Digital Hammurabi Progress Report for the 49e Rencontre Assyriologique Internationale Nineveh, The British Museum, London, 7-11 July 2003

Dean Snyder
dean.snyder@jhu.edu

The Digital Hammurabi Project is developing new hardware and software aimed at enhancing research on cuneiform documents (figure 1). The first nine months of our $1.65 million, three-year grant from the U.S. National Science Foundation has seen progress on all research fronts. We are concurrently developing new 3D surface scanning hardware technology, new computer algorithms for visualizing 3D models of cuneiform tablets, a new international standard computer encoding for cuneiform text, and the beginnings of a digital archive and library for 3D cuneiform tablets.

The goals of the project are to provide all researchers with the tools necessary for the remote, virtual autopsy of cuneiform tablets, the conversion of 3D tablets to 2D for print, a standard encoding for font and natural language processing software, and file formats for long term archiving, preservation, and transmission of 3D cuneiform tablets.

SUMERO-AKKADIAN CUNEIFORM ENCODING

The Digital Hammurabi Project hosted a very successful second international Initiative for Computer Encoding (ICE2) conference at Johns Hopkins University, Baltimore, Maryland, USA, June 5-7, 2003. Cuneiform scholars, Unicode experts, software engineers, and font architects from four countries were in attendance. Consensus was reached on several methodological issues and a working group was announced. The working group, consisting of Miguel Civil, Jerrold Cooper, Karljuergen Feuerherm, Madeleine Fitzgerald, Eckart Frahm, Cale Johnson, Matthew Stolper, and Steve Tinney, and coordinated by Dean Snyder, has already begun work on compiling a cuneiform character inventory, based on grapheme lists provided by the Cuneiform Digital Library Initiative, the Pennsylvania Sumerian Dictionary Project (incorporating lists provided by Civil), and various published and unpublished sign lists (including draft materials from Rykle Borger's forthcoming Mesopotamisches Zeichenlexikon). The preliminary Unicode proposal will be posted on the Internet in mid-September for public review and comment. We intend to make a formal Sumero-Akkadian cuneiform computer encoding proposal to the international meeting of the The Unicode Technical Committee, which is hosted this year by the Digital Hammurabi Project at Johns Hopkins University, November 2003.

3D SURFACE SCANNER

After extensive surveys of the state of the art in high resolution 3D surface scanning hardware technologies, and numerous discussions with cuneiformists, tablet curators, and 3D scanner manufacturers, vendors, and software engineers we have refined our requirements for scanning cuneiform tablets. We have designed and fabricated precisely calibrated aluminum and ceramic targets in order to objectively establish resolution and accuracy claims. We have submitted these targets, as well as cuneiform tablets, for scanning to leading manufacturers and vendors, and have performed our own scanning with an array of technologies - in particular ultra-high resolution scans (6 micron, 167 points per mm) of cuneiform tablets.

Our current thinking on requirements for scanning cuneiform tablets include the following points: • we must offer greater spatial resolution than is currently available, capturing tablet geometry at 25 micron* resolution, or 40 points per mm (figures 3-5); • we need to capture color, BRDF, and polarization data, possibly at lower resolutions than the geometric data; • we need to scan an average cuneiform tablet in under 2 minutes; • scanners must be portable - for example, fitting in two pieces of luggage, and totaling less than 100 pounds; • we cannot clamp the tablets in any way - gravity will be our only holder; • scanning and hole filling must be automated, with only one manual re-positioning for each tablet; • run-time variable resolution scanning is a plus; • no scanning aids, such as fiducial markers, may be attached to tablets; • 3D displays may be utilised for greatly enhanced readability of cuneiform tablets.

In order to fulfill these requirements, Digital Hammurabi engineers, physicists, and instrumentation specialists at the Johns Hopkins Applied Physics Laboratory are actively and concurrently investigating several technologies - tri-color laser range-finding, conoscopic interferometric range measurement, photometric stereo, and amplitude modulated laser diodes operating at the telecom wavelength of 1550nm. We believe an optimal solution for 3D surface scanning of cuneiform will depend on multiple technologies working together.

One of the promising technologies for our application is the tri-color laser range-finding technology invented by Marc Rioux at the National Research Council of Canada. We are currently pursuing a formal research and development collaboration with the NRC and Arius3D of Canada, and other parties to be announced.

3D VISUALIZATION SOFTWARE

We are designing and coding ground-breaking software for the visualization of 3D cuneiform tablets, software that will allow researchers interactively to change tablet viewing position, lighting conditions, shadows, detail, and feature enhancements. Our software is cross-platform and freely available for Linux, Macintosh OSX, and Windows XP. (We are using Java for the user interface and C++ with OpenGL for the 3D processing and rendering, and are writing to the leading graphics video cards, nVidia and ATI.)

A 100mm by 150mm cuneiform tablet, when scanned at 25 micron intervals (40 points per mm), will yield about 1 gigabyte of data (40,000 square millimeters times 1600 points per square millimeter times 16 bytes = 1,024,000,000 bytes). Data objects of this size present challenging problems for internet transmission and remote visualization. To address these issues, we are, among other efforts, developing parameterization algorithms that will allow efficient and accurate data compression (and, as a by-product, enable tablet unfolding onto a flat plane). We are also developing algorithms that will calculate the next optimal position a scanner should take in order to fill holes left by previous scans.

We are developing, improving, and adapting various rendering algorithms that will provide new visualization tools for cuneiformists: • accessibility shading computes the maximum radius of a sphere that can fit in each depression in order to approximate how much light reaches there; it finds narrow areas and darkens them; • curvature-based shading darkens areas of greater curvature, or slope; • local depth-directed shading computes the local depth of each point from a statistically constructed surface in its neighborhood and uses that to modify the amount of light reaching that point for shading - this assumes that on average deeper recesses receive less light (figure 2); • shadow computation algorithms precisely determine if light reaches a given point on the tablet surface and darkens areas not illuminated by light, allowing researchers to see shadows raking across and in and out of wedges while rotating tablets on the computer screen.

DIGITAL CUNEIFORM ARCHIVE & LIBRARY

In anticipation of the need to store, archive, and access a petabyte (one million gigabytes) of cuneiform data, we have begun planning for a very large data center at Johns Hopkins University. Apple Computer has already donated us a dual processor xServe computer server and a 2.5 terabyte RAID hard disk storage array. And we are developing XML schemas aimed at data longevity and recoverability for cuneiform bibliographic markup, cuneiform text markup, and cuneiform 3D data formats.

---

*A micron, or more accurately a micrometer, is a millionth of a meter, a thousandth of a millimeter. Typical current 3D surface scanning technologies suitable for cuneiform tablets scan at 100 micron intervals, or 10 data points per millimeter (254 data points per inch). We are currently scanning at 40 micron intervals, or 25 data points per millimeter (635 data points per inch); our goal is 25 micron scan intervals or 40 data points per millimeter (1016 data points per inch).

Figure 1: 3D tablet - Cuneiform tablets are 3 dimensional; the wedges themselves are 3 dimensional; and the writing streams can be 3 dimensional. [100 micron 3D scan by the National Research Council of Canada; rendering by Digital Hammurabi.]

Figure 2: Depth-directed shading - an unrealistic 3D feature enhancement technique whereby deeper areas receive less light. [100 micron 3D scan data by the National Research Council of Canada; rendering by Digital Hammurabi.]

Figure 3: 3D tablet resolutions - The area framed here is enlarged to the right. [Digital photograph of cast of A32099, Oriental Institute, Chicago, a Neo-Babylonian tablet of manumission for Laqiptu and her three children.]

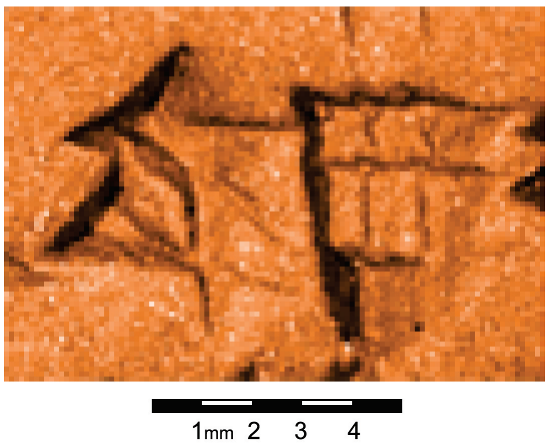

Figure 4: 100 micron resolution

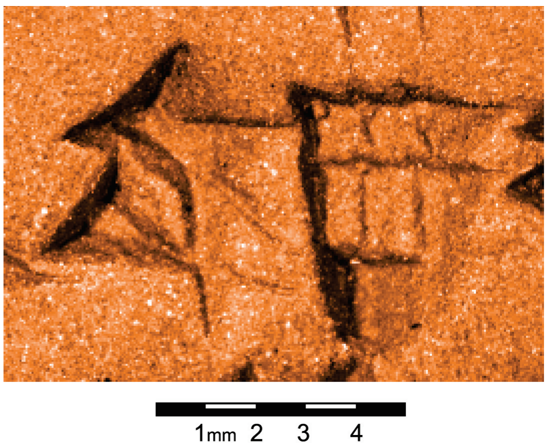

Figure 5: 40 micron resolution

Scan Resolution is one of the most important issues affecting digital tablet autopsy. Fig. 4 represents a 100 micron* scan - i.e., 3D data was collected at intervals of 1/10 of a mm; fig. 5 shows our current high resolution capability, 40 micron resolution, or intervals of 1/25 of a mm. We are targetting 25 micron resolution.

At 100 micron resolution, the vertical wedge of the UD sign, which is 4.4 mm tall, has 44 data points along its length, while its upper and lower diagonal wedges have 32 and 28 data points, respectively, along the widths of their wedge heads. The 40 micron scan has 110 data points along the length of the vertical wedge, and 80 and 70 data points, respectively, along the widths of the wedge heads. More data points mean greater readability, particularly for broken or difficult to read signs.

[Photometric stereo scans made by the Johns Hopkins Applied Physics Laboratory]