Curve Encoded Compression and Transmission

Sending Document Images to Low-Bandwidth Users

Document Images

Digital Libraries

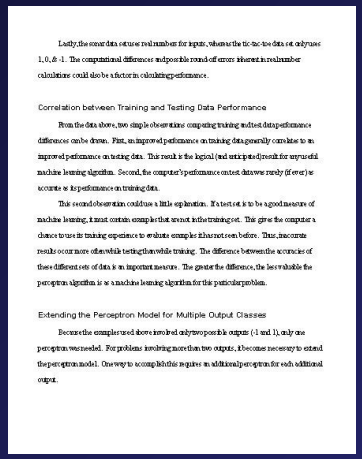
- Wide Distribution
- Easy Access
- Less “Shelf” Storage

Digital Media

- Text Transcripts
- Document Images

Genealogical Document Images

- Handwriting (no OCR)
- Mostly Bi-tonal (but needs grayscale)
- “Browsing” Operations



Lastly, the sonar data returns real number for inputs, whereas the tic-tac-toe data set only uses 1, 0, & -1. The computational differences and possible round-off errors inherent in real-number calculations could also be a factor in calculating performance.

Correlation between Training and Testing Data Performance

From the data above, two simple observations comparing training and test data performance differences can be drawn. First, an improved performance on training data generally correlates to an improved performance on testing data. This result is the logical (and anticipated) result for any useful machine learning algorithm. Second, the computer's performance on test data was rarely (if ever) as accurate as its performance on training data.

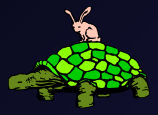
This second observation could use a little explanation. If a test set is to be a good measure of machine learning, it must contain examples that are not in the training set. This gives the computer a chance to use its training experience to evaluate examples it has not seen before. Thus, inaccurate results occur more often while testing than while training. The difference between the accuracies of these different sets of data is an important measure. The greater the difference, the less valuable the perceptron algorithm is as a machine learning algorithm for this particular problem.

Extending the Perceptron Model for Multiple Output Classes

Because the examples used above involved only two possible outputs (-1 and 1), only one perceptron was needed. For problems involving more than two outputs, it becomes necessary to extend the perceptron model. One way to accomplish this requires an additional perceptron for each additional output.

Challenges

Large File Sizes



Slow Connection Speeds

How do we give researchers the ability to browse through family history document images quickly despite “low bandwidth” connection speeds?

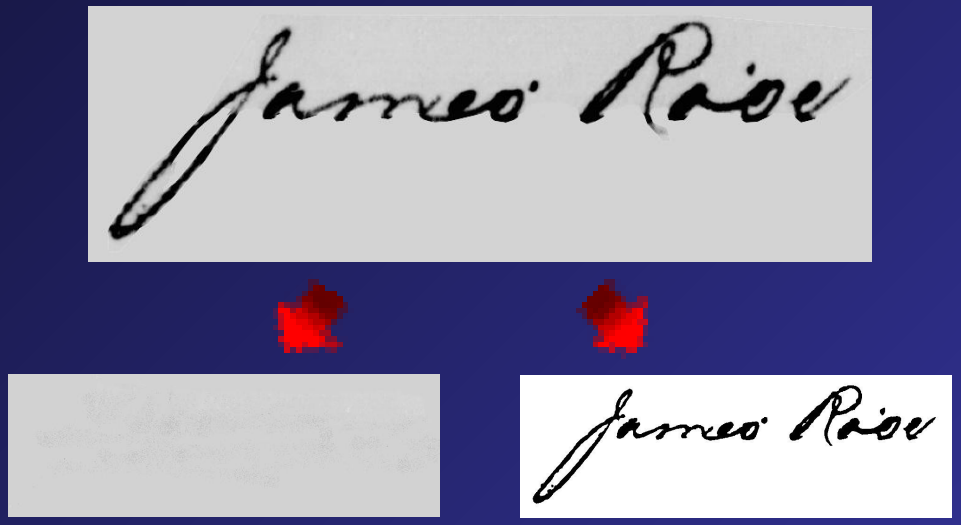
Approach One: Image Compression

Transform
- JPEG
- Wavelet

Context
- GIF
- CCITT-G4

Codebook
- JBIG2
- JB2

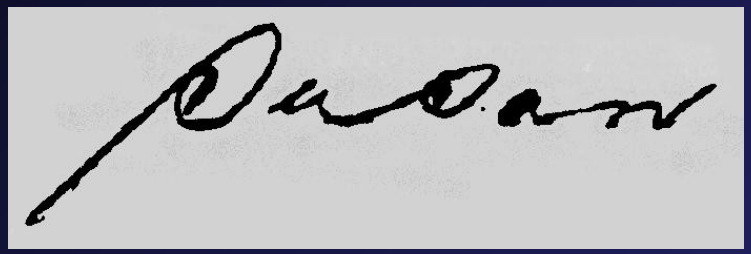
"Hybrid" Strategies
- DjVu (Bottou et al. '98)
- SLIm (http://research.microsoft.com/dpu/)
- DigiPaper (Huttenlocher et al. '00)

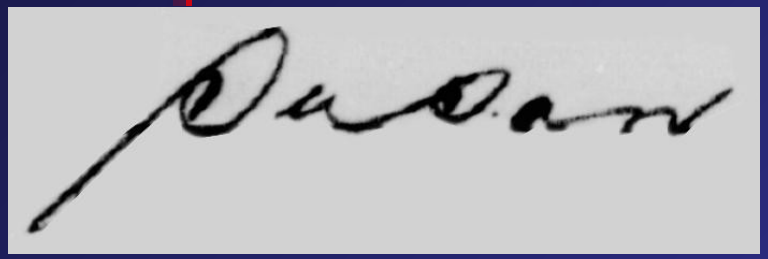


Approach Two: Progressive Transfer

Content Progressive

Example: DjVu (Bottou et al. '98)

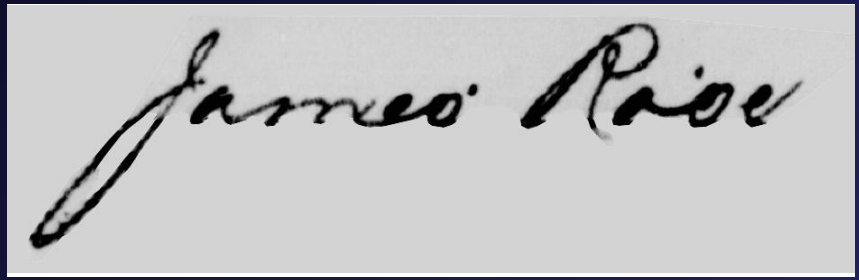




Quality Progressive

Example: JITB (Kennard '03)

Curve Encoded Compression and Transmission (CECAT)



Compression

1) Extract Foreground Mask from Image
2) Detect and Mark the Contours
3) Encode Contours as 1st – 3rd Order Bezier Curves
4) Group Curves by Locality & Priority

Transmission

1) Transfer & Fill Most Important Contours
2) Transfer Rest of Foreground
3) Add Grayscale Variations to Foreground
4) Transfer Background Color Image

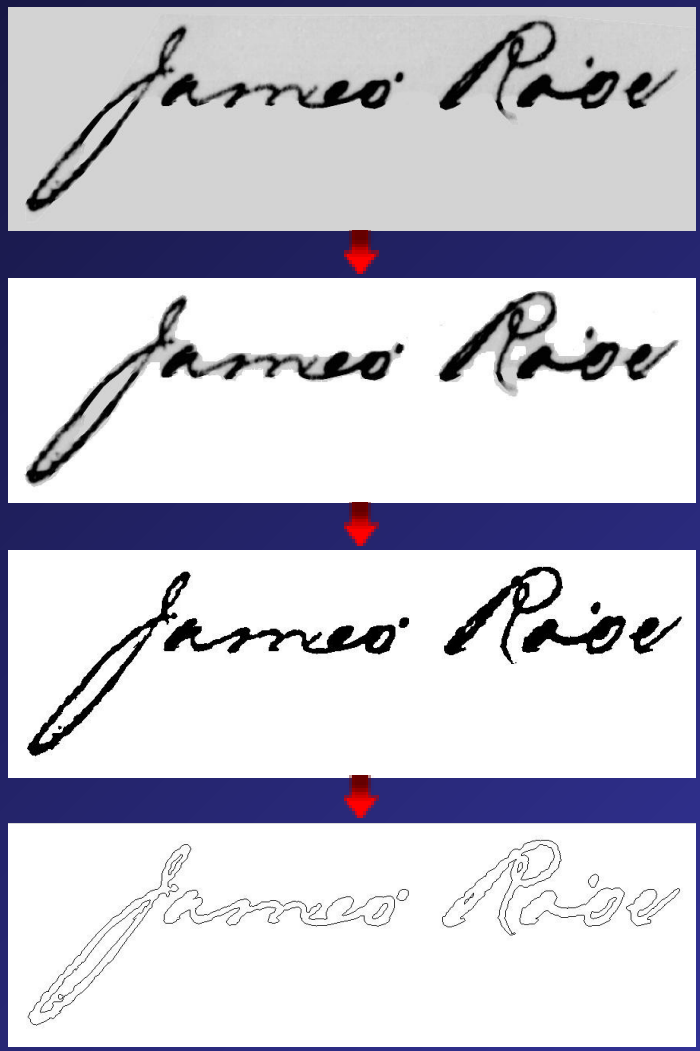
Preprocessing: From Image to Contours

1) Convert to Grayscale

2) Apply Median Filter
(Hutchison ’04)

3) Thresholding Operation
(Niblack ’85)

4) Contour Detection
(Witten et al. ’94)

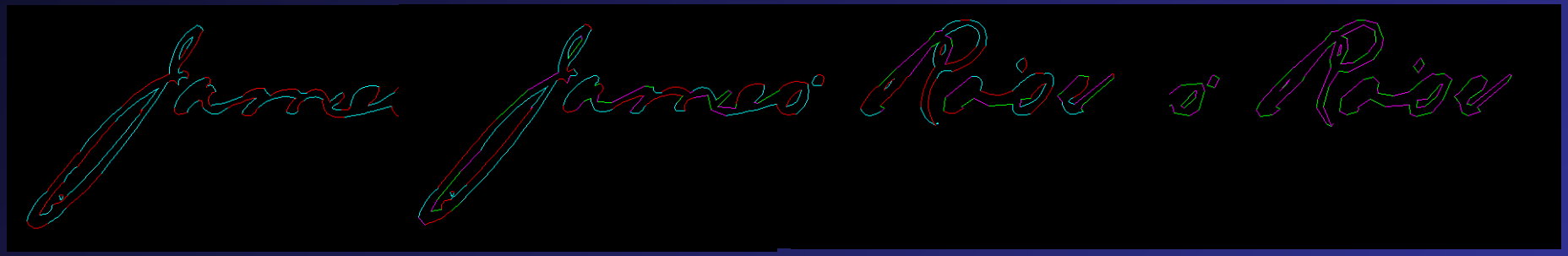


Finding a Parametric Fit to Contours

Curve Order	Bezier Curve Parametric Representation	File Size
1st (Line)	$\mathbf{p}(u) = (1-u)\mathbf{p}_0 + u\mathbf{p}_1$	4 bytes
2nd (Quadratic)	$\mathbf{p}(u) = (1-u)^2\,\mathbf{p}_0 + 2u(1-u)\mathbf{p}_1 + u^2\,\mathbf{p}_2$	6 bytes
3rd (Cubic)	$\mathbf{p}(u) = (1-u)^3\,\mathbf{p}_0 + 3u(1-u)^2\,\mathbf{p}_1 + 3u^2(1-u)\mathbf{p}_2 + u^3\,\mathbf{p}_3$	8 bytes

$\mathbf{p}(u)$ = points on the curve ($u \in [0, 1]$) $\mathbf{p}_n$ = Bezier control points

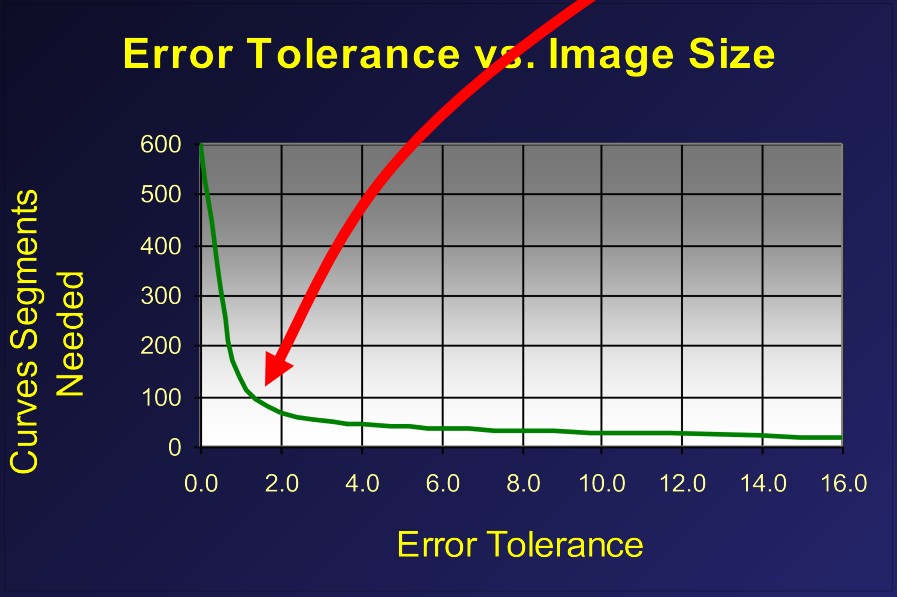
Results Using Least-Squares-Best-Fit Algorithm

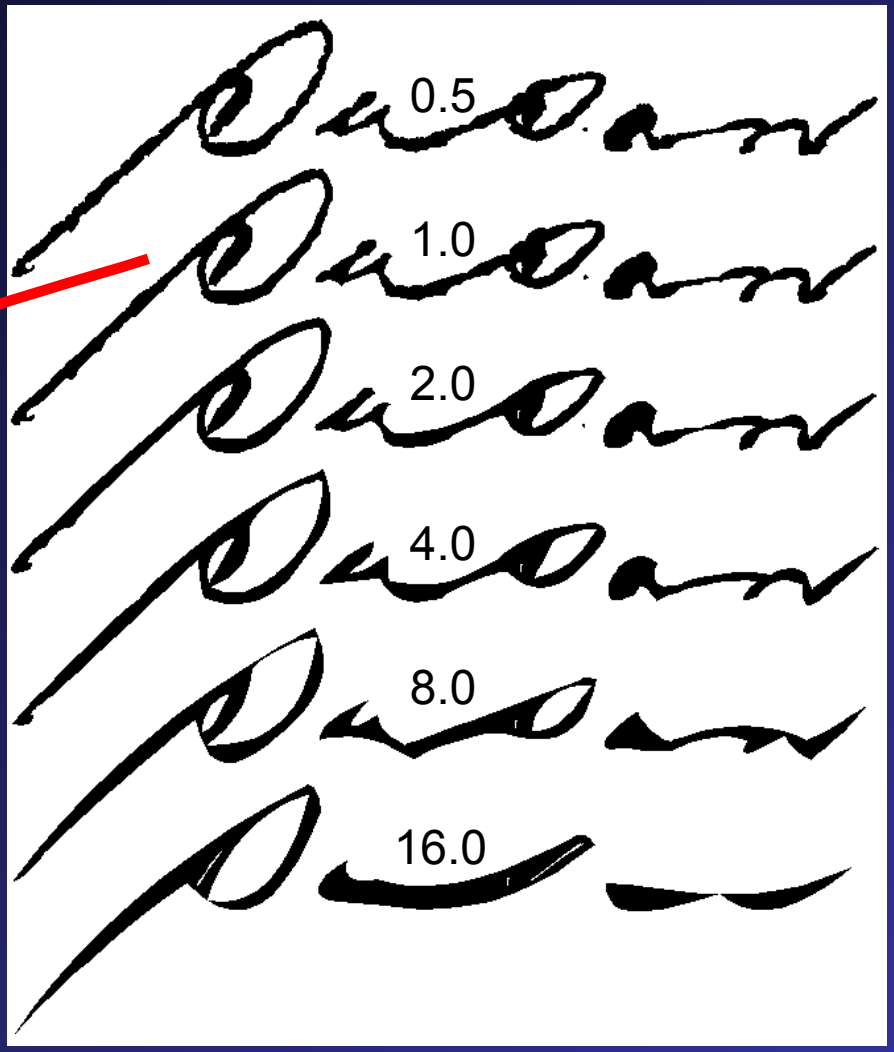


122 Quadratics (max 732 bytes) & 59 Lines (max 354 bytes) 77 Quadratics (max 462 bytes) & 22 Lines (max 132 bytes) 228 Quadratics (max 1368 bytes) & 59 Lines (max 354 bytes)

Lossy Compression: Error Tolerance

Error Metric: Maximum Pixel Distance Between Points on the Contour and the Parametric Curve



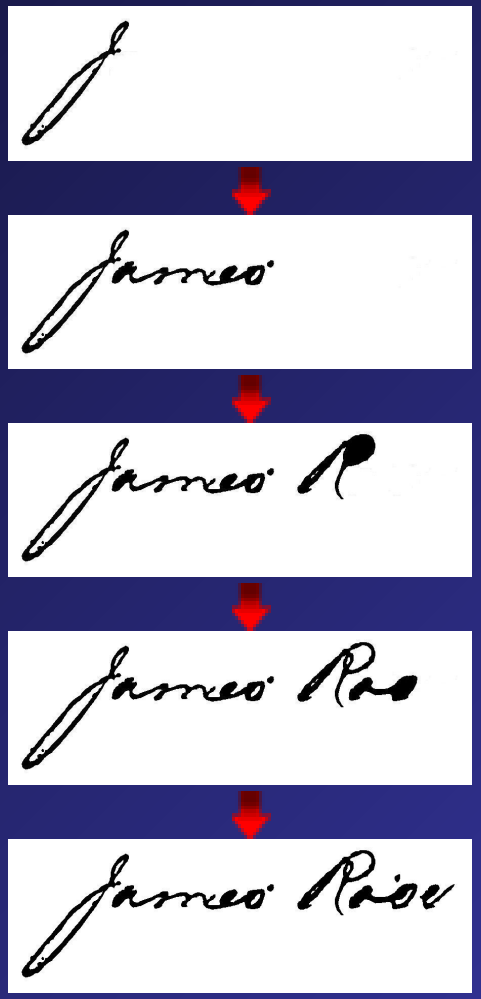


Progressive Transfer: Foreground

Encoding Strategy: Sort Parametric Curves According to Locality and/or Priority

Transfer Strategy: Send (and Fill) the Most Important Sets of Contours First

Demonstration

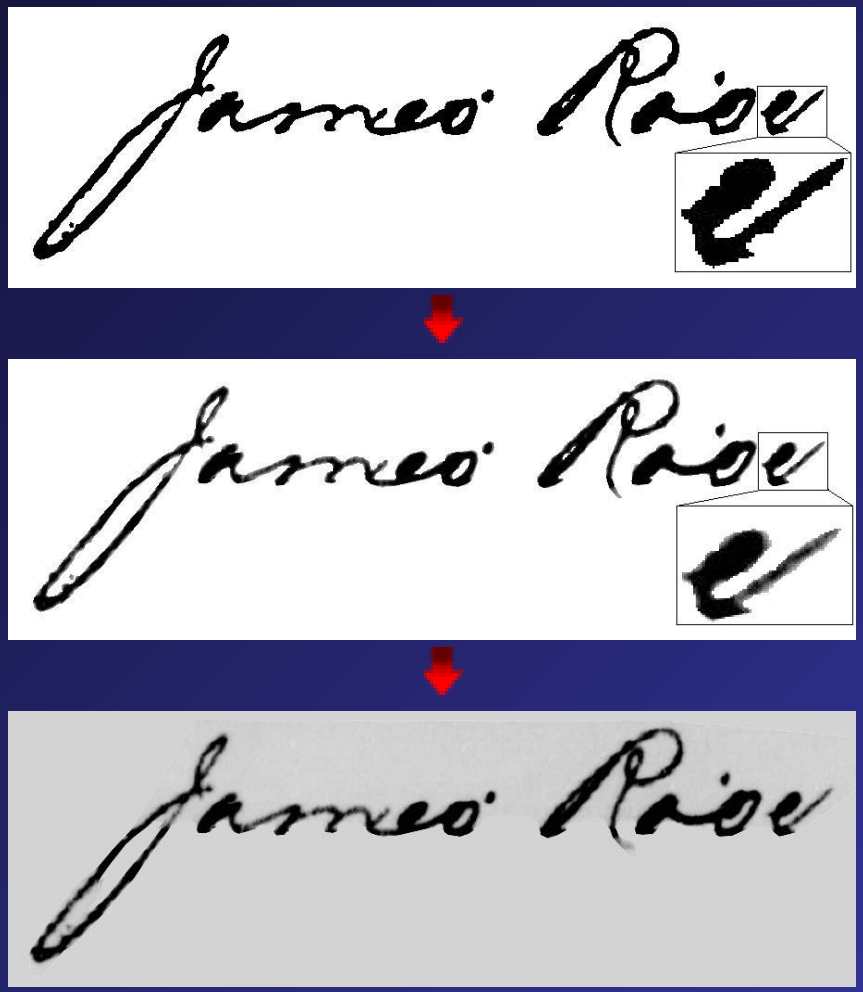


Progressive Transfer: Background

1) Foreground Mask Complete

2) Foreground Grayscale Data

3) Background Color Image



References

DjVu

- http://www.djvuzone.org/home.html

DigiPaper

- http://www.dlib.org/dlib/january00/moll/01moll.html

Contour Following

- Ian H. Witten et al. *Managing Gigabytes*. Van Nostrand Reinhold: New York. 1994

Niblack Thresholding

- Wayne Niblack. *An Introduction to Digital Image Processing*. Prentice-Hall International, 1985.

Just-In-Time-Browsing

- Douglas J. Kennard. *Just-In-Time Browsing for Digital Images*. Thesis Presented to BYU: February 2003

Quadratic Contour Compression

- Michael D. Smith. *Handwriting Compression using Quadratic Curves*. BYU CS 750 Project Write-Up. November 29, 2003

Median Filter Background Removal

- Luke A. D. Hutchison et al. *Fast Registration of Tabular Document Images Using Fourier-Mellin Transform*. In *Proceedings of DIAL04*, pages 253-269, January 2004.