundary or an existing edge will no longer participate for larger radii. The domain of a pattern \( S \) is fully characterized if all changes to \( SEQ_r(S) \) are known for varying \( r \).

To compute the characterization we divide \( V(S_P, S_Q) \) in regions (in 2D-parameter space) such that every region is critically determined by a specific \( s \in S \). For the elements of \( S_P \) we use the intersection of the furthest point Voronoi polygon, a structure well known in computational geometry [23], with \( V(S_P, S_Q) \). For elements in \( S_Q \) we use the closest point Voronoi polygon.

**Definition 7. (s-Region)**

\[
L(s) = \begin{cases} 
V(S_P, s) \cap V(S_P, S_Q) & (s \in S_P) \\
V(s, S_Q) \cap V(S_P, S_Q) & (s \in S_Q) 
\end{cases}
\]

In Fig. 6 the \( s \)-regions are shown for a simple arbitrary pattern.

The \( s \)-regions provide a disjunct covering of \( V(S_P, S_Q) \), which will be called the \( L_P \)-diagram for set \( S_P \) and the \( L_Q \)-diagram for set \( S_Q \).

![Fig. 6. For every element of \( p \in S_P \) the intersection of the generalized Voronoi polygon of \( S \) with the furthest point Voronoi polygon corresponding to \( p \) yields the \( s \)-region of \( p \). For \( q \), the corresponding \( s \)-region is highlighted (a). In a similar way (b) shows the \( s \)-regions of elements in \( S_Q \) with the \( q \)-region of \( q_4 \) highlighted.](image)
Changes in the \( r \)-sequence are related to straight edges and vertices in the two diagrams. To that end we list all admissible transitions of the \( r \)-sequence, when increasing \( r \) (see Table II). For reasons of simplicity we make the assumption, common in Voronoi related results, that no four points in the pattern \( S \) are co-circular.

### TABLE II

This table gives a characterization of all possible isolated changes to \( \text{SEQ}_n(S) \) occurring, when \( r \) passes through one of the values in the arc spectrum of \( S \). The change is determined by the configuration of the points defining the minimum or maximum distance to an edge in the \( L_p \)- or \( L_Q \)-diagram depending on the set(s) the elements stem from and on the type of triangle they form. The symbols \( \alpha \) and \( \gamma \) denote a (possibly) empty sequence of points from \( S \).

<table>
<thead>
<tr>
<th>configuration</th>
<th>point set ( T )</th>
<th>( r = p(T) - e )</th>
<th>( r = p(T) + e )</th>
<th>case</th>
</tr>
</thead>
<tbody>
<tr>
<td>obtuse triangle</td>
<td>( [p', p, q] )</td>
<td>( \alpha, p', q, \gamma )</td>
<td>( \alpha, p, q', \gamma )</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>( [q', q, p] )</td>
<td>( \alpha, q, p', \gamma )</td>
<td>( \alpha, p, q', \gamma )</td>
<td>2</td>
</tr>
<tr>
<td>acute triangle</td>
<td>( [p, q'] )</td>
<td>( \alpha, q, q', \gamma )</td>
<td>( \alpha, q', q, \gamma )</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>( [q, q', p] )</td>
<td>( \alpha, q, q', \gamma )</td>
<td>( \alpha, q', q, \gamma )</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>( [q', q, p] )</td>
<td>( \alpha, q, q', \gamma )</td>
<td>( \alpha, q', q, \gamma )</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>( [q, q', p] )</td>
<td>( \alpha, q, q', \gamma )</td>
<td>( \alpha, q', q, \gamma )</td>
<td>6</td>
</tr>
<tr>
<td>right triangle</td>
<td>( [p, p'] )</td>
<td>( \alpha, p, p' )</td>
<td>( \alpha, p', p )</td>
<td>7</td>
</tr>
<tr>
<td></td>
<td>( [q, q'] )</td>
<td>( \alpha, q, q' )</td>
<td>( \alpha, q', q )</td>
<td>8</td>
</tr>
<tr>
<td>minimal distance</td>
<td>point pair</td>
<td>( \alpha, \gamma )</td>
<td>( \alpha, \gamma )</td>
<td>9</td>
</tr>
<tr>
<td></td>
<td>( [p, q] )</td>
<td>( \alpha, p, q )</td>
<td>( \alpha, p, q )</td>
<td>10</td>
</tr>
</tbody>
</table>

This table and the \( L_p \) and \( L_Q \)-diagrams form the basis for the algorithm to compute the characterization of a pattern \( S \). The pseudocode for this algorithm is given in Fig. 7.

The algorithm first constructs the \( L_p \)- and \( L_Q \)-diagram. From there the edges in the diagram are assigned a direction, such that the distance to the points of which the edge is the perpendicular bisector is increasing. Edges which properly contain the point of minimum distance cannot be assigned such a direction and therefore such edges are split into two separate edges by introduction of an artificial vertex at the point of minimum distance. Thus, following an edge in its given direction is always corresponding to increasing \( r \).

Recursively all vertices of the diagrams are treated, starting from the vertex defining the minimal disk. At every vertex with all incoming edges labeled, the change to the sequence of edges in \( \partial M_c(S) \) is applied, (where the actual change is determined from Table II) and then the algorithm has to choose the appropriate diagram to continue with. If the vertex considered lies in the interior of \( V(S_p, S_Q) \) the same diagram has to be chosen. However, if the vertex lies on the boundary of \( V(S_p, S_Q) \) the choice depends on whether the center of the newly created arc is an element of \( S_p \) or \( S_Q \). Then the algorithm treats all endpoints of the outgoing edges. If one of the diagrams contains artificial vertices, which do not define the minimal disk, some preprocessing has to be done as such a vertex does not have incoming edges and thus, would otherwise never be processed.

### III. SHAPE PARAMETER ESTIMATION

In the previous section the arc domain of a pattern described in parameter space was considered; now we consider the estimation of the radius or curvature from the digital pattern resulting from digitizing a circular arc.

#### A. The Geometric Minimum Variance Bound

By definition all circular arcs in the arc domain yield the same digital pattern \( S \). However, only one estimate \( \hat{g} \) for the shape parameter of the arc can be assigned to the pattern \( S \). As the true shape parameter \( g \) varies over different elements in \( B(S) \), an uncertainty is unavoidably introduced in the estimation of the shape parameter of the preimage. This is quantified in the domain variance.

**Definition 8. (Domain Variance)**

\[
\text{VAR}_g(S) = \int_{w \in B(S)} p(\omega / S)(\hat{g}(S) - g(\omega))^2 d\omega.
\]
Table III
Characterization of the Arc Domain of the Example Pattern S Used
in Fig. 6. Shown are the r-values for which changes in $\text{SEQ}_r(S)$
occur. The points $T$ causing the change (where $r = \text{mp}(T)$, the case in
Table II corresponding to the configuration of $T$, and finally the
changing $r$-sequence.

<table>
<thead>
<tr>
<th>$r$</th>
<th>$T$</th>
<th>case</th>
<th>$\text{SEQ}_r(S)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>18.6</td>
<td>$[5,6,7,8]$</td>
<td>7</td>
<td>$[5,6,7,8]$</td>
</tr>
<tr>
<td>17.0</td>
<td>$[5,6,7,8]$</td>
<td>1</td>
<td>$[5,6,7,8]$</td>
</tr>
<tr>
<td>18.5</td>
<td>$[5,6,7,8]$</td>
<td>4</td>
<td>$[5,6,7,8]$</td>
</tr>
<tr>
<td>19.7</td>
<td>$[5,6,7,8]$</td>
<td>8</td>
<td>$[5,6,7,8]$</td>
</tr>
<tr>
<td>20.2</td>
<td>$[5,6,7,8]$</td>
<td>2</td>
<td>$[5,6,7,8]$</td>
</tr>
<tr>
<td>22.0</td>
<td>$[5,6,7,8]$</td>
<td>9</td>
<td>$[5,6,7,8]$</td>
</tr>
<tr>
<td>25.5</td>
<td>$[5,6,7,8]$</td>
<td>2</td>
<td>$[5,6,7,8]$</td>
</tr>
<tr>
<td>31.6</td>
<td>$[5,6,7,8]$</td>
<td>4</td>
<td>$[5,6,7,8]$</td>
</tr>
</tbody>
</table>

where $p(\omega|S)$ is the probability of $\omega$ being the original leading
to the pattern $S$ and $g((m, r)) = r$ when estimating the radius of
the arc and $g((m, r)) = 1/r$ for curvature estimation.

We seek the estimator $\hat{\omega}_{\text{mv}}(S)$ minimizing the domain variance, i.e.,

**Definition 9.**

$$\hat{\omega}_{\text{mv}}(S) = \arg\min \left\{ \text{VAR}_{\omega}(S) \right\}$$

This turns out to be the expectation $g$ over the arc domain.

**Theorem 1.**

$$\hat{\omega}_{\text{mv}}(S) = \int_{\omega \in \mathcal{B}(S)} p(\omega|S) g(\omega) d\omega$$

**Proof.** From the definition of the domain variance and $\hat{\omega}_{\text{mv}}$
we have:

$$\hat{\omega}_{\text{mv}} = \arg\min \left\{ \int_{\omega \in \mathcal{B}(S)} p(\omega|S)(\hat{\omega} - g(\omega))^2 d\omega \right\}$$

Rewriting the expression, taking constants not depending on
$\omega$ out of the integration, and using

$$\int_{\omega} p(\omega|S) d\omega = 1$$

yields:

$$\hat{\omega}_{\text{mv}}(S) = \arg\min \left\{ \hat{\omega}^2 - 2\hat{\omega} \int_{\omega \in \mathcal{B}(S)} p(\omega|S) g(\omega) d\omega 
+ \int_{\omega \in \mathcal{B}(S)} p(\omega|S) g^2(\omega) d\omega \right\}$$

We are looking for a stationary point with respect to $\hat{\omega}$ of
the above term within the curly braces. Taking the first deriv-ative with respect to $\hat{\omega}$ and substituting $\hat{\omega} - \hat{\omega}_{\text{mv}}$ leads to
the following condition on $\hat{\omega}_{\text{mv}}(S)$:

$$2\hat{\omega}_{\text{mv}}(S) - 2\int_{\omega \in \mathcal{B}(S)} p(\omega|S) g(\omega) d\omega = 0$$

The second derivative with respect to $\hat{\omega}$ of the term within
the curly braces is a positive constant (2) not depending on
$\hat{\omega}$. Hence, the stationary point is a minimum and thus,

$$\hat{\omega}_{\text{mv}}(S) = \int_{\omega \in \mathcal{B}(S)} p(\omega|S) \hat{\omega}(\omega) d\omega$$

As the domain variance is a deterministic measure, the
domain variance of $\hat{\omega}_{\text{mv}}$ is the minimal variance one can reach for
any possible estimator. This should be discriminated from the
bound on the minimum variance of parameter estimates from
stochastic data [5] (MVBB, or Cramér-Rao lower bound). In
this bound the variance is a stochastic measure. The latter
bound is a genuine lower bound only if one restricts the class
of estimators to be linear and unbiased. Hence, we define the
Geometric variant of the Minimum Variance Bound and de-
note it by GMVB_

**Definition 10. (Geometric Minimum Variance Bound)**

$$\text{GMVB}_g(S) = \text{VAR}_{\omega}(\hat{\omega}_{\text{mv}}(S))$$

The bound quantifies the ability of the pattern $S$ to discrimi-
nate among variations in the shape parameter of the preimage
of the arc, i.e., it expresses the maximum achievable precision
in shape parameter estimation.

To find an expression for the bound we make the following
three assumptions:

1) Features only depend on $r$
2) There is no preferred position
3) Radius and position are independent.

Now, consider the following moment generating function
for features $g$ which only depend on $r$ (assumption 1):

$$T^g_S = \int_{r \in \mathcal{A}(S)} \tilde{p}(r|S) g^2(r) dr$$

where the function $\tilde{p}(r|S)$ is proportional to the conditional probability function on $r$ given the pattern $S$. It gives
the "number" of centers corresponding to this value of $r$, weighted by the prior probability on $r$.

The equation has important applications [31]. First, consider $T^g_S$, this gives a measure for the total number of pre-
images of the pattern. It is the normalization factor needed to
make $\tilde{p}$ a proper probability density function. Second, it al-

tows to express $\hat{\omega}_{\text{mv}}$ and GMVB as:

$$\hat{\omega}_{\text{mv}} = \frac{T^1_S}{T^0_S} \quad \text{GMVB}_g(S) = \frac{T^2_S}{T^0_S} \left( \frac{T^1_S}{T^0_S} \right)^2$$

In the next section we consider the computation of $T^4_S$.

**B. The Computation of the Moment Generating Function**

We seek the probability on the radius $r$, given that the dig-
itzation of the arc leads to the digital pattern $S$. We assume no
a priori information on position (assumption 2). Thus, within
the domain of $S$, the centers have a uniform distribution and
the probability on $r$ is proportional to the area $\mathcal{A}$ in domain
space of the set of arc centers $\mathcal{M}(S)$. For the time being we
leave open the choice for the prior distribution on $r$. Using the
independence of position and radius (assumption 3) we find:

$$\hat{\omega}_{\text{mv}}(S) = \int_{\omega \in \mathcal{B}(S)} p(\omega|S) \hat{\omega}(\omega) d\omega$$

In [7] the equivalent theorem for straight lines is proved in a different
way. In their proof they need to make the assumption that the estimator is
linear. Our result shows that this assumption is not needed.
and it follows that we have to find an expression for the area of $M_r(S)$.

The boundary of the set $M_r(S)$ is a polygon with curved edges. Let the vertices of $M_r(S)$ be given by $v_i$, and let $\phi_i$ be the size of the edge (in radians) connecting $v_i$ and $v_{i+1}$; then the area of $M_r(S)$ is expressed using dot products as [26]:

$$A(M_r(S)) = \sum_i \left\{ \frac{1}{2} (v_i \cdot (v_{i+1} \cdot \frac{1}{2} r^2 (\phi_i - \sin \phi_i) \right\}$$

with

$$\left( \begin{array}{c} x \\ y \\ \end{array} \right) = \left( \begin{array}{c} y \\ -x \\ \end{array} \right)$$

The first term of the equation, summed over all $i$, gives the area of the straight polygon through the vertices $v_i$. The second term corrects for the area of the segment between the curved arc and the chord from $v_i$ to $v_{i+1}$, denoted by $(v_i, v_{i+1})$. The area of the segment is added if the curved edge is convex (the center is in $S_r$) and subtracted if it is concave (the center is in $S_Q$).

We obtain a summation over contributions from individual arcs in the boundary of $M_r(S)$. As noted, each curved edge in $M_r(S)$ is defined by three subsequent points in $SEQ_r(S)$. Let $(s', s, s'')$ be such a triplet of points and recall that $s$ defines the center and $s'$ and $s''$ in conjunction with $s$ determine the start and endpoint of the edge.

The start point $v'$ of the arc lies on the perpendicular bisector of $s'$ and $s$. To be precise, the start point is an intersection of the circles centered at $s'$ and $s$ with radius $r$. The two intersections of these circles are given by (see Fig. 8(a)):

$$v_{s',s}(r) = s + \frac{1}{2} (s', s) \pm \sqrt{\frac{1}{4} \| (s', s) \|^2 - \frac{1}{2} \| (s', s) \|^2}$$

In [32] it is shown that whenever $s, s'$ are in the same set of $S$ (i.e., $s, s' \in S_P$ or $s, s' \in S_Q$) one should take the minus case, otherwise one should take the plus. To find the end point $v''$ one simply replaces $s'$ by $s$ and $s$ by $s''$.

The size of the arc connecting $v'$ to $v''$ is calculated as:

$$\phi_{s',s''}(r) = \cos^{-1} \left( \frac{(s, v_{s',s}(r)) \cdot (s, v_{s',s''}(r))}{r^2} \right)$$

The geometric construction used to derive the equation is exemplified in Fig. 8(b).

The contribution $A_r$ of the triplet $(s', s, s'')$ to the area of the curved polygon $M_r(S)$ is given by:

$$A_r((s', s, s'')) = \frac{1}{2} v_{s',s}(r) \cdot v_{s',s''}(r) \pm \frac{1}{2} r^2 (\phi_{s',s''}(r) - \sin \phi_{s',s''}(r))$$

Now, let $T_r(S)$ be the set of all triplets defining edges in the boundary of $M_r(S)$ i.e., all triplets obtained by (cyclically) taking three subsequent points in $SEQ_r(S)$. Then it follows from (12) that:

$$p(r / S) \propto \bar{p}(r) \sum_{a \in T_r(S)} A_r(a)$$

Every triplet of points defining an edge in the boundary of $M_r(S)$ does so for a distinct range of $r$-values, see Section B. Thus, the term under the integral sign in the moment generating function (10) changes (for increasing $r$) whenever an edge is introduced in the boundary of $M_r(S)$ or one annihilated, i.e., whenever a change to $SEQ_r(S)$ occurs. It is more convenient to rewrite the equation in terms of the triplets defining edges and their associated minimum and maximum radius. So, let $A(S)$ be the set of all triplets occurring in $SEQ_r(S)$ for some value of $r$ and let $r_{\min}(a)$ be the value of $r$ at which the edge defined by triplet $a$ occurs in $M_r(S)$ and let $r_{\max}$ be the value at which the edge is annihilated.

We can now summarize the main theoretical result of this paper in the following theorem.

**Theorem 2.**

$$T^*_r(S) = \sum_{a \in A(S)} \left[ \int^{r_{\max}(a)}_{r_{\min}(a)} \tilde{p}(r) A_r(a) g(r) \, dr \right]$$

where $T^*_r(S)$ is the moment generating function (10) for the feature $g(r)$ (e.g., the radius) to be estimated. $p(r)$ is the a priori distribution of $r$. $A(S)$ is the set of all triplets of points defining edges in $M_r(S)$ for some value of $r$ and $a(a)$ is the contribution of triplet $a$ to the area of $M_r(S)$, which is nonzero only for $r \in [r_{\min}(a), r_{\max}(a)]$.

The characterization of the arc domain presented in the preceding section provides us with all the information needed to compute the set $A(S)$ and the associated $r_{\min}$ and $r_{\max}$ values. From this a numerical integration method is used to calculate the moment generating function using Theorem 2. This in turn allows for the computation of the optimal estimator $g_{\mathbf{opt}}$ and the GMVB using (11).

In Figs. 9 and 10 the lower bound and optimal radius estimate are illustrated for example patterns using a uniform a priori distribution on $r$. 
C. The Positional Error

The Geometric Minimum Variance Bound gives a bound to the precision of the measurement for a given pattern of points. Now consider the estimation of shape properties from a given continuous disk of radius \( r \).

When the continuous disk is placed at some fixed position on the grid and a window containing \( n \) Freemancodes is moved along the resulting digital contour, optimal estimates of the radius or curvature will vary for different parts of the contour. Further, when the disk is placed at different positions with respect to the grid, the optimal estimates in the windows vary as the digital contour of the disk is not the same for each position. As the measurement error for a pattern only depends on the codes in the Freemanchain and not on its start position we can fix the start position of all Freemanchains to the origin and see whether it contains arcs of radius \( r \). The expected positional error is thus defined as:

**Definition 11. (Positional Error)**

\[
E^p_\theta(r) = \sum_{F \in F^s} p(F^*/r) (g(r) - \hat{g}(F^*))^2
\]

As the centers are uniformly distributed we find the following conditional probability for a specific Freemanchain \( F \).

\[
p(F/r) = \frac{\mathcal{A}(M_r(F))}{\sum_{F^* \in F^s} \mathcal{A}(M_r(F^*))}
\]

(18)

The estimator \( \hat{g}_{mv} \) has the smallest expected positional error (for varying \( r \)) among all possible estimators:

**Proposition 1.**

\[
\hat{g}_{mv} = \arg \min_{\hat{g}} \left\{ \int_r p(r) E^p_\theta(r) \, dr \right\}
\]

**Proof.** From (11) we have:

\[
\int_r p(r) E^p_\theta(r) \, dr = \int_r p(r) \left\{ \sum_{F^* \in F^s} p(F^* / r) (g(r) - \hat{g}(F^*))^2 \right\} \, dr
\]

(19)

In order to change the order of integration and summation we apply Bayes’ rule to \( p(F/r) \) and find:

\[
p(F/r) = \frac{\tilde{p}(r) \mathcal{A}(M_r(F))}{\int_r \tilde{p}(r) \mathcal{A}(M_r(F)) \, dr}
\]

\[
p(r) \propto \tilde{p}(r)
\]

Let again \( \tilde{p}(r) \) denote the unnormalized a priori probability on \( r \). Then the terms in the equation are given by:

\[
p(r / F) = \frac{\tilde{p}(r) \mathcal{A}(M_r(F))}{\int_r \tilde{p}(r) \mathcal{A}(M_r(F)) \, dr}
\]

\[
p(F) \propto \sum_{F^* \in F^s} \left\{ \int_r \tilde{p}(r) \mathcal{A}(M_r(F^*)) \, dr \right\}
\]

Rearranging terms we find:

\[
p(F / r) = \frac{\mathcal{A}(M_r(F))}{\sum_{F^* \in F^s} \left\{ \int_r \tilde{p}(r) \mathcal{A}(M_r(F^*)) \, dr \right\}}
\]

By changing the order of summation and integration (noting that the right hand side is a constant times \( \mathcal{A}(M_r(F)) \)) we find:

\[
\int_r p(r) E^p_\theta(r) \, dr = \sum_{F^* \in F^s} \left\{ \int_r p(r) \mathcal{A}(M_r(F^*)) (g(r) - \hat{g}(F^*))^2 \, dr \right\}
\]

As we have a sum of positive terms the total expected positional error is minimized if each individual term is minimized. It follows from Theorem 1 that each term is minimized if \( \hat{g}_{mv}(F) \) is used which proves the proposition.
Hence, the positional error $E_i$ for estimator $\hat{g}_{mv}$ quantifies the ultimate bound on the error for shape parameter measurement on the arc when placed at random positions.

To study the error it is convenient to split the error in a bias and a deviation part. The positional bias tells us how the digitization affects the apparent shape of the continuous preimage.

**DEFINITION 12. (Positional Bias)**

\[
\text{BIAS}^n(r) = E \{\hat{g}_{mv}(F) \} - g(r)
\]

where $E$ denotes expectation over all elements in $\mathcal{F}^n$, i.e.,

\[
E[h] = \sum_{F \in \mathcal{F}^n} p(F | h) h(F).
\]

The positional deviation quantifies the ultimate precision with which we can measure the shape parameters of the arc.

**DEFINITION 13. (Positional Deviation)**

\[
\text{DEV}^n(r) = \sqrt{E[\hat{g}_{mv}(F)^2]} - \left( E[\hat{g}_{mv}(F)] \right)^2
\]

IV. RESULTS

We calculated $\text{BIAS}^n(r)$ and $\text{DEV}^n(r)$ for $n = 7, 8, 9$ with $r \leq \text{MRR}(n)$ for the convex patterns. The results are summarized in Fig. 11, for estimation of the radius of the predigitized arc, and in Fig. 12 for curvature estimation. Both were calculated with uniform priors on the feature to be estimated. They cannot be directly compared as the priors on the radius of the original arcs are different.

From the figures we conclude that the observed radius of the arc increases if we use too few Freeman codes in the chain as confirmed by the curvature estimates obtained. Curvature estimates are unbiased, a consequence of the fact that for curvature estimation the a priori probability on $r$ is proportional to $1/r^2$ and hence favors smaller radii. As expected, more precise estimates are found when $n$ is larger. However, for the range of radii used, the deviation for $n = 9$ is still between 2% and 9% in both radius and curvature estimation. To find better estimates, one needs larger values of $n$, hence considering larger arcs. As an alternative one can use denser sampling but then one has to change $n$ accordingly to keep the arc in the window of the same size.

Now consider the estimation of the radius $r$ of a disk when given its full digital boundary. Then $n$ is no longer constant but equals the number of points in the digital boundary which varies when the disk is placed at random positions on the grid. The values for bias and deviation are derived from Monte Carlo experiments placing the disk at random position with respect to the grid a 1,000 times. Results are presented in Fig. 13.

It follows that for a full disk the estimation of the radius is virtually unbiased. Deviation is below 1% if the radius is larger than four grid units.

Fig. 11. The $\text{BIAS}^n(r)$ and $\text{DEV}^n(r)$ for radius estimation from (open) circular arcs with $r$ in the range 2.0 to MRR($n$) and $n = 7, 8, 9$. Both expressed as fraction of the true radius.

Fig. 12. The $\text{BIAS}^n(r)$ and $\text{DEV}^n(r)$ for curvature estimation from (open) circular arcs with $r$ in the range 2.0 to MRR($n$) and $n = 7, 8, 9$. Expressed as fraction of the true curvature.

Fig. 13. Optimal estimation of the radius of a full circular disk, when repeatedly placed at random positions with respect to the grid. The bias and deviation are indicated, both relative to the true radius.

V. SUMMARY AND CONCLUSION

In this paper we have considered two topics. First, we have characterized the set of all continuous circular arcs which give rise to a given digitization pattern. Second, we have derived optimal estimates for the curvature and radius of the continuous arc which due to the inherent loss of information in the digitization process cannot be improved.

The local curvature or local radius of an object is estimated by moving a window along the digital contour. Optimal
estimation can only be achieved if all circular arcs which give rise to the specific digital pattern in the window are considered. The set of parameters of arcs which give rise to the pattern is called the arc domain of the pattern.

A domain can be classified as one of six types (straight, strictly convex, infinite convex, strictly concave, infinite concave, and noncircular) which is calculated for all possible Freemancodes. Based on the domain types one can find the maximal recognizable radius MRR(n) which poses a limit on the radius of continuous arcs one can estimate using a window of n Freemancodes. The type classification can further be used to decompose a given digital contour of an object into straight, convex, and concave parts. A full characterization of the domain of any specific digital pattern in \((x_{cenr}, y_{cenr}, \text{radius})\)-space is given in Section II.B based on the changes in the sequence of curved edges bounding the set of centers in the domain for varying \(r\).

The arc domain of a given digital pattern by definition is the set of all predigitized arcs which are mapped to this pattern. As the radius of elements in the arc domain varies, an unavoidable error is made in the radius or curvature estimation. This error is bounded by the equivalent of the Cramér-Rao bound expressing the influence of digitization rather than stochastic noise. It is called the Geometric Minimum Variance Bound (Definition 10). Computation of the GMVB using the characterization of the arc domain is based on the moment generating function given in Theorem 1. The unique estimator achieving this minimal variance is the expectation of the shape feature over the arc domain (Theorem 1). Theorem 2 again provides the basis for computing the estimator in terms of the characterization.

In contrast to methods commonly used (see [34]) the method presented in this paper allows to incorporate prior knowledge on the distribution of the shape parameter of the circular arcs, always achieving optimal precision. The shape of the prior distribution is arbitrary. In the absence of other prior information a uniform distribution of the feature can be assumed (as is used in this paper).

Practical bounds on the precision in measuring curvature or the radius using uniform priors were presented in Section IV. It is found that the relative deviation in the digitization limited optimal shape parameter measurement of arcs placed at random positions with \(r \leq 6\) grid units, using a window of \(n = 9\) Freemancodes, is between 2% and 9%. For digitizations of full disks (and hence varying \(n\)) the deviation is below 1% for \(r \geq 4\) grid units.

The bounds derived cannot directly be compared to the more practical curvature measurement methods discussed in [33], [34]. In the reference arcs of larger radii were considered and larger number of Freemancodes were used with a weighting function on the points in the window. However, the presented method does not suffer from any of the problems associated with the other estimators which ignored the two dimensional character of the curve and the special structure of digitized circular arcs. The performance is also better than arc fitting as full knowledge of the digitization process is incorporated. The precision of estimation of the new estimator, even when using small chains, clearly supports these observations.

In practical situations, the theoretical best estimator can be calculated from the digital pattern directly or it can be precomputed for all patterns of a given size and stored in a lookup table. For chains of nine Freemancodes, the table has about 4,500 entries. All other 33,000,000 chains of nine codes are not the digitization of a circular arc.

This method is particularly useful for problems where some prior knowledge on the distribution of radii is known and where there is a noise-free sampling. It can lead to optimal performance for a given measurement error. For example, in the surveillance of circular holes in industrial objects performance will be critical, the sampling can be arranged such that it is noise-free, and the distribution \(p(r)\) is known. Under these circumstances the grid size can be reduced to the bare minimum given the maximum tolerable measurement error.

ACKNOWLEDGMENTS

The remarks made by the reviewers have clearly improved the presentation of the paper. This work was supported by the Dutch Ministry of Economic Affairs through SPIN grants "3D Computer Vision" and "3D Image Analysis."

REFERENCES