Editor: Mike Potel

Using Mid-Range Laser Scanners to Digitize Cultural-Heritage Sites

Adam P. Spring, Caradoc Peters, and Tom Minns
Combined Universities in Cornwall

Here, we explore new, more accessible ways of modeling 3D data sets that both professionals and amateurs can employ in areas such as architecture, forensics, geotechnics, cultural heritage, and even hobbyist modeling. To support our arguments, we present images from a recent case study in digital preservation of cultural heritage using a mid-range laser scanner.

Our appreciation of the increasing variety of methods for capturing 3D spatial data inspired our research. Available methods include photogrammetry, airborne lidar, sonar, total stations (a combined electronic and optical survey instrument), and mid- and close-range scanning.[1] They all can produce point clouds of varying density. In our case study, the point cloud produced by a mid-range scanner demonstrates how open source software can make modeling and disseminating data easier. Normally, researchers would model this data using expensive specialized software, and the data wouldn't extend beyond the laser-scanning community.

The Case Study

Our subject was the Wheal Unity Engine House in Cornwall, UK. This 19th-century building contained a steam-powered beam engine that helped pump mine shafts free of water.

We chose a mid-range terrestrial laser scanner for three reasons. First, it's designed for rapid acquisition of surface data, unlike total stations, which focus solely on capturing individual points. Second, it has the appropriate data capture range in terms of distance for terrestrial land and building recording. Finally, unlike photogrammetry, it isn't affected by surface texture. (However, it's still affected somewhat by surface reflectants.) So, it can deal with both shiny and regular surfaces.

To model the case study, we selected Blender, an open source program.

The Laser Scanner

We used the Leica HDS3000 (see Figure 1), a time-of-flight (ToF) mid-range laser scanner with a built-in mechanical accuracy of +/–5 mm, a 360-degree horizontal scan window, and a 270-degree vertical scan window. (For a short history of laser scanning, see the related sidebar.) We captured all data using a 3R-rated (a safety classification based on the wavelength and power of the energy produced) green laser fixed to a wavelength of 532 nanometers.

Besides the mechanical accuracy and field of view, other immediate variables related to the scanner's interior orientation affect data capture. These include the range at which we can collect data (realistically, 150 meters maximum) and the signal-return strength, which is affected by both the range and the surface reflectants' quality. The signal-return strength, combined with the mechanical accuracy, affects data noise in the point cloud. This directly affects the accurate representation of key features, such as surface texture, during modeling. To optimize point-cloud registration (a process in which individual points are joined together to recreate the Engine House), users should employ targets purchased with the scanner that are coated in a retroreflective material.

The Target Environment

Measurement accuracy and the subsequent modeling requirements were linked directly to our application's data-processing phase. In this case, the application was a model that required detailed recording of surface textures obtained from high-resolution scans.

The primary goal of data capture and modeling the engine house was to place an example from the Cornwall and West Devon Mining Landscapes World Heritage Site (WHS) in a digital context. The engine house model could then be used in open-access immersive environments such as Second Life for edutainment and WHS site promotion. Using the physically liberating aspect of digital media, disabled users or visitors could access the WHS freely. Furthermore, information could appear on message or video boards in Second Life to assist interpretation. If greater realism was

0272-1716/10/$26.00 © 2010 IEEE

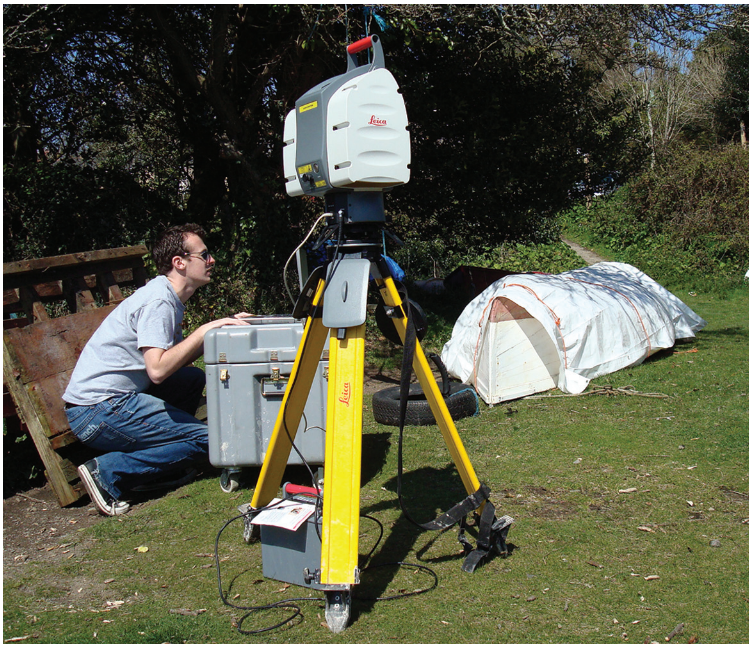

Figure 1. The Leica HDS3000 scanner used in our case study. This scanner is a time-of-flight, mid-range laser scanner with a built-in mechanical accuracy of +/–5 mm, a 360-degree horizontal scan window, and a 270-degree vertical scan window.

required, then textures—for example, weathered granite—could be draped over the detailed mesh created from the point cloud.

We captured four exterior scans of the engine house at 30-mm resolution, with an interior scan at 15-mm resolution. This generated a data cloud of approximately one million points. The data's rate and size would present problems for any user without access to expensive software and specialized knowledge. The capture process produces an overkill of points. This ironically generates data too rich and dense, which makes working with the data problematic in the current processing-power climate. The user can control the density somewhat by setting the point resolution before scanning, but any reduction in point density must be balanced against the resolution of the registered point cloud in its totality. The registration process—derived from Paul Besl and Neil McKay's *iterative closest point* algorithm—also affects point cloud accuracy.[2]

You can achieve decimation of points in an open source environment during data processing. However, a workflow centered on cultural-heritage preservation would benefit considerably from a more balanced number of points during data capture. This would lessen further alteration of all data, such as decimation or unification, and the associated effects on the data.

Modeling in Blender

We chose Blender for several reasons. In addition to its support community and ease of access, we like its animation functionality. Also, it's a logical

Laser Scanning

Commercial mid-range laser scanning has been available since 1998. Since 2007, its range of uses and applications has become more widely known. However, despite its rapid development as a technology and growth as a commercial market, it still remains at a quasi beta phase. Scanners' interior orientation (laser system and type of beam used to acquire data) vary between manufacturers, who offer a time-of-flight, phase shift, or triangulation laser system for data capture, as well as their own modeling software and file formats.[1]

The history of such commercial laser scanners makes them an appropriate subject for our case study. In 1999, Ben Kacyra, Jerry Dimsdale, and Mark Brunkhart's US Patent 5988862, which they had submitted in 1996, came into force. This patent for laser scanning was based on well-tested technologies that radar specialists had used for decades. The main reason for their invention and submission of the patent was a gap in the commercial survey market for a terrestrial instrument with hardware requirements intended for rapid acquisition of coordinated data at millimeter accuracy (at the time the available technology was either too slow or set for centimeter accuracy). A further use for laser scanning was to avoid danger when recording industrial accidents or contaminated areas.

The significance of this for cultural heritage was that this new technology could record surfaces with greater rapidity than conventional technologies such as total stations could. Where buildings or other historic structures were in danger of collapse, laser scanner operators no longer had to get close to them to produce detailed records.

Key to the development of laser scanning were communities freely exchanging and disseminating information. Three of the five main manufacturers of mid-range laser scanners have common research origins. These companies are Leica, which bought its rights from Cyrax (Kacyra and Dimsdale's company); Riegl, whose founder was a former colleague of Kacyra's from before 1992 (hence, the Cyclone software works with scan data from both Leica and Riegl scanners); and Topcon, who cooperated with and then bought out Voxis, a company founded by Dimsdale in 2003.

Reference

1. T. Kremen, B. Koska, and J. Pospísil, "Verifying Terrestrial Laser Scanner Quality," *GIM Int'l*, vol. 21, no. 8, 2007, pp. 28–29.

open source environment for working with point clouds. We wanted to demonstrate that data captured by costly techniques doesn't have to be expensive to process. We also wanted to encourage greater synthesis and application of scan data by providing tools to process it.

Normally, Blender interprets each point in the cloud as a separate object because it works with absolute or fixed coordinates. We had to reverse this viewpoint so that Blender redefined each section of the scanned area and recognized it as one object with multiple points. The Blender extraction method retains the original relative positions of all point cloud data, while grouping it into manageable blocks.

Blender has a very useful feature set, but it isn't designed to work with point clouds' large data sets that laser scanning creates. Instead, the program works with distinctive geometric shapes such as polyhedra. We discovered a way around this problem by splitting the point cloud into more manageable sections. These were imported into Blender one by one, and they were converted into meshes of geometric shapes. Underneath these shaped meshes, the original points remain, should reprocessing or reinterpretation be required.

For more on Blender, see the related sidebar.

Processing the Scan Data

Importing the scan data into Blender takes five steps (see Figure 2):

1. Register all scan worlds using the scanner manufacturer's software—for example, Cyclone (Leica's scanner operating software) or RiSCAN (Riegl's scanner operating software). A scan world is an area captured by the scanner that is later combined with other areas captured, if required, to make up the final point cloud. To put scan worlds together, there must be overlap of an area in each one.
2. Segment the fully registered point cloud by using the reference plane function in the manufacturer's software tools menu. You can achieve the same process in other software such as AutoCAD using a DXF (Drawing Interchange Format) copy of the point cloud. Drape the reference plane over the point cloud. You can extract individual sections in the plan-view perspective as DXF files—each retaining all *x*, *y*, and *z* coordinates as in the original integrated modeling program (IMP) file.
3. Extract each section as a DXF file. Import the sections into Blender.
4. Open individual files in Blender and select as many points as possible. Go to the bottom menu and use the join objects function.
5. Repeat the process on all files, and then link the objects together to reassemble the point cloud.

Blender

Blender began in 1995 as a 3D toolkit for a Dutch animation studio called NeoGeo. The studio set up a subsidiary, Not a Number (NaN), to develop and market the program as a free cross-platform 3D creation suite. At that time, companies normally charged thousands of US dollars for such software. Profits from additional products and services made it possible to offer the program for free. Because the program was free and offered fast animation, demand for it spread to the wider computing public rather than just to wealthy corporations. The company failed owing to market realities and its limited capabilities at the time.

However, in 2002, Ton Roosendaal, an original NeoGeo cofounder, founded the Blender Foundation because of the user community's enthusiasm and commitment. The foundation has an official website (www.blender.org), a community website (www.blender.org/community/user-community), a wiki (en.wikipedia.org/wiki/Blender_(software)), and a news website (www.blendernation.com).

We used the points from the scans, photos, or other point data to align particular functionalities in Blender, such as mesh tools, polygons, and plane cubes. Basic Blender functions such as extrude, remove doubles, and face were sufficient for this modeling. However, lacing together vast numbers of points requires a suitable macro or script. Where relevant, we also used illustrations and photographs to make decisions, which also helped us recreate texture and internal parts such as machinery that had rusted away or been removed.

The process demonstrated that working with point data in an open source environment is possible without a plug-in or an added design feature for programs such as Blender. Familiar difficulties of modeling with large data sets, such as time and effort, still apply. Here, we've demonstrated that processing point data doesn't have to be expensive but does require an awareness of the available tools and lateral thinking.

The process isn't just technical. Once scan data, photo data, or any other digital data has been captured, it becomes static, unaltered by time, and representative of the moment in which an object or scene is captured. It's also representative of the processes and performance that went into its capture. As outlined in the concept of empirical provenance (which we discuss in the next section), the link between captured and processed data is ever present but might not always be apparent when data capture or interaction occurs. Working with

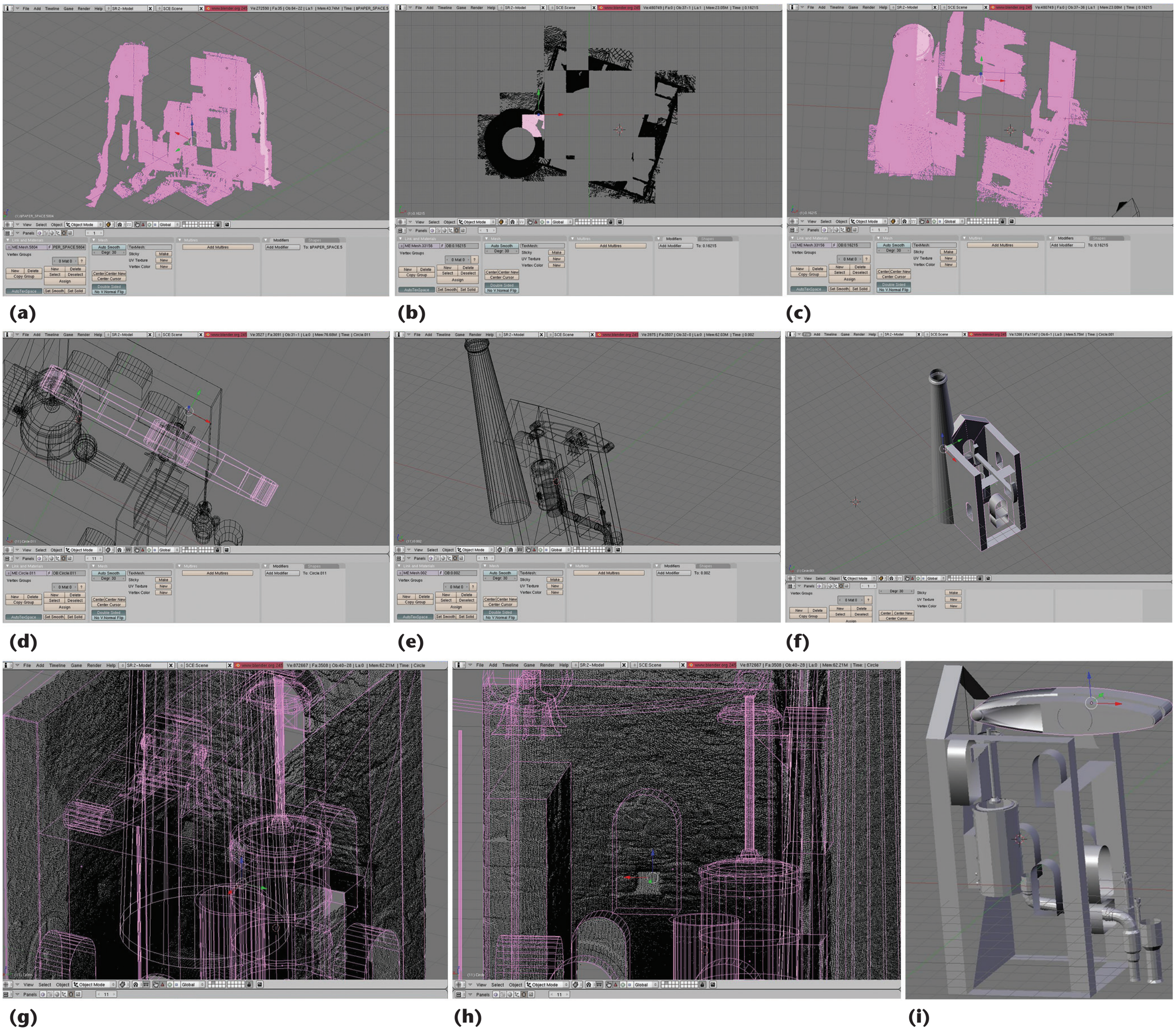

Figure 2. Stages of processing of point cloud data in blender. (a) Point cloud conversion in Blender. (b) Reconstructing a modified point cloud section using absolute coordinates. (c) A polyhedral frame for modeling. (d) Reconstructing the Wheal Engine House's inner workings. (e) A skeletal frame with the engine incorporated. (f) Reconstruction before adding the surface texture. (g) Using texture from point density to aid modeling. (h) Inner workings and external frame construction. (i) A layered process for model construction.

data sets involves further processing or interpretation that reflects the environment and conditions in which it's conducted. The tools and the person using them will always directly affect the results. In our case study, practicalities relating to the available processing power and the problems of working with large data sets were overriding factors determining our approach.

Long-Term Data Management

Eleven years after Stewart Brand's "Escaping the Digital Dark Age,"[3] retrieval of information as recent as 10 years old is still, if not more, problematic. For example, you could view digital information's exponential growth as alarming or exciting. Recent research on predictability models for Internet growth and case studies, such as Guo-Qing Zhang's "Evolution of the Internet and Its Cores,"[4] provide insight on the state of digital information, which doubles in size every 5.32 years. Digital survey tools, such as mid-range laser scanners, aren't detached from this situation.

It's no coincidence that new methods of managing increased data production have emerged over the past 11 years. Each method gives a range of sustainable preservation strategies for all relevant data sets—providing guidance on what information to keep. Management strategies such as the CIDOC

Empirical provenance for mid-range laser scanning	**Empirical provenance for modeling**
■ Scanner model (including technical specifications)	■ Software used
■ Point or target registration	■ Steps undertaken to make changes
■ Software used	■ Any changes in file format
■ Scan position	■ Any changes in resolution
■ Scan direction	
■ Field notes (for example, environmental conditions)	
■ Photographs and videos of the recording process	
■ The resolution selected	
■ Original individual scans before decimation or any modification before combination with other scans and before overdraping	
■ Original photographic panorama before being overdraped on the scan	
■ Any conversions—for example, Cyclone and subsequent file format changes	

Figure 3. Checklists for establishing the empirical provenance for mid-range laser scanning and subsequent modeling.

Conceptual Reference Manual (CIDOC CRM), empirical provenance, Dublin Core Metadata Initiative (DCMI), and RecorDIM (Recording, Documentation, and Information Management), whether intentional or otherwise, highlight the flexibility, dynamic flow, and organic characteristics of digital-data development and growth. Worryingly, they also reflect digital data's ephemeral nature and the difficulties in ensuring its long-term survival.

Empirical provenance (sometimes called empirical origin) is a solution for giving data sets a cultural memory; it originates in Immanuel Kant's *Critique of Pure Reason*. It concerns the need to establish where information comes from and how it was produced. This helps avoid misconceptions due to corruption of data sets caused by human error and misunderstanding. The role people play in creating digital data sets, along with processing and archiving, becomes easier to follow and understand.

Cultural Heritage Imaging (CHI)—a California nonprofit corporation—has outlined key information for empirical-provenance records for camera-based techniques (see http://www.c-h-i.org/featured_projects/feat_pub_04.html), based on Mark Mudge and his colleagues' scheme.[5] On the basis of CHI's outline, we've developed checklists for mid-range laser scanning and modeling (see Figure 3).

Open source and low-cost modeling are good ways to preserve data through accessible distribution in the same way that plain text and PDF files have assisted the flow of basic text and 2D images. Accessible distribution ensures many copies are available and that data will less likely be lost. However, wide distribution can lead to loss of data integrity, so it's important that the empirical provenance accompanies the data outlining its creation. The process's relative simplicity makes it accessible to a wide audience, thus bridging the gap between specialists and nonspecialists.

***Adam P. Spring** is codirector of Archaeovault. Contact him at adamspring@gmail.*

***Caradoc Peters** is program leader of archaeology at the University of Plymouth, Truro College Campus. Contact him at rutcpeters@plymouth.ac.uk.*

***Tom Minns** is a research fellow at Combined Universities in Cornwall. Contact him at tminns@btinternet.com.*

Contact department editor Mike Potel at potel@wildcrest.com.

References

1. A.P. Spring, C. Peters, and A. Wetherelt, "3D Laser Scanning and Its 2D Partners: Looking Forward to a Harmonious Future Together," *GeoInformatics*, vol. 11, no. 8, 2008, pp. 50–54; http://fluidbook.microdesign.nl/geoinformatics/08-2008.
2. P.J. Besl and N.D. McKay, "A Method for Registration of 3-D Shapes," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 14, no. 2, 1992, pp. 239–256.
3. S. Brand, "Escaping the Digital Dark Age," *Library J.*, vol. 124, no. 2, 1999, pp. 46–49.
4. G.-Q. Zhang, "Evolution of the Internet and Its Cores," *New J. Physics*, vol. 10, no. 12, 2008, pp. 1–11.
5. M. Mudge, M. Ashley, and C. Schroer, *A Digital Future for Cultural Heritage*, July 2009; www.c-h-i.org/events/CIPA_2007.pdf.