Computer Science Department of The University of Auckland
CITR at Tamaki Campus (http://www.citr.auckland.ac.nz)

CITR-TR-83 February 2001

Minimum-Length Polygons In Approximation Sausages

Tetsuo Asano[1], Yasuyuki Kawamura[1], Reinhard Klette[2],
And Koji Obokata[1]

Abstract

The paper introduces a new approximation scheme for planar digital curves. This scheme defines an approximating sausage 'around' the given digital curve, and calculates a minimum-length polygon in this approximating sausage. The length of the polygon is taken as an estimator for the length of the curve being the (unknown) preimage of the given digital curve. Assuming finer and finer grid resolution it is shown that this estimator converges to the true perimeter of an *r*-compact polygonal convex bounded set. This theorem provides theoretical evidence for practical convergence of the proposed method towards a 'correct' estimation of the length of a curve. The validity of the scheme has been verified through experiments on various convex and non-convex curves. Experimental comparisons with two existing schemes have also been made.

[1] School of Information Science, JAIST, Asahidai, Tatsunokuchi, 923-1292, Japan
[2] CITR, University of Auckland Tamaki Campus, Building 731, Auckland, New Zealand.

Minimum-Length Polygons in Approximation Sausages

Tetsuo Asano[1], Yasuyuki Kawamura[1], Reinhard Klette[2], and Koji Obokata[1]

[1] School of Information Science, JAIST
Asahidai, Tatsunokuchi, 923-1292 Japan
t-asano@jaist.ac.jp
[2] CITR, University of Auckland, Tamaki Campus, Building 731
Auckland, New Zealand
r.klette@auckland.ac.nz

Abstract. The paper introduces a new approximation scheme for planar digital curves. This scheme defines an approximating sausage 'around' the given digital curve, and calculates a minimum-length polygon in this approximating sausage. The length of this polygon is taken as an estimator for the length of the curve being the (unknown) preimage of the given digital curve. Assuming finer and finer grid resolution it is shown that this estimator converges to the true perimeter of an r-compact polygonal convex bounded set. This theorem provides theoretical evidence for practical convergence of the proposed method towards a 'correct' estimation of the length of a curve. The validity of the scheme has been verified through experiments on various convex and non-convex curves. Experimental comparisons with two existing schemes have also been made.

Keywords: Digital geometry, digital curves, multigrid convergence, length estimation.

1 Introduction and Preliminary Definitions

Approximating planar digital curves is one of the most important topics in image analysis. An approximation scheme is required to ensure convergence of estimated values such as curve length toward the true length assuming a digitization model and an increase in grid resolution. For example, the *digital straight segment approximation method* (DSS method), see [3, 8], and the *minimum length polygon approximation method* assuming one-dimensional grid continua as boundary sequences (MLP method), see [9], are methods for which there are convergence theorems when specific convex sets are assumed to be the given input data, see [6, 7, 10]. This paper studies the convergence properties of a new minimum length polygon approximation method based on so-called *approximation sausages* (AS-MLP method).

Motivations for studying this new technique are as follows: the resulting DSS approximation polygon depends upon starting point and the orientation of the boundary scan, it is not uniquely defined, but it may be calculated for any given digital object. The MLP approximation polygon is uniquely defined, but it assumes a one-dimensional grid continua as input which is only possible if the given digital object does not have cavities of width 1 or 2. The new method leads to a uniquely defined polygon, and it may be calculated for any given digital object.

Let r be the *grid resolution* defined as being the number of grid points per unit. We consider *r-grid points* $g^r_{i,j} = (i/r, j/r)$ in the Euclidean plane, for integers i, j. Any r-grid point is assumed to be the center point of an *r-square* with *r-edges* of length $1/r$ parallel to the coordinate axes, and *r-vertices*.

The digitization model for our new approximation method is just the same as that considered in case of the DSS method, see [4–6]. That is, let S be a set in the Euclidean plane, called *real preimage*. The set $C_r(S)$ is the union of all those r-squares whose center point $g^r_{i,j}$ is in S. The boundary $\partial C_r(S)$ is the *r-frontier* of S. Note that $\partial C_r(S)$ may consists of several non-connected curves even in the case of a bounded convex set S. A set S is *r-compact* iff there is a number $r_S > 0$ such that $\partial C_r(S)$ is just one (connected) curve, for any $r \geq r_0$. This definition of r-compactness has been introduced in [6] in the context of showing multigrid convergence of the DSS method.

The validity of the proposed scheme has been verified through experiments on various curves, which are described in Section 5. It has also been compared with the existing schemes in convergence and computation time.

2 Approximation Scheme

Given a connected region S in the Euclidean plane and a grid resolution r, the r-frontier of S is uniquely determined. We consider r-compact sets S, and grid resolutions $r \geq r_S$ for such a set, i.e. $\partial C_r(S)$ is just one (connected) curve. In such a case the r-frontier of S can be represented in the form $P = (v_0, v_1, \ldots, v_{n-1})$ in which the vertices are clockwise ordered so that the interior of S lies to the right of the boundary. Note that all arithmetic on vertex indices is modulo n.

Let δ be a real number between 0 and $1/(2r)$. For each vertex of P we define forward and backward shifts: The *forward shift* $f(v_i)$ of v_i is the point on the edge (v_i, v_{i+1}) at the distance δ from v_i. The *backward shift* $b(v_i)$ is that on the edge (v_{i-1}, v_i) at the distance δ from v_i.

For example, in the approximation scheme as detailed below we will replace an edge (v_i, v_{i+1}) by a line segment $(v_i, f(v_{i+1}))$ interconnecting v_i and the forward shift of v_{i+1}, which is referred to as the *forward approximating segment* and denoted by $L_f(v_i)$. The *backward approximating segment* $(v_i, b(v_{i-1}))$ is defined similarly and denoted by $L_b(v_i)$. Refer to Fig. 1 for illustration. Now we have three sets of edges, original edges of the r-frontier, forward and backward approximating segments. Let $0 < \delta \leq .5/r$. Based on these edges we define a connected region $A^\delta_r(S)$, which is homeomorphic to the annulus, as follows:

Given a polygonal circuit P describing an r-frontier in clockwise orientation. By reversing P we obtain a polygonal circuit Q in counterclockwise order. In

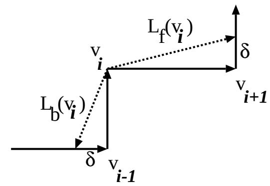

Fig. 1. Definition of the forward and backward approximating segments associated with a vertex v_i.

the initialization step of our approximation procedure we consider P and Q as the *external* and *internal* bounding polygons of a polygon P_B homeomorphic to the annulus. It follows that this initial polygon P_B has area contents zero, and as a set of points it coincides with $\partial C_r(S)$.

Now we 'move' the external polygon P 'away' from $C_r(S)$, and the internal polygon Q 'into' $C_r(S)$ as specified below. This process will expand P_B step by step into a final polygon which contains $\partial C_r(S)$, and where the Hausdorff distance between P and Q becomes non-zero. For this purpose, we add forward and backward approximating segments to P and Q in order to increase the area contents of the polygon P_B.

To be precise, for any forward or backward approximating segment $L_f(v_i)$ or $L_b(v_i)$ we first remove the part lying in the interior of the current polygon P_B and updating the polygon P_B by adding the remaining part of the segment as a new boundary edge. The direction of the edge is determined so that the interior of P_B lies to the right of it.

Definition 1. *The resulting polygon P_B^δ is referred to as the* approximating sausage *of the r-frontier and denoted by $A_r^\delta(S)$.*

The width of such an approximating sausage depends on the value of δ. It is easy to see that as far as the value of δ is at most half of the grid size, i.e., less or equal $1/(2r)$, the approximating sausage $A_r^\delta(S)$ is well defined, that is, it has no self-intersection. It is also immediately clear from the definition that the Hausdorff distance from the r-frontier $\partial C_r(S)$ to the boundary of the sausage $A_r^\delta(S)$ is at most $\delta \leq 1/(2r)$.

We are ready to define the final step in our AS-MLP approximation scheme for estimating the length of a digital curve. Our method is similar to that of the MLP as introduced in [9].

Definition 2. *Assume a region S having a connected r-frontier. An* AS-MLP curve *for approximating the boundary of S is defined as being a shortest closed curve $\gamma_r^\delta(S)$ lying entirely in the interior of the approximating sausage $A_r^\delta(S)$, and encircling the internal boundary of $A_r^\delta(S)$.*

It follows that such an AS-MLP curve $\gamma_r^\delta(S)$ is uniquely defined, and that it is a polygonal curve defined by finitely many straight segments. Note that this curve depends upon the choice of the approximation constant δ. An example of such a shortest closed curve $\gamma_r^\delta(S)$ is given in Fig. 2, with $\delta = .5/r$.

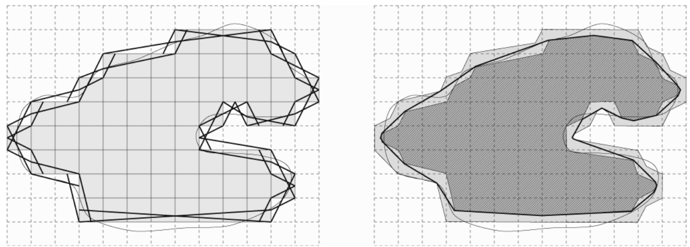

Fig. 2. Left: construction of approximating sausage. Right: approximation by shortest internal path.

3 Properties of the Digital Curve

We discuss some of the properties of the approximating polygonal curve $\gamma_r^\delta(S)$ defined above, assuming that $\partial C_r(S)$ is a single connected curve.

Non-selfintersection: The AS-MLP curve $\gamma_r^\delta(S)$ is defined as being a shortest closed curve lying in the approximating sausage. Since it is obvious from the definition that the sausage has no self-intersection, so does the curve.

Controllability: The width of an approximating sausage can be controlled by selecting a value of δ, with $0 < \delta \leq .5/r$.

Smoothness: Compared with the other two approximation schemata DSS and MLP, our approximating curve is 'more smooth' in the following sense: the angle associated with a corner of an approximating polygon is the smaller one of its internal angle and external angle. We consider the minimum angle of all these angles associated with a corner of the AS-MLP curve. Similarly, such minimum angles may be defined for approximating DSS and MLP curves. It holds that the minimum AS-MLP angle is always greater than or equal to the minimum DSS or minimum MLP angle, if a convex set S has been digitized. Note that 'no small angle' means 'no sharp corner'.

Linear complexity: Due to the definition of our curve $\gamma_r^\delta(S)$ the number of its vertices is at most twice that of the r-frontier.

Computational complexity: Assuming that a triangulation of an approximating sausage is given, linear computation time suffices to find a shortest closed path: we can triangulate an approximating sausage in linear time since the vertices of the sausage can be calculated only using nearby segments. So, linear time is enough to triangulate it. Then, we can construct an adjacency graph, which is a tree, representing adjacency of triangles again in linear time. Finally, we can find a shortest path in linear time by slightly modifying the linear-time algorithm for finding a shortest path within a simple polygon.

Figure 3 gives visual comparisons of the proposed AS-MLP method with two existing schemes DSS and MLP.

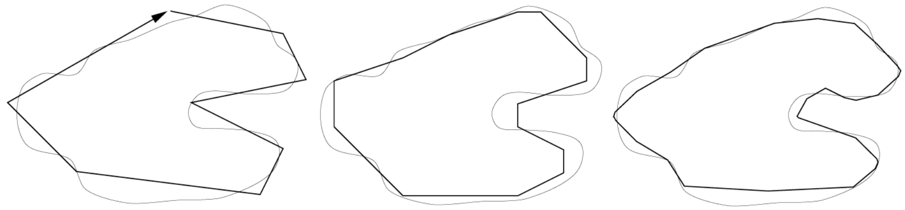

Fig. 3. Original region with DSS (left), MLP (center), and proposed approximation using $\delta = .5/r$ (right).

4 Convergence Theorem

In this section we prove the main result of this paper about the multigrid convergence of the AS-MLP curve based length estimation of the perimeter of a given set S.

Theorem 1. *The length of the approximating polygonal curve $\gamma_r^\delta(S)$ converges to the perimeter of a given region S if S is a r-compact polygonal convex bounded set and $0 < \delta \leq .5/r$.*

We sketch a proof of this theorem with an investigation of geometric properties of the r-frontier of a convex polygonal region S.

We first classify r-grid points into interior and exterior ones depending on whether they are located inside of the region S or not. Then, CH_{in} is defined to be the convex hull of the set of all interior r-grid points. CH_{out} is the convex hull of the set of those exterior r-grid points adjacent horizontally or vertically to interior ones. See Fig. 4 for illustration.

Lemma 1. *The difference between the lengths of CH_{in} and CH_{out} is exactly $4\sqrt{2}/r$.*

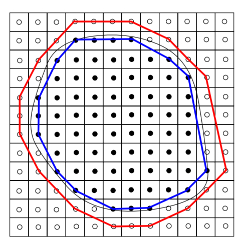

Fig. 4. Interior r-grid points (filled circles) and exterior points (empty circles) with the convex hulls CH_{in} of a set of interior points and CH_{out} of a set of exterior points adjacent to interior ones.

Now, we are ready to state the following lemma which is of crucial importance for proving the convergence theorem.

Lemma 2. *Given an r-compact polygonal convex bounded set S, the approximating polygonal curve $\gamma_r^\delta(S)$ is contained in the region bounded by CH_{in} and CH_{out}, for $0 < \delta \leq .5/r$.*

Let CH be the convex hull of the set of vertices of the approximating polygonal curve $\gamma_r^\delta(S)$. The convex hull CH is also bounded by CH_{in} and CH_{out}. Obviously, the vertices of CH are all intersections of approximating segments. Furthermore, exterior intersections do not contribute to CH, where external (internal, resp.) intersections are those on the external (internal, resp.) boundary of the approximating sausage. Therefore, we can evaluate the perimeter of CH. An increase in distance of an internal intersection from the boundary of CH_{in} corresponds to an increase in length of an approximating segment, and a decrease of distance of its associated intersection to the inner hull CH_{in}. Thus, such an intersection is farthest at a corner defined by two unit edges. Thus, the maximum distance from CH_{in} to CH is bounded by $\sqrt{2}/6$, which implies that the perimeter of CH is bounded by $\sqrt{2}\pi/3$.

Lemma 3. *Let CH be the convex hull of all internal intersections defined above. Then, the approximating polygonal curve $\gamma_r^\delta(S)$ lies between the two convex hulls CH_{in} and CH. The maximum gap between CH_{in} and CH is bounded by $\sqrt{2}/6$, and for their perimeter we have*

$$Perimeter(CH) \leq Perimeter(CH_{in}) + 4\sqrt{2}/r. \tag{1}$$

So, if the approximating polygonal curve $\gamma_r^\delta(S)$ is convex, then we are done. Unfortunately, it is not always convex. In the remaining part of this section we evaluate the largest possible difference on lengths between $\gamma_r^\delta(S)$ and CH.

Lemma 4. *The approximating polygonal curve $\gamma_r^\delta(S)$ is concave when two consecutive long edges of lengths d_{i-1} and d_i with intervening unit edge satisfy $d_i > 3d_{i-1} + 1$.*

By analysis of the possible differences from the convex chain, we obtain the following theorem.

Theorem 2. *Let S be a bounded, convex polygonal set. Then, there exists a grid resolution r_0 such that for all$r \geq r_0$ it holds that any AS-MLP approximation of the r-frontier $\partial C_r(S)$, with $0 < \delta \leq .5/r$, is a connected polygon with a perimeter l_r and*

$$|Perimeter(S) - l_r| \leq (4\sqrt{2} + 8 * 0.0234)/r = 5.844/r. \tag{2}$$

5 Experimental Evaluation

We have seen above that the perimeter estimation error by AS-MLP is bounded in theory by $5.8/r$ for a grid resolution r, for convex polygons. To illustrate

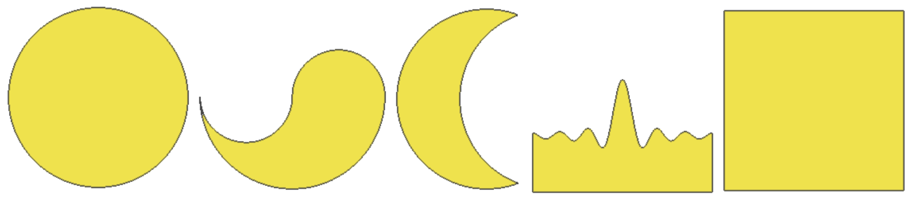

Fig. 5. Experimental objects.

its practical behavior we report on experiments on various curves, which are described below. Although we have restricted ourselves to convex objects in the preceding proof, we took non-convex curves as well in these experiments. Figure 6 illustrates a set of objects used for experiments as suggested in [5].

CIRCLE: the equation of the circle is

$$(x-0.5)^2+(y-0.5)^2=0.4^2.$$

YINYANG: the lower part of the yinyang symbol is composed by arcs of 3 half circles: the lower arc is a part of CIRCLE, and the upper arcs are parts of circles whose sizes are half of CIRCLE.

LUNULE: this object is the remainder of two circles, where the distance between both center points is 0.28.

SINC: the sinc equation corresponding to the upper curve is

$$y = sin\left(\frac{\pi x}{4\pi x}\right).$$

SQUARE: the edges of the isothetic SQUARE are of length 0.8.

5.1 Two Existing Approximation Schemes

We sketch both existing schemes which are used for comparisons, where the DSS and MLP implementation reported in [4] has been used for experimental evaluation. First, the digital straight segment (DSS) algorithm traces an r-frontier, i.e. vertices and edges on $\partial C(S)$, i.e. a boundary of $C(S)$, and detects a consecutive sequence of maximum length DSSs. The sum of the lengths of these DSS is used as DSS curve length estimator. The DSS algorithm runs in linear time.

The minimum-length polygon (MLP) approximation needs two boundaries, of set $I(S)$ and of set $O(S)$, as input. Roughly saying, $I(S)$ is the union of r-squares that are entirely included in S, in other words, all four r-vertices of such a square are included in a convex set S; and $O(S)$ is obtained by 'expanding' $I(S)$ by a dilation using one r-square as structuring element. The MLP algorithm calculates the shortest path in the area $O(S)\setminus I(S)$ circumscribing the boundary of $I(S)$. The length of such a shortest path is used as MLP curve length estimator. The MLP algorithm also takes linear time.

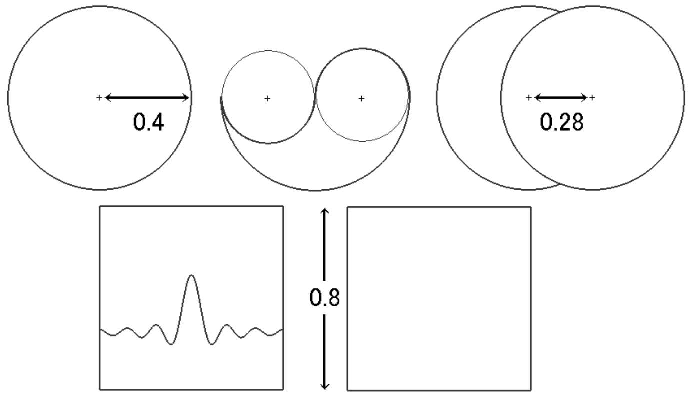

Fig. 6. Test sets drawn in unit size.

In the experiments we computed the errors of three approximation schemes for the specified objects digitized in grid resolutions $r = 32 \sim 1024$. For DSS and AS-MLP, $C(S)$ was used as a digitized region, where $C(S)$ is a set of pixels whose midpoints are included in S. For MLP, $I(S)$ and its expansion was used.

5.2 Experiments

Following the given implementations of DSS, and MLP, also our new AS-MLP scheme has been implemented in C++ for comparisons. We have computed the curve length error in percent compared to the true perimeter of a given set S. The error E_{DSS} of the DSS estimation scheme is defined by

$$E_{DSS} = \frac{P(S) - P(DSS_S)}{P(S)}$$

where $P(S)$ is the true perimeter of S and $P(DSS_S)$ is the perimeter of the approximation polygon given by the DSS scheme. E_{MLP} and E_{ASMLP} are analogously defined.

Figure 7 shows the errors for all five test curves, the boundaries of CIRCLE, YINYANG, LUNULE, SINC, and SQUARE in that order, from top to bottom. The diagrams for DSS, MLP, and AS-MLP are arranged from left to right, in each row of the figure. The graphs illustrate that AS-MLP has smaller errors in general than MLP has, but DSS is the best among the three.

6 Conclusion

We proposed a new approximation scheme for planar digital curves and analyzed its convergence to the true curve length by stating a theorem for convex sets. To

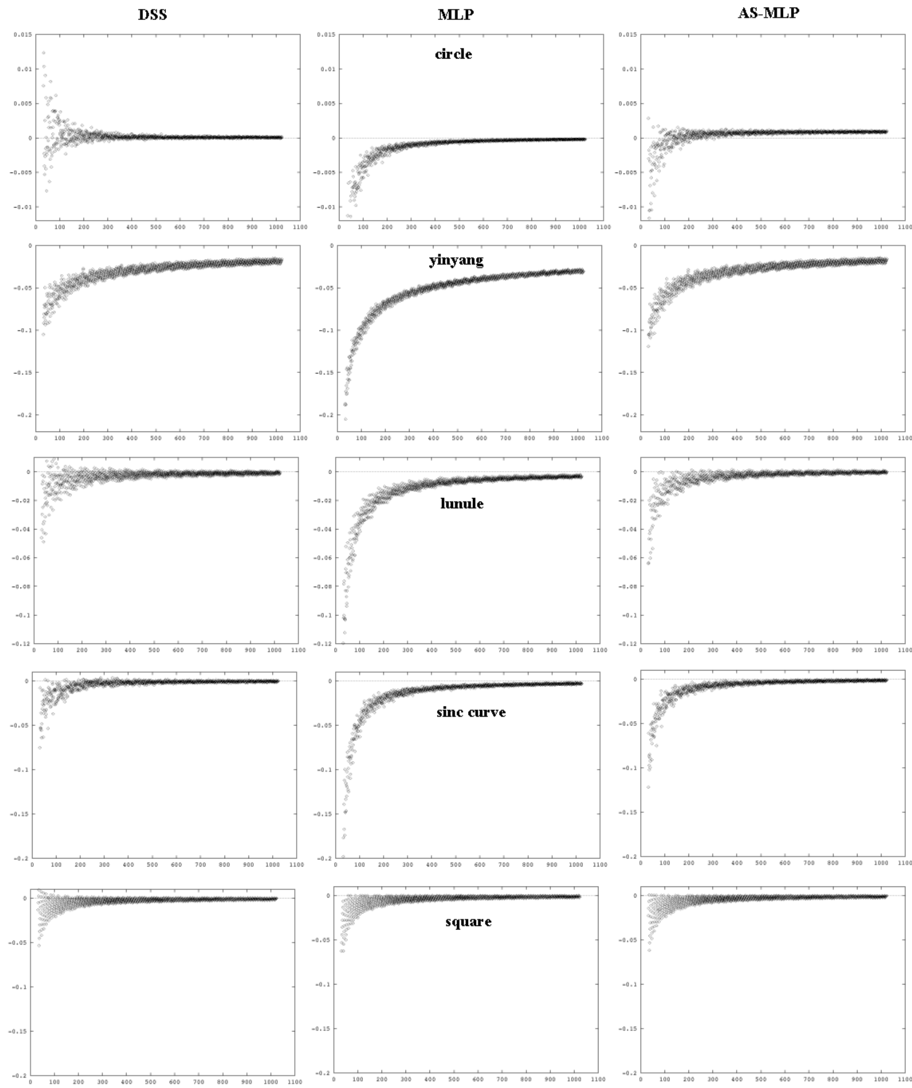

Fig. 7. Estimation errors: DSS in left column, MLP in the middle, and AS-MLP on the right; top row for circle, followed by lower part of yinyang, lunule, sinc curve, and square.

verify its practical performance we have implemented this scheme and tested it on various curves including non-convex ones. The results reflected the theoretical analysis of the three schemes, that is, DSS is the best in accuracy, and our scheme is in the middle. The AS-MLP approximation curves are smoother (see our definition above) than the MLP or DSS curves,

Acknowledgment

The used C++ programs for DSS and MLP are those implemented for and described in [4], and the test data set has been copied from [5].

References

1. T. Asano, Y. Kawamura, R. Klette, and K. Obokata. A new approximation scheme for digital objects and curve length estimations. in *Proc. of Image and Vision Computing New Zealand*, Hamilton, pages 26–31, 2000.
2. T. Buelow and R. Klette. Rubber band algorithm for estimating the length of digitized space-curves. in *Proc. IEEE Conf. ICPR*, Barcelona, Vol. III, pages 551-555, 2000.
3. A. Hübler, R. Klette, and K. Voss. Determination of the convex hull of a finite set of planar points within linear time. *EIK*, 17:121–139, 1981.
4. R. Klette, V. Kovalevsky, and B. Yip. On the length estimation of digital curves. in *Proc. Vision Geometry VIII*, Denver, SPIE-3811, pages 117–129, 1999.
5. R. Klette and Ben Yip. The length of digital curves. *Machine GRAPHICS & VISION*, 9:673–703, 2000.
6. R. Klette and J. Žunić. Multigrid convergence of calculated features in image analysis. *J. Mathem. Imaging and Vision*, 173–191, 2000.
7. V. Kovalevsky and S. Fuchs. Theoretical and experimental analysis of the accuracy of perimeter estimates. in *Robust Computer Vision* (W. Förstner, S. Ruwiedel, eds.), Wichmann, Karlsruhe, pages 218–242, 1992.
8. A. Rosenfeld. Digital straight line segments. *IEEE Trans. Comp.*, 28:1264–1269, 1974.
9. F. Sloboda, B. Zaťko, and P. Ferianc. Minimum perimeter polygon and its application. in *Theoretical Foundations of Computer Visi/-on* (R. Klette, W.G. Kropatsch, eds.), Mathematical Research **69**, Akademie Verlag, Berlin, pages 59-70, 1992.
10. F. Sloboda, B. Zaťko, and J. Stoer. On approximation of planar one-dimensional continua. in *Advances in Digital and Computational Geometry*, (R. Klette, A. Rosenfeld and F. Sloboda, eds.) Springer, pages 113–160, 1998.