Measures of Evaluation

goal: comparative evaluation of length estimators with respect to different performance criteria (accuracy, stability, runtime, and so forth)

performance measures:

(1) the (absolute) *relative error* $|T - E|/T$ (in percent) between the estimated (E) and the true (T) curve length;

(2) if the length estimator is based on calculating a polygonal approximation of the curve, then the *tradeoff measure* = relative error $\times$ number of generated line segments; and

(3) *runtimes* measured with respect to a uniform implementation basis

The tradeoff measure characterizes the *efficiency of convergence* (assuming it actually happens): we 'pay' for a smaller relative error with a larger number of generated line segments, but this increase should stay in the order as the error decreases

Test Data

We select a class of test curves in the Euclidean plane (e.g., within the unit square $[0, 1] \times [0, 1]$) of known length, and digitize these at different grid resolutions (say, between $h = 32$ and $h = 1,024$).

Example:
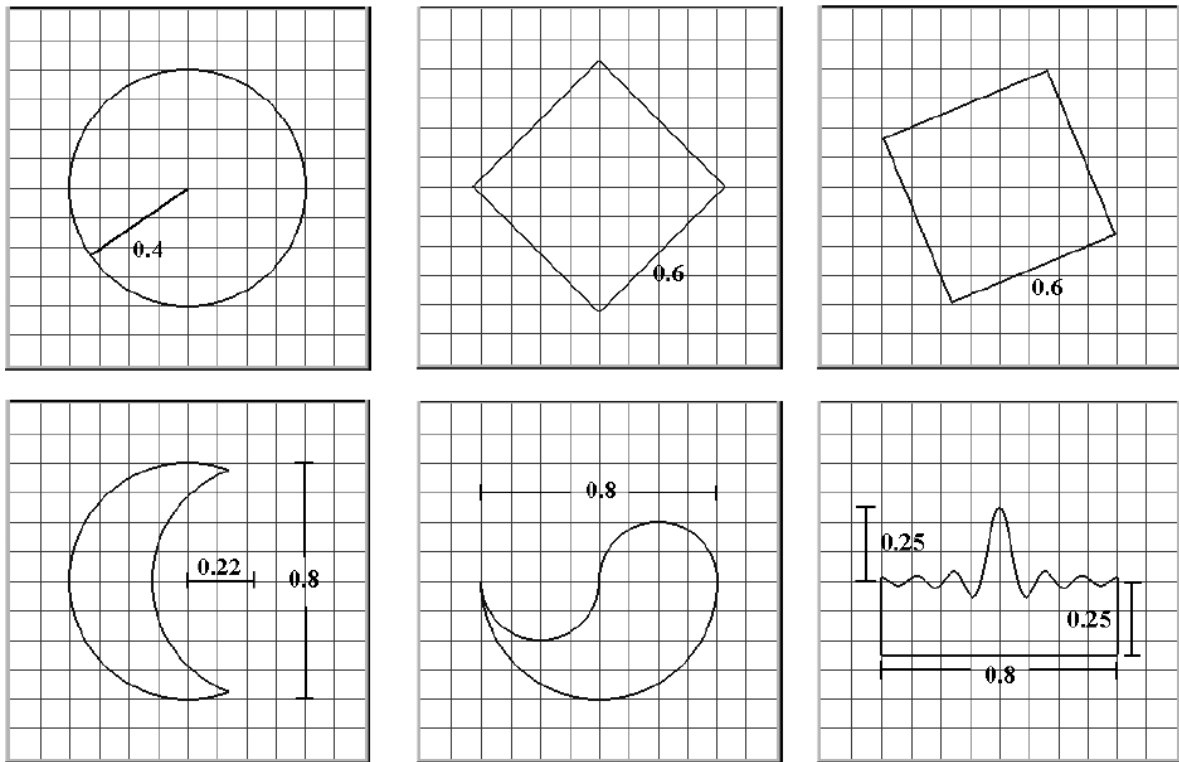
A circle, a square rotated by 45° and by 22.5^{o}, a lunule generated by two overlapping circles of identical size, the (lower) yin-part in the Chinese yinyang symbol, and the function graph of the *sinc function*

$$y = sin(16\pi \cdot x)/(64\pi \cdot x)$$

within a bounded interval symmetric to the y-axis.

The True Length ("Groundtruth")

For our experiments, we have to know the true perimeters of the chosen set of regions. In the example it is 0.8π for the disk, 2.4 for the squares, 0.8π for the lunule, 0.8π for the yin-region, and (calculated by numerical methods) $2.792696\ldots$ for the region partially bounded by the graph of the given sinc function.

Digitizations

Option 1: Gauss digitization, and use the border (i.e., grid-point model) for local methods or DSS-approximations, or the frontier (of the union of all 2-cells contained in $G_h(S)$) for 4-DSS-approximations or MLP calculations.

Option 2: approximate inner Jordan digitization by accepting all those squares for which all four vertices are in the given region; consider the frontier of the union of all identified 2-cells as input for 4-DSS-approximations or MLP calculations, or as 4-curve in the frontier grid and as input (after mapping it into an 8-curve) for local methods or DSS-approximations.

Option 3: calculate the grid-intersection digitization of the given curves (frontiers of regions)

The following test data are obtained by applying Option 1.

Sliding Mean

Experiments show that relative errors are not monotonously increasing or decreasing, for increases in grid resolution. It appears to be practical to apply a sliding mean defined by its size $2k + 1$: map all calculated relative errors e_h at grid resolution h into means of errors $e_{(h-k)}, \ldots, e_h, \ldots, e_{h+k}$.
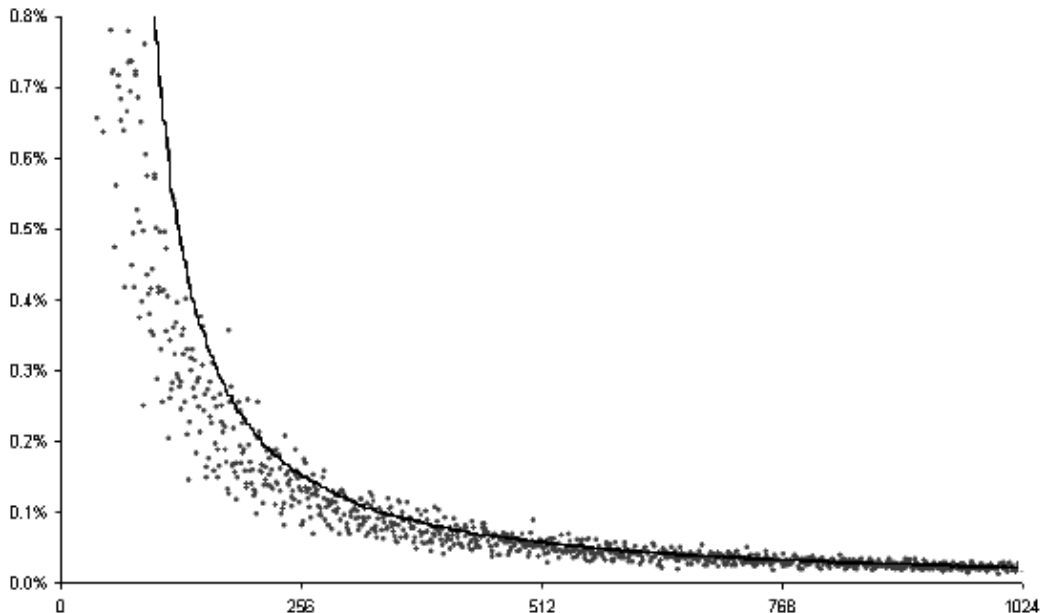

Relative errors (dots) for the circle and MLP, and the sliding mean (curve) using size $2k + 1 = 31$.

Combining measurements for several curves: first we calculate e_h as being the mean of all relative errors of all participating curves for grid resolution h, then we calculate the sliding mean of these combined relative errors

One More Algorithm

The following diagrams also illustrate the *approximating-sausage* estimator E_{aps}. This is a special variant of an MLP-based estimator.

Let the h-frontier of S be represented by $\Pi = \langle v_0, v_1, \ldots, v_{n-1} \rangle$ where the vertices are in clockwise order and the interior of S lies to the right. We define the *forward shift* $f(v_i)$ of v_i as the point on the edge (v_i, v_{i+1}) at distance $1/(2h)$ from v_i, and the *backward shift* $b(v_i)$ of v_i as the point on the edge (v_{i-1}, v_i) at distance $1/(2h)$ from v_i.
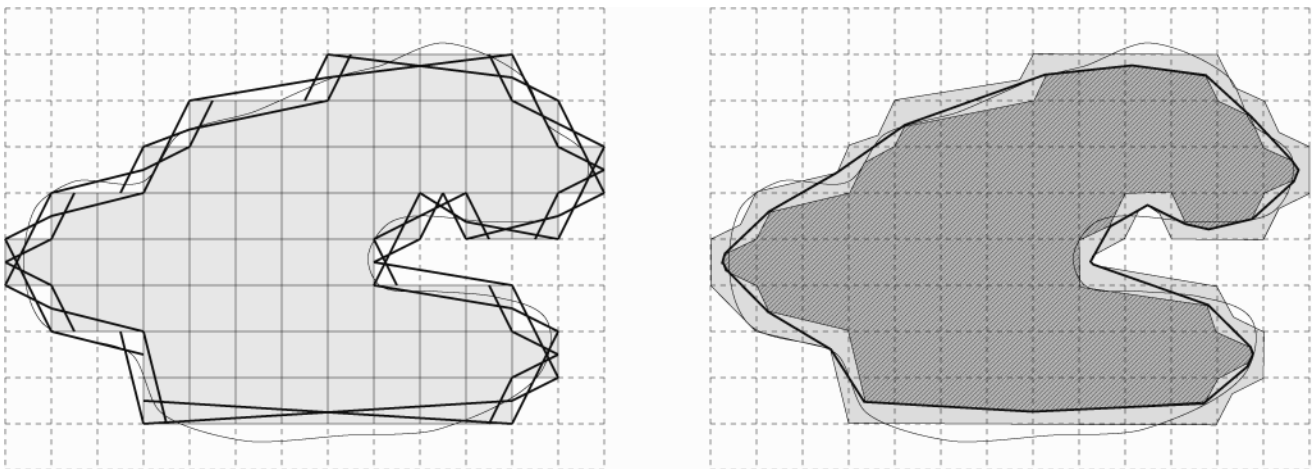
The approximating-sausage is a connected region that is topologically equivalent to an annulus (see textbook for details), and E_{aps} is defined by the total length of the MLP contained in this approximating sausage (and which appears in general to be a "smoother" approximation than the standard MLP discussed before).
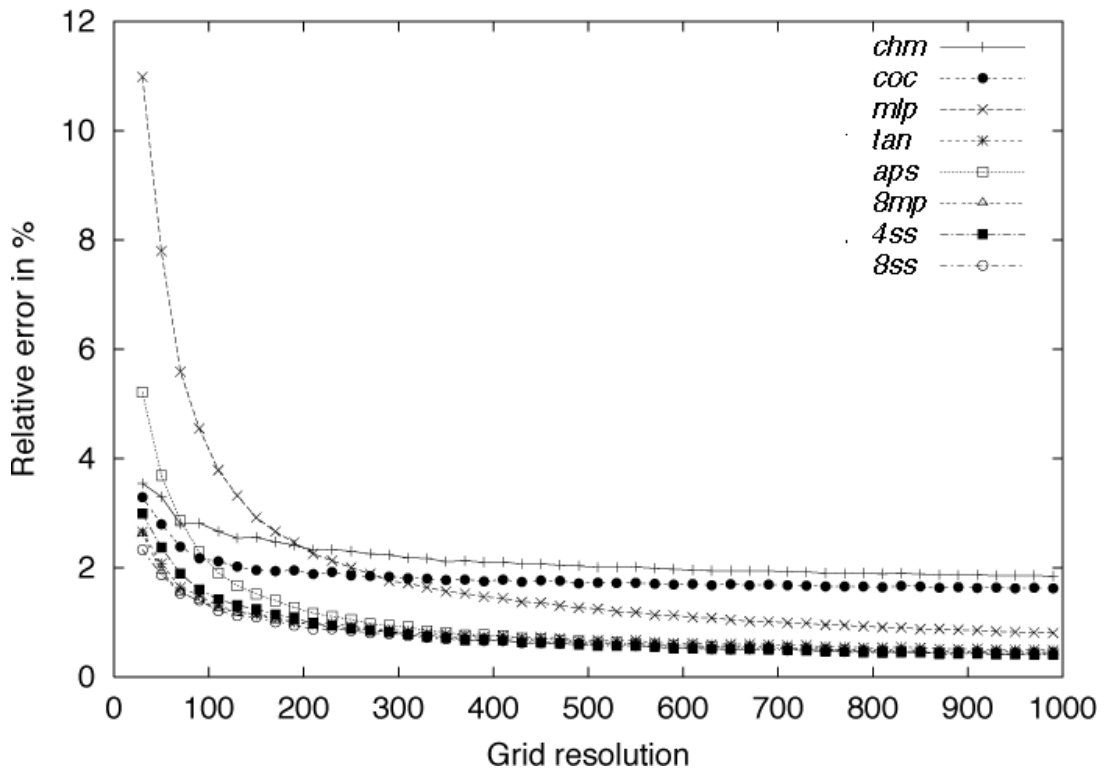
"Obvious convergence" of all estimators for our set of six test curves: however, for chm and coc, the errors do not converge to zero. The following logarithmic scale shows this even better:
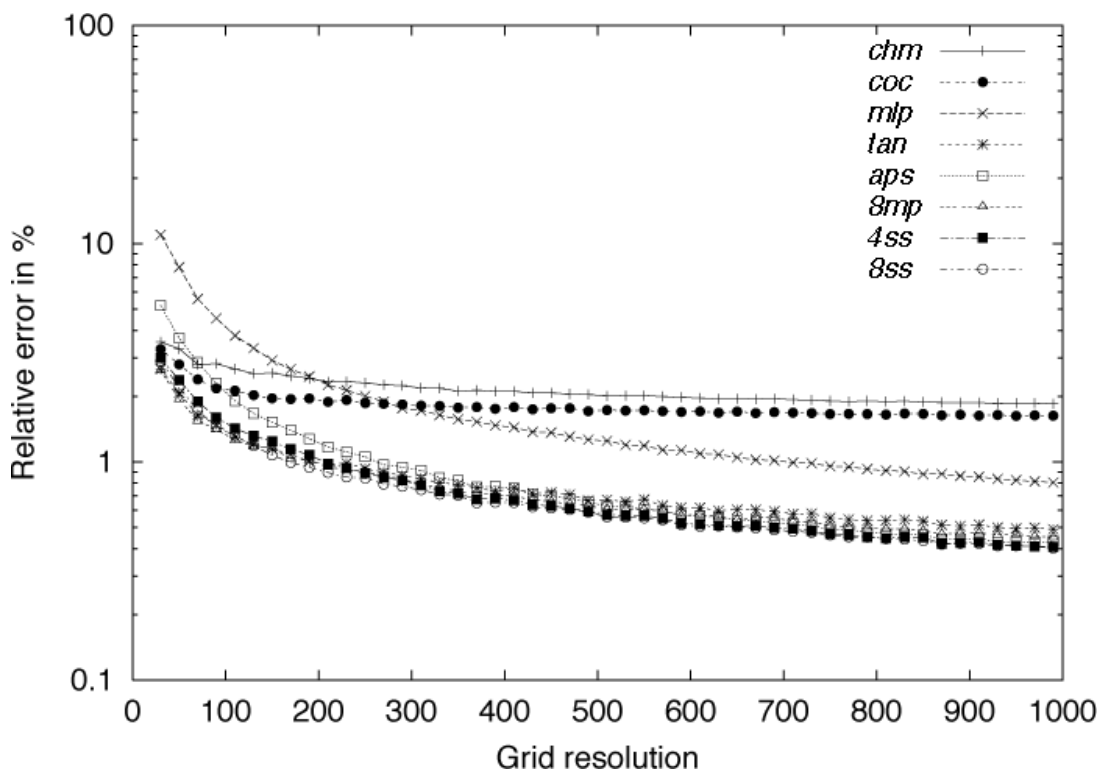
We study the **rotation invariance** of length measurement.

Here, a square of fixed size was rotated in a grid of resolution 128. The figure shows its estimated perimeter as a function of rotation angle. Except for chm and coc, the estimates are relatively orientation-independent.
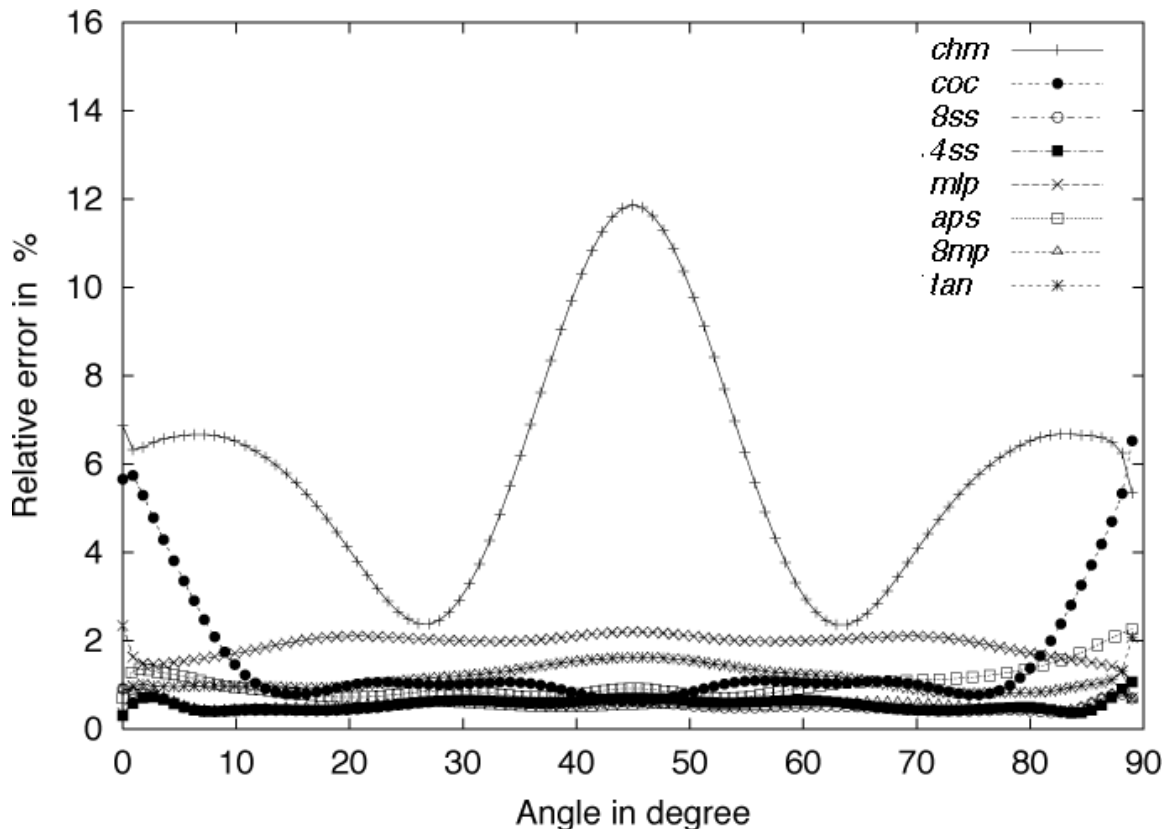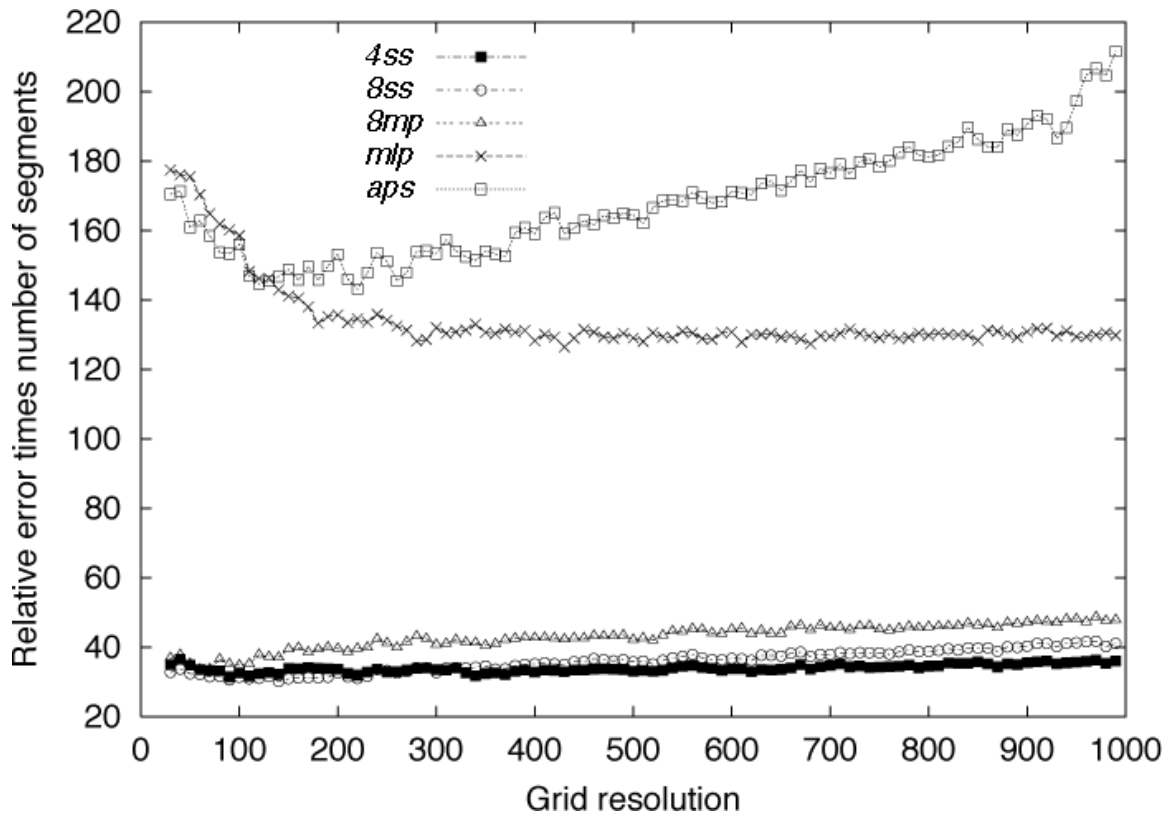
Trade-off plots for the DSS and MLP estimators (here, too, the errors are calculated for each curve and combined into a mean error for a given grid size, and the plots are generated by taking sliding means of size 31).

possible conclusion: 4ss reduces the error with the least relative increase in numbers of segments, and aps has a strong increase in numbers of segments (and provides a "smooth approximation" to the original curve)
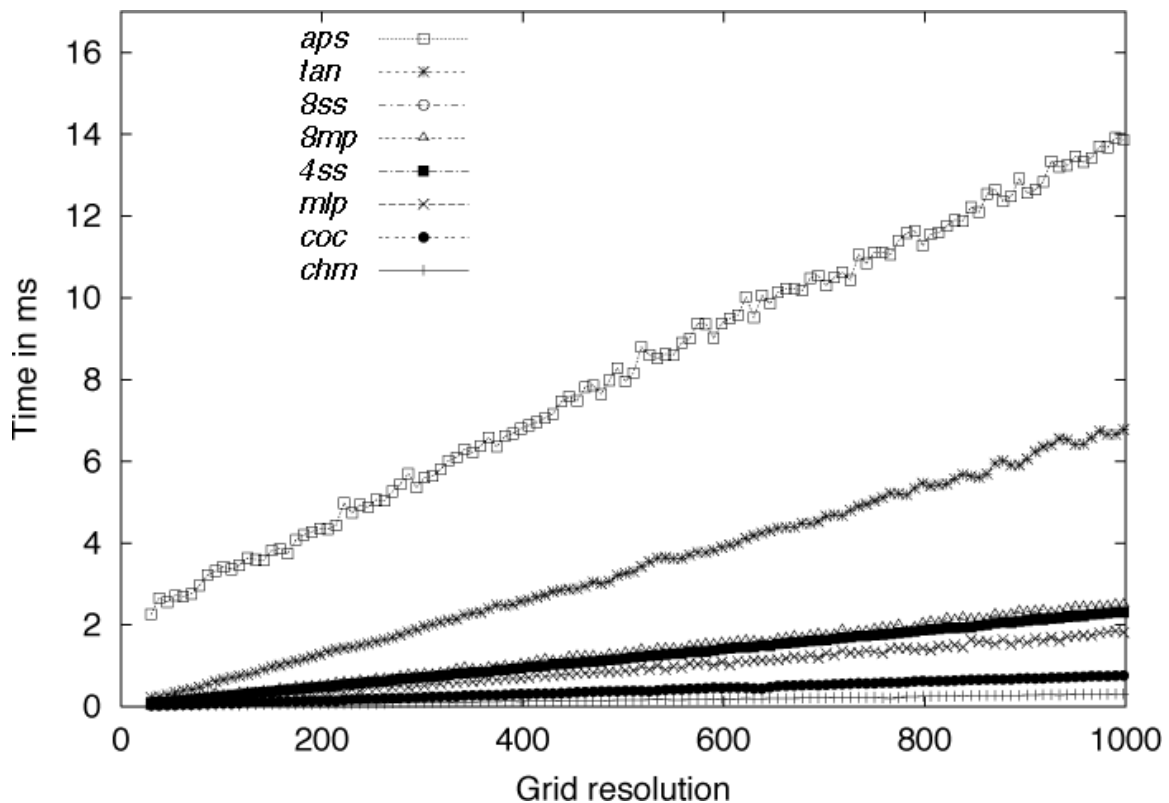
Runtimes of the DSS- and MLP-based estimators on an Ultra 10 Sparc workstation

Obviously, the local estimators are fastest.

MLP (implemented as described in this lecture before) is the fastest global estimator, but 4ss and 8ss come close to it.

The aps estimator has not yet been optimized; faster implementations may be possible.

Theoretically the tan estimator has a linear asymptotic runtime implementation, but the estimator that was tested had quadratic runtime; optimization is needed here, too.

Conclusions from Experiments

Local estimators are not multigrid convergent even for our simple test data set. However, they are (regarding the errors in length estimation) competitive for small grid resolutions. The non-invariance with respect to rotation (see example of square) is very critical for classifications of regions based on perimeter values.

For the five global estimators, we obtained experimental confirmations of known theoretic convergence results.

Interestingly, the runtimes of the polygonal DSS and MLP estimators were only slightly greater than those of the local estimators; hence the use of local estimators is not justified by a runtime argument.

The choice of a global estimator may depend on the available software. Studies involving more extensive test data might be useful in the selection of the most efficient estimator for a given application.

Conclusions from Theoretical Insights

multigrid convergence: Is the estimator multigrid convergent at least for convex curves? (If so, we are also interested in its convergence speed.)

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

unique: Is the result independent of initialization (starting point, tracing orientation, and so forth)?

3D extension: Can the estimator be extended to digital curves in 3D space?

Method	Multigrid	Discrete	Unique	3D extension
8ss	Yes	Yes	No	Exists
8mp	Yes	Yes	No	NYR
4ss	Yes	Yes	No	NYR
mlp	Yes	Yes	Yes	Difficult
aps	Yes	No	No	?
tan	Yes	Yes	Yes	NYR
local	No	Possibly	Yes	Exist

NYR = not yet reported, but should be straightforward

Difficult = there (not yet satisfying) approximative solutions

? = unclear how the "aps-idea" can be applied to curves in 3D space

Coursework

Related material in textbook: Section 10.3.4 (and optional Section 10.2.5 if interested in more details on the approximating sausage approach).

A.20. [5 marks] Consider a finite set of ellipses as test curves, specified by two radii a and b (let $a \leq b$) and an angle α of the main axis (assuming $a < b$) with the positive x-axis. Assume that these are centered in the unit square $[0, 1] \times [0, 1]$.

Digitize these curves for varying grid resolutions h, with $30 \leq h \leq 1000$.

Apply two length estimators of your choice, one local and one global method, for these digital curves and discuss

(i) the impact of decreasing ratios a/b on relative errors, and

(ii) rotation invariance of length estimation for an ellipse satisfying $a = b/5$.