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

CITR-TR-105   December 2001

A Comparative Evaluation of Length Estimators

David Coeurjolly[1] and Reinhard Klette[2]

Abstract

Relative convex hulls have been suggested for multigrid-convergent surface area estimation. Besides the existence of a convergence theorem there is no efficient algorithmic solution so far for calculating relative convex hulls. This article discusses an approximative solution based on minimum-length polygon calculations. It is illustrated that this approximative calculation also proves (experimentally) to provide a multigrid convergent measurement.

[1] Laboratoire ERIC, Université Lumière Lyon 2, 5 av. Pierre-Mendès France, 69676 Bron CEDEX, France. david.coeurjolly@univ-lyon2.fr

[2] Center for Image Technology and Robotics Tamaki Campus, The University of Auckland, Auckland, New Zealand. r.klette@auckland.ac.nz

A Comparative Evaluation of Length Estimators

David Coeurjolly
Laboratoire ERIC
Université Lumière Lyon 2
5 av. Pierre-Mendès France
69676 Bron CEDEX, FRANCE
david.coeurjolly@univ-lyon2.fr

Reinhard Klette
CITR Tamaki, The University of Auckland
Tamaki Campus, Building 731
Auckland, New Zealand
R.klette@auckland.ac.nz

Abstract

The paper compares previously published length estimators having digitized curves as input. The evaluation uses multigrid convergence (theoretical results and measured speed of convergence) and further measures as criteria. The paper also suggests a new gradient-based method for length estimation.

1 Introduction

The digitization of curves or boundaries has been studied in image analysis for about 40 years [1]. Since these first studies, many algorithms have been proposed to estimate the length of a digitized curve. Some approaches are based on local metrics such as the chamfers metrics, other approaches are based on polygonalizations of digital curves, e.g. directed on subsequent calculations of maximum-length digital straight segments (DSSs) or of minimum length polygons (MLPs), and we also propose a new length estimator based on calculated gradients.

The computational problem is as follows: The input is a sequence of chain codes $i(0), i(1), \ldots$ with $i(k) \in A = \{0, \ldots, 7\}$, $k \geq 0$. An *off-line DSS recognition algorithm* decides for finite words $u \in A^\star$ whether u is a DSS or not. An *on-line DSS recognition algorithm* reads successive chain codes $i(0), i(1), \ldots$ and specifies the maximum length $k \geq 0$ such that $i(0), i(1), \ldots, i(k)$ is a DSS, and $i(0), i(1), \ldots, i(k), i(k+1)$ is not. An off-line algorithm is *linear* iff it runs in $\mathcal{O}(n)$ time, i.e. it performs at most $\mathcal{O}(|u|)$ basic computation steps for any input word $u \in A^\star$. An on-line algorithm is linear iff it uses *on the average* a constant number of operations for any incoming chain code symbol.

An obvious benefit of local metrics based approaches is that they support linear on-line implementations.

Linear off-line algorithms for DSS recognition were published in 1981 in [2] and in 1982 in [5]. A linear off-line algorithm for cellular straight segment recognition, based on convex hull construction, is briefly sketched in [4]. Two linear on-line algorithms for DSS recognition were published in 1982 in [3]; one of them is an on-line version of the off-line algorithm published in [2].

The general problem of decomposing a digital curve into a sequence of DSSs, which includes DSS recognition as a subproblem, is discussed in, e.g. [9, 10, 13, 20]. Obviously, linear on-line DSS recognition algorithms will support linear decomposition algorithms, but linear off-line algorithms will only allow quadratic run-time behavior.

An MLP-approximation [15, 21] calculates a minimum-length polygon circumscribing a given (closed) inner boundary (given by a sequence of chain codes), and being in the interior of an outer boundary (typically in Hausdorff-Chebyshev distance 1 to the inner boundary). Gradient-based length estimation uses input data as in case of DSS approaches. The notions on-line, off-line, or linear time apply for MLP or gradient-based approaches as defined for the case of DSS algorithms.

In this article, we present a comparative evaluation of length estimators. We are especially interested in evaluating these algorithms (and underlying methodologies) with respect to the accuracy of length estimation. Multigrid convergence is one option of characterizing this accuracy, and experimental studies provide another way for performance testing. Our experiments are directed on illustrating accuracy and stability of the chosen algorithms on convex and non-convex curves (see Fig. 1 for the used test data). These Jordan curves are digitized for increases in grid resolution, allowing to study and illustrate experimental multigrid convergence. In our experiments we use a grid size ranging from 30 to 1000.

Figure 1. Test data set as proposed in [20].

Our studies on multigrid convergence require digitizations of planar Jordan curves up to a given grid resolution: we assume an orthogonal grid with grid constant $0 < \theta \leq 1$ in the Euclidean plane $\mathbb{R}^2$, *i.e.* θ is the uniform spacing between grid points parallel to one of the coordinate axes. Let $r = 1/\theta$ be the *grid resolution*, and the *r-grid* $\mathbb{Z}_r^2$ has resolution r, defined by *r-points* whose coordinates are $(\theta \cdot i, \theta \cdot j)$, with $i, j \in \mathbb{Z}$.

Now, we consider a Jordan curve $\gamma : [0,1] \to \mathbb{R}^2$, bounding a set S. Let $D_r(\gamma)$ be a *r-digitization* of γ in $\mathbb{Z}_r^2$, defined by *r-grid-intersection digitization*, or by a digital border in $\mathbb{Z}_r^2$. Common models are Gauss digitization (i.e. union of all r-grid squares with centroid in S), and inner or outer Jordan digitization (i.e. union of all r-grid squares contained in the interior of S, or having a non-empty intersection with S). See, e.g., [23] for details.

We denote by $\mathcal{F}(\gamma) \in \mathbb{R}$ a property of curve γ, which is the length $l(\gamma)$ of γ in this article. We denote by $\mathcal{E}$ an estimated feature. Assume that $\mathcal{E}$ is defined for digitizations $D_r(\gamma)$, for $r > 0$. The estimated feature $\mathcal{E}$ is said to be *multigrid convergent* iff $\mathcal{E}(D_r(\gamma))$ converges to $\mathcal{F}(\gamma)$, for $r \to \infty$. More formally:

$$|\mathcal{E}(D_r(\gamma)) - \mathcal{F}(\gamma)| \leq \epsilon(r)$$

with $\lim_{r \to \infty} \epsilon(r) = 0$. The order $O(1/\epsilon(r))$ denotes the *speed* of this convergence. Multigrid convergency of estimated features is a standard constraint in numerical mathematics for discrete versions of 'continuous' features.

2 Local metrics

Local metrics were historically the first attempts towards a solution of the length estimation problem. These algorithms can be viewed as shortest path calculations in weighted adjacency graphs of pixel locations. Weights have been designed with the intention of approximating the Euclidean distance. For example, horizontal and vertical moves in the orthogonal r-grid may be weighted by $1/r$, and diagonal moves may be weighted by $\sqrt{2}/r$. More generally, a chamfer metrics definition first lists elementary moves and then associates weights to each move, see [6, 7, 11].

In order to make length estimation as accurate as possible, the use of statistical analysis has been suggested to find those weights which minimize the mean square error between estimated and true length of a straight segment. For example, [7] presents a best linear unbiased estimator (BLUE for short) for straight lines, and defines the following length estimator:

$$\mathcal{E}_{BLUE}(D_r(\gamma)) = \tfrac{1}{r} \cdot (1.059 \cdot n_i + 1.183 \cdot n_d) \,, \quad (1)$$

where n_i is the number of isothetic steps and n_d of diagonal steps in the r-grid. We include this estimator into our comparative study. It ensures a superlinear convergence $O(r^{-1.5})$ of asymptotic length estimation [10] in case of digitized straight lines, but fails to be multigrid convergent for more general situations of digitized arcs or curves.

3 Polygonal approaches

Many algorithms have been published for the DSS recognition problem. These approaches are based on characterizations of digital lines, such as syntactic chain code properties [2, 5], arithmetical properties defining tangential lines [13], properties of feasible regions in the (dual) parameter space [8, 10], or use linear programming tools such as the Fourier-Motzkin algorithm [14]. All these algorithms present a solution for deciding whether a given sequence of r-grid points is a DSS, or even for segmenting a digital curve into a sequence of maximum-length DSSs. The length estimator $\mathcal{E}_{DSS}$ is then defined by the length of the obtained poly-line.

In our comparative study we include two representative implementations of DSS-based length estimators: if the digital curve is defined as an 8-curve, we use the Debled-Reveillès algorithm [13] and call it the $\mathcal{E}_{DR-DSS}$ estimator. If the digital curve is defined as a 4-curve, we consider a length estimator based on Kovalevsky's algorithm [9] and call it the $\mathcal{E}_{VK-DSS}$ estimator. We use the algorithm as detailed in [20].

These two DSS-based length estimators are known to be multigrid-convergent for convex Jordan curves γ [12, 23, 25].

MLP-based length estimators consider a situation where a given simple digital curve C is described by two discrete curves γ_1 and γ_2, bounding sets S_1 and S_2 respectively, such that S_2 is contained in the interior S_1° of set S_1, and C is contained in $B = S_1 \setminus S_2^\circ$. The task consists in calculating that MLP which is contained in B and circumscribes γ_2. The length estimator $\mathcal{E}_{MLP}$ is then defined by the length of this (uniquely defined [15, 18]) MLP.

In our comparative study we include two representative implementations of MLP-based length estimators: the *grid-continua MLP* approach of [15, 18] has been derived for the model of using inner and outer Jordan digitization, and defining B to be the difference set between outer and inner Jordan digitization. We use the MLP algorithm as reported in [20]. We call it this $\mathcal{E}_{SZ-MLP}$ estimator. As another MLP-method we include the *approximation-sausage MLP* approach of [21, 24] defining the $\mathcal{E}_{AS-MLP}$ estimator. We use the algorithm provided by the authors of this approach.

These two MLP-based length estimators are known to be multigrid-convergent for convex Jordan curves γ [15, 21, 24].

4   New gradient-based approach

We use a gradient integration process to estimate the length of a curve. Let $\vec{n} : [0,1] \to \mathbb{R}^2$ denote the gradient or normal vector field associated with curve γ. The length of γ can be expressed as

$$l(\gamma) = \int_\gamma \vec{n}(s)ds \qquad (2)$$

The main idea of our gradient-based approach consists in using discrete estimates of products $\vec{n}(s)ds$.

The discrete tangent on a digital curve was proposed in [16], and it is based on a chosen DSS definition (see Fig. 2): the discrete tangent at point p of a digital curve is the longest DSS centered at point p. Note that this definition is applicable whatever DSS definition (and related recognition algorithm) is chosen. A straightforward application of a linear on-line DSS recognition algorithms (at any point p) leads to an $O(n^2)$ solution. However, the optimization proposed in [19] allows to compute all tangents in linear time, assuming a cellular approach (i.e. a grid square has four $0-$cells as its vertices etc.) for defining pixel locations. The Vialard algorithm computes the discrete tangent and thus the discrete normal vector at each $0-$cell of the curve.

We define the normal vector $\tilde{n}$ associated to a $1-$cell as the mean vector of both vectors calculated at its two bounding $0-$cells. We also define an elementary normal vector n_{el} to a $1-$cell as the unit vector orthogonal to this $1-$cell (see Fig. 2). Hence, the discrete version of eq. (2) is:

$$\mathcal{E}_{TAN}(D_r(\gamma)) = \sum_{s\in\mathcal{S}} \tilde{n}(s) \cdot n_{el}(s) \qquad (3)$$

where '·' denotes the scalar product and $\mathcal{S}$ is the set of all $1-$cells of $D_r(\gamma)$, which is assumed to be an alternating sequence of $0-$cells and $1-$cells. The main idea
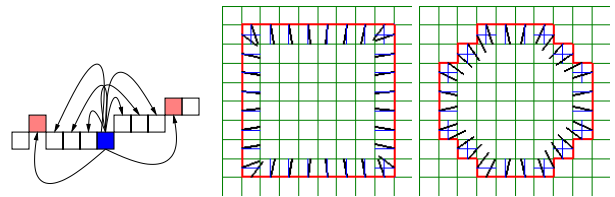

Figure 2. Illustration of the Discrete tangent calculus and examples of discrete normal vector computations for a square and a circle: black (bold) vectors are estimated normals and blue (thin) vectors are elementary normals.

of this approach is to compute the contribution of each $1-$cell to the global length estimation by projecting the $1-$cell according to the direction of the normal. In [26] it is shown that length estimator $\mathcal{E}_{TAN}$ is multigrid convergent for convex Jordan curves.

5   Evaluation

We compare the chosen length estimator algorithms based on practical experiments and available theoretical results. Note that all are linear on-line algorithms. Theoretical studies should answer the following questions:

multigrid : Is the estimator multigrid convergent for convex curves (if 'yes', we are also interested in an analysis of convergence speed)?

discrete : Does the core of the algorithm only deal with integers?

unique : Does the result depend on initialization?

3D extension : May the approach be extended to length estimation of digital curves in 3D space?

Table 1 informs on the situation.

We consider two measures: the relative error in percent between estimated and true curve length, and for DSS- and MLP-approaches, also a *trade-off measure* defined as the product of relative error times the number of generated segments. In [18, 20] it has been called the *efficiency of convergence*.

In Fig. 3 and Fig. 4, experimental convergence is evident for all methods. Errors are calculated for all curves, transformed into a mean value for a given grid size, and curves are generated by sliding means (of 30 values) along different grid sizes. The trade-off measure is presented for polygonal approaches. In Fig. 5, runtimes of polygonalization-based estimators compared to a local-metric estimator are presented.

Method	multigrid	discrete	unique	3D extension	References
DR-DSS	yes	yes	no	[25]	[13]
VK-DSS	yes	yes	no	-	[9]
SZ-MLP	yes	yes	yes	[22]	[15, 21]
AS-MLP	yes	no	no	-	[24]
TAN	yes	yes	yes	-	this article
Local metrics	no	depends	yes	[17]	[6, 7]

Table 1. Length estimators used in this comparison.
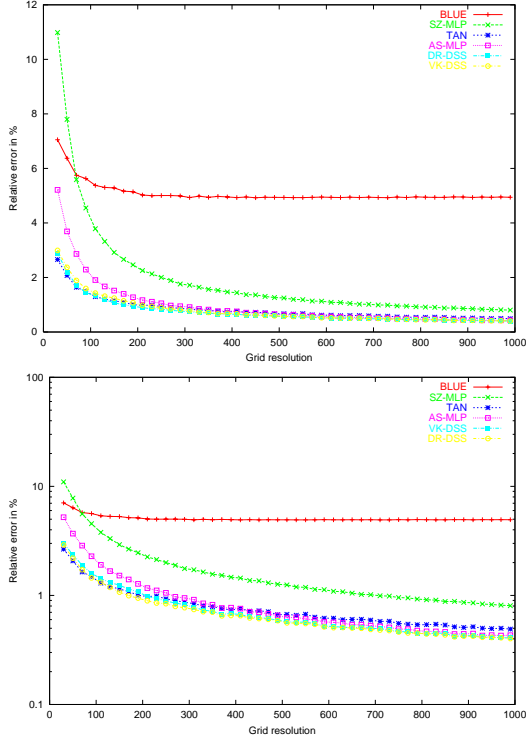
Figure 3. Evident multigrid convergence of length estimators: sliding means of relative errors and same results on a logarithmic scale.
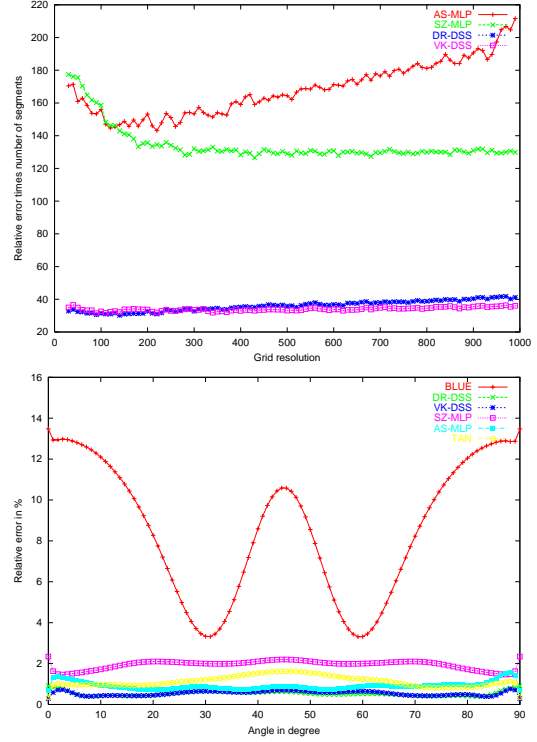

Figure 4. Up: **the trade-off measure for polygonal approaches,** $down$: **relative errors for a rotating square.**

Figure 4 shows the behavior of theses estimators when a square curve of fixed shape is rotated in a grid of resolution 128.

6 Conclusions

The experiments show that the local-metrics approach is not multigrid convergent for the test data set, but all five global approximation methods confirm known theoretical convergence results by experimental

evidence. Furthermore, the increase in run-times of the studied polygonal methods is only minor compared to that of local-metric algorithms. Hence, the use of an (incorrect) local-metric algorithm is also not justified by a run-time argument. The choice of a global method may depend on preferences defined by the context of an image processing software package, and the authors may recommend any of the five studied methods. Studies on test data might be useful for selecting the most efficient implementation for a given application context.
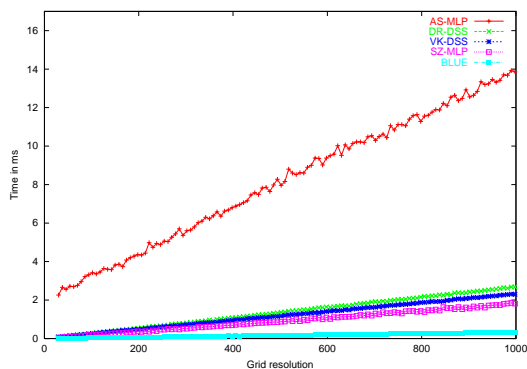
Figure 5. Run-times of polygonalization-based estimators on an Ultra 10 Sparc Workstation.

References

[1] H. Freeman. Techniques for the digital computer analysis of chain-encoded arbitrary plane curves. In *Proc. Natl. Elect. Conf.*, volume 17, pages 421–432, 1961.

[2] A. Hübler, R. Klette, and K. Voss. Determination of the convex hull of a finite set of planar points within linear time. *EIK*, 17:121–140, 1981.

[3] E. Creutzburg, A. Hübler and V. Wedler. On-line Erkennung digitaler Geradensegmente in linearer Zeit. In *Proc. GEOBILD'82*, Wiss. Beiträge der FSU Jena, pages 48–65, 1982.

[4] C. E. Kim. On cellular straight line segments. *Computer Graphics Image Processing*, 18:369–391, 1982.

[5] L. D. Wu. On the chain code of a line. *IEEE Trans. Pattern Analysis Machine Intelligence*, 4:347–353, 1982.

[6] G. Borgefors. Distance transformations in digital images. *Computer Vision, Graphics, and Image Processing*, 34:344–371, 1986.

[7] L. Dorst and A. W. M. Smeulders. Length estimators for digitized contours. *Computer Vision, Graphics, and Image Processing*, 40:311–333, 1987.

[8] J. Koplowitz and A. M. Bruckstein. Design of perimeter estimators for digitized planar shapes. *IEEE Trans. Pattern Analysis Machine Intelligence*, PAMI-11:611–622, 1989.

[9] V. A. Kovalevsky. New definition and fast recognition of digital straight segments and arcs. In *Proc. 10th Intl. Conf. on Pattern Recognition*, pages 31–34, 1990.

[10] L. Dorst and A.W.M. Smeulders. Discrete straight line segments: parameters, primitives and properties. In: R. Melter, P. Bhattacharya, A. Rosenfeld (eds): *Ser. Contemp. Maths.*, Amer. Math. Soc. 119:45–62, 1991.

[11] B. Verwer. Local distances for distance transformations in two and three dimensions. *Pattern Recognition Letters*, 12:671–682, 1991.

[12] V. Kovalevsky and S. Fuchs. Theoretical and experimental analysis of the accuracy of perimeter estimates. In *Robust Computer Vision*, Wichmann, Karlsruhe, pages 218–242, 1992.

[13] I. Debled-Rennesson and J. Reveillès. A linear algorithm for segmentation of digital curves. *Int. J. Pattern Recognition and Artificial Intelligence*, 9:635–662, 1995.

[14] J. Françon, J.-M. Schramm, and M. Tajine. Recognizing arithmetic straight lines and planes. In *Proc. DGCI*, LNCS 1176, 141–150. Springer, Berlin, 1996.

[15] F. Sloboda and B. Zaťko. On one–dimensional grid continua in R^2. Technical report, Institute of Control Theory and Robotics, Bratislava, 1996.

[16] A. Vialard. Geometrical parameters extraction from discrete paths. In *Discrete Geometry for Computer Imagery*, pages 24–35, 1996.

[17] A. Jonas and N. Kiryati. Length estimation in 3-D using cube quantization. *J. Mathematical Imaging and Vision*, 8:215–238, 1998.

[18] F. Sloboda, B. Zaťko, 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, Singapore, pages 113–160, 1998.

[19] F. Feschet and L. Tougne. Optimal time computation of the tangent of a discrete curve : Application to the curvature. In *Discrete Geometry for Computer Imagery*, pages 31–40, 1999.

[20] R. Klette, V. Kovalevsky, and B. Yip. Length estimation of digital curves. In *Vision Geometry VIII - SPIE*, pages 117–129, Denver, July 1999. (extended version: R. Klette and B. Yip. The length of digital curves. *Machine Graphics & Vision*, 9:673–703, 2000.).

[21] 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*, pages 26–31, New Zealand, Hamilton, 2000.

[22] R. Klette and T. Bülow. Minimum-length polygons in simple cube-curves. In *Discrete Geometry for Computer Imagery*, pages 467–478, 2000.

[23] R. Klette and J. Žunić. Multigrid convergence of calculated features in image analysis. *J. Mathematical Imaging Vision*, 13:173–191, 2000.

[24] T. Asano, Y. Kawamura, R. Klette, and K. Obokata. Minimum-length polygons in approximation sausages. In *Visual Form IV*, LNCS 2059, Springer, pages 103–112, 2001.

[25] D. Coeurjolly, I. Debled-Rennesson, and O. Teytaud. Segmentation and length estimation of 3d discrete curves. In *Digital and Image Geometry*. to appear, LNCS, Springer, 2001.

[26] D. Coeurjolly and O. Teytaud. Multigrid convergence of discrete differential estimators: discrete tangent and length estimation. Research Report Laboratoire ERIC, RR-0101, http://eric.univ-lyon2.fr/RR/RR-0101.ps.gz, 2001.