

Curve Encoded Compression and Progressive Transfer for Document Images

Christopher Nelson and William Barrett
BYU Computer Science Department
Email: cbn7@email.byu.edu

Browsing Family History Document Images

By making scanned images of important genealogical documents available over the Internet, digital libraries can make Family History research much easier. Even now, many U.S. Census images are available for on-line indexing and viewing (see Figure 1.1). Imagine being able to access the contents of the Family History Library in Salt Lake from the comfort of your own home!

Browsing through document images, however, is a time consuming process. Most conventional methods for downloading images require the researcher to wait until the entire image has been downloaded before it can be displayed. Because most genealogical research involves scanning through a large number of images very quickly, a delay of even five or six seconds per image is unacceptable. To complicate matters, despite the increased use of broadband services, in 2002 over 70% of the all Internet access made from homes used dial-up connections (56K modems) [1].

Improved image compression is one way to help researchers download and browse through document images quickly. The advantage of this approach is quite obvious: smaller file sizes means shorter download time. Current document image compression techniques such as DjVu separate the image into text (foreground) and images (background). By applying image compression techniques specifically designed for each part of the image, DjVu gives very good compression results [2].

To send “reading level detail” interactively over low-bandwidth connections, however, much smaller file sizes are needed than can be provided by even the most sophisticated compression strategies. Fortunately, researchers don’t need to see the entire image before they can start scanning through the important pieces of text. By sending the important “text” portions of a document image first and filling in the remaining detail later, researchers can browse through document images at an interactive speed. This technique is known as Progressive Image Transfer (PIT).

We propose a novel method for representing document images in a parametric form. This method, named CECAT (Curve Encoded Compression and Transmission) compares favorably to current image compression and Progressive Image Transfer strategies. Because many documents (especially handwritten genealogical documents) contain a wide variety of shapes, we claim that using a parametric fit to connected component



Figure 1.1: Image of 1910 U.S. Census

contours can provide greater compression than current codebook approaches. By using first to third-order parametric curves, we believe that we can better compress the high resolution foreground mask used by current “hybrid” compression strategies (e.g. DjVu). In addition, we also believe that CECAT can also be used to enhance the Progressive Image Transfer strategy known as Just-In-Time Browsing (JITB), which will help meet the needs of researchers who are performing document browsing operations [3].

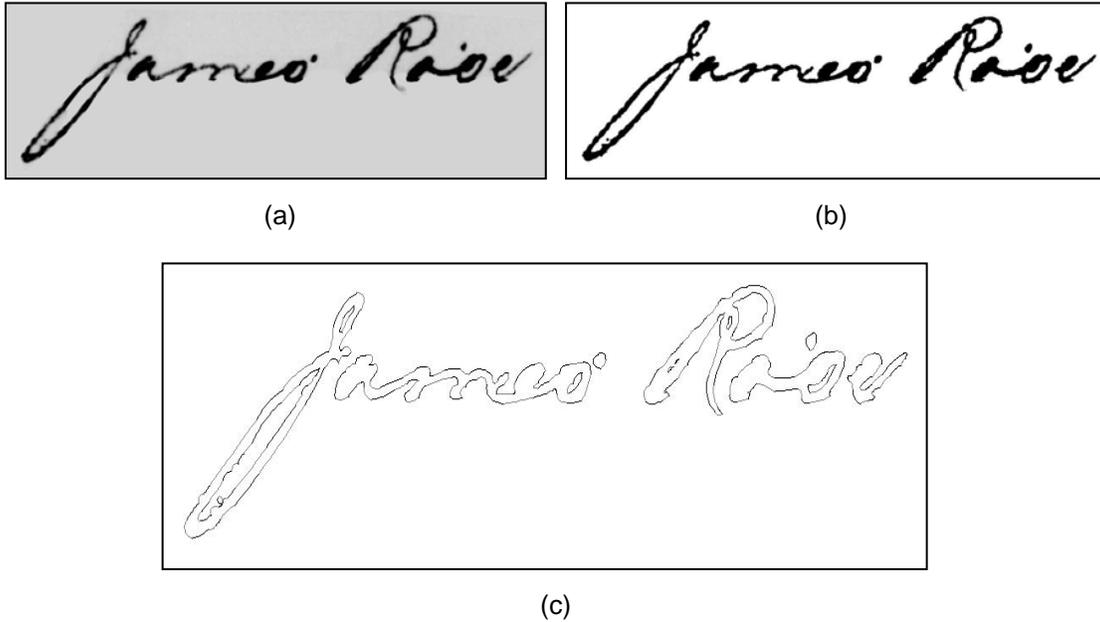


Figure 1.2 CECAT Compression Strategy. (a) Original Image (b) After Background Removal and Thresholding (c) After Contour Extraction and Parametric Fit

CECAT Phase One – Curve-Encoded Image Compression

First consider the image shown in Figure 1.2a shown above. By applying a median filter to remove the background data [4] followed by a modified version of the Niblack thresholding operation [5], we have created a binary foreground mask for this image (Figure 1.2b). Uncompressed, this foreground mask is 878 KB in size. By compressing this with high-quality JPEG compression, this image can be compressed down to 40.8 KB.

Once we have the foreground mask, CECAT marks the contours for each connected component using a simple contour-following algorithm. The contours are then fitted with first to third order Bezier curves (lines, quadratics, and cubics). See Equations 1.11 – 1.13 for a brief description of this parametric representation [6]. Figure 1.2c shows the image as represented by these parametric curves. These contours are then sorted by locality and compressed. Using the MDS encoding scheme [7] on these curves reduces the image size to less than 1 KB!

Bezier Curve Parametric Representations

$$1^{\text{st}} \text{ Order (Line): } \mathbf{p}(u) = (1-u)\mathbf{p}_0 + u\mathbf{p}_1 \quad (1.11)$$

$$2^{\text{nd}} \text{ Order (Quadratic): } \mathbf{p}(u) = (1-u)^2\mathbf{p}_0 + 2u(1-u)\mathbf{p}_1 + u^2\mathbf{p}_2 \quad (1.12)$$

$$3^{\text{rd}} \text{ Order (Cubic): } \mathbf{p}(u) = (1-u)^3\mathbf{p}_0 + 3u(1-u)^2\mathbf{p}_1 + 3u^2(1-u)\mathbf{p}_2 + u^3\mathbf{p}_3 \quad (1.13)$$

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

CECAT Phase Two – Progressive Image Transfer

Because parametric curves are sorted by locality (nearby contours being grouped together), the first stage of CECAT's Progressive Image Transfer strategy involves transmitting and filling these groups of contours. The most important contours (determined at the time the image is encoded) are sent first. Figure 1.3a shows an example of this contour transfer. Once the binary foreground mask is complete, grayscale data is restored into the foreground mask as shown in Figure 1.3b. Once this is complete, the background data is finally added to the image (as shown in Figure 1.3c).

By sending the most important readable sections of the image first, CECAT makes it possible for researchers to quickly scan through the compressed document images, even with low bandwidth connection speeds. If the document appears to be of importance to the researcher, he/she can wait for the entire image download to complete. If the document is not relevant, however, the researcher can quickly move on to the next document image with a minimal amount of time lost.

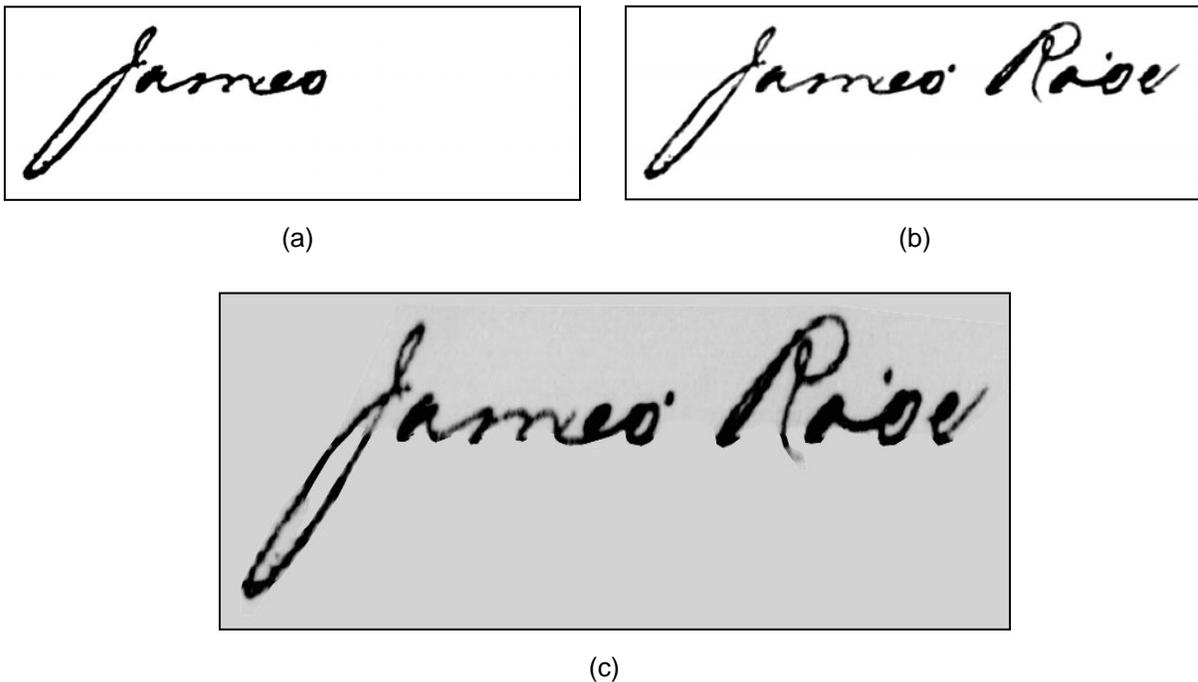


Figure 1.3 CECAT Progressive Transfer Strategy. (a) First Contours Transferred (b) All Contours and Grayscale Data Transferred (c) Background Data Transferred

References

- [1] NTIA and ESA of the U.S. Department of Commerce. A nation online: How Americans are expanding their use of the internet. Feb 2002. Found at URL <http://www.ntia.doc.gov/ntiahome/dn/>, visited on 2 March 2004.

- [2] Leon Bottou, Patrick Haffner, Paul G. Howard, Partice Simard, Yoshua Bengio, Yann Le Cun. High Quality Document Image Compression with DjVu. *Journal of Electronic Imaging*, Vol 7, No 3, pp 410-425, SPIE, 1998.
- [3] Douglas J. Kennard. Just-In-Time Browsing for Digital Images. Thesis Presented to BYU: February 2003.
- [4] Luke A. D. Hutchison, William A. Barrett. Fast Registration of Tabular Document Images Using Fourier-Mellin Transform. In *Proceedings of the IEEE First International Workshop of Document Image Analysis for Libraries (DIAL)*, Palo Alto, CA, pages 253-269, January 2004.
- [5] Wayne Niblack. *An Introduction to Digital Image Processing*. Prentice-Hall International, 1985.
- [6] Michael E. Mortenson. *Geometric Modeling*. John Wiley & Sons, New York, 1985.
- [7] Michael D. Smith. Handwriting Compression using Quadratic Curves. Brigham Young University Computer Science 750 Project Write-Up. November 29, 2003.