A System for High-Volume Acquisition and Matching of Fresco Fragments: Reassembling Theran Wall Paintings

Benedict J. Brown*   Corey Toler-Franklin*   Diego Nehab†   Michael Burns*   David Dobkin*
Andreas Vlachopoulos‡   Christos Doumas§   Szymon Rusinkiewicz*   Tim Weyrich*

Figure 1: *Our acquisition system, deployed at the Akrotiri Excavation, Thera. We use a flatbed scanner to capture high-resolution images and normals of wall painting fragments (shown at left), and multiple 3-D scanners to acquire geometry. A single user can operate up to four scanners simultaneously, while a second user operates the flatbed scanner and verifies processing results. This yields a throughput of approximately 10 fragments per hour. Our matching algorithm correctly finds the only two matches in this data set using the scanned 3-D geometry.*

Abstract

Although mature technologies exist for acquiring images, geometry, and normals of small objects, they remain cumbersome and time-consuming for non-experts to employ on a large scale. In an archaeological setting, a practical acquisition system for routine use on *every* artifact and fragment would open new possibilities for archiving, analysis, and dissemination. We present an inexpensive system for acquiring all three types of information, and associated metadata, for small objects such as fragments of wall paintings. The acquisition system requires minimal supervision, so that a single, non-expert user can scan at least 10 fragments per hour. To achieve this performance, we introduce new algorithms to robustly and automatically align range scans, register 2-D scans to 3-D geometry, and compute normals from 2-D scans. As an illustrative application, we present a novel 3-D matching algorithm that efficiently searches for matching fragments using the scanned geometry.

1 Introduction

Computer-based acquisition and processing of 3-D shape and reflectance data has proven its potential to revolutionize certain kinds of research in the humanities and social sciences. However, previous "computational humanities" projects have necessarily involved significant manual labor and nontrivial participation by computer scientists. This is because the tools for high-throughput 3-D scanning, calibrated reflectance measurement, systematic organization

---

*Princeton University.
  {bjbrown,ctoler,mburns,dpd,smr,tweyrich}@cs.princeton.edu
†Microsoft Research.
  nehab@microsoft.com, work conducted while at Princeton University
‡Akrotiri Excavation, Thera
§National University of Athens and Akrotiri Excavation, Thera

of acquired data, and applications such as 3-D shape search have been in the realm of research, rather than production. As these technologies mature, their availability to non-computer experts can be broadened by solving the crucial problems of *scalability* and *usability*. This paper focuses on the associated technical challenges.

We describe how to integrate acquisition and processing tools recently developed in the computer graphics field into a system suitable for large-scale archaeological documentation and reconstruction. Although we present our work in the context of a case-study, described below, our methods for rapid acquisition by non-experts, automatic and robust 3-D reconstruction, user-focused workflow, and scalable 3-D search are applicable to many archaeological excavations and cultural heritage projects, and could be deployed broadly with minimal incremental development effort.

We focus on the specific problem of documenting and reconstructing fragments of wall paintings from the site of Akrotiri on the volcanic island of Thera (modern-day Santorini, Greece). An eruption destroyed the ancient civilization, burying the remains of a flourishing Late Bronze Age (c. 1630 B.C.) settlement in ash, similar to Pompeii. Excavations have yielded an unparalleled trove of artifacts and information from the prehistoric Aegean, including numerous wall paintings ranging from naturalistic and narrative scenes to abstract motifs [Doumas 1992].

Uniquely, much of the original plaster material remains at Akrotiri; from the tens of thousands of small lime-plaster fragments uncovered at the site (typically less than 10 cm across and less than 1 cm thick), it is possible to largely reconstruct interior walls of buildings and thereby recover important clues to the culture, technology, and architecture of ancient Thera. The fragments tested in this paper come from a huge wall painting (approximately 5.20 m × 3.20 m) depicting symmetrical pairs of spirals (Figure 15); the original wall was not preserved at all. Thus, the restoration of this composition is also important to studying the architecture and function of the collapsed third floor of the building [Vlachopoulos 2008, 454, Figure 41.47–50].

At Akrotiri, as at other excavations, recent computer graphics research may significantly improve the quality of artifact documentation and reduce the human labor involved in matching the "jigsaw puzzle" of fragments (currently estimated at 75% of the total human effort at the site), freeing up time for other important tasks including conservation and restoration. For example, 3-D scanning may provide more complete documentation of the state of excavated fragments; registration algorithms may be used to align the 3-D

scans with 2-D color scans or photographs; both the 3-D and 2-D data may be stored in a database together with relevant metadata such as the location where the fragment was excavated and the tray in which it is stored; and automatic image and 3-D search methods may suggest fragment matches. All told, the time to reconstruct a complete wall may be reduced from years to months.

The key challenge to building an effective acquisition system is minimizing manual labor without sacrificing data quality. Analysis tools such as matching algorithms face the same challenge. From this perspective, scalability and usability issues remain with each technology and algorithm suggested above. Therefore, we make the following contributions:

- An inexpensive and rapid 3-D model acquisition system that produces complete models of at least 10 fragments per hour with a single non-expert operator. The most important component is a robust registration algorithm for automatic, high-quality range scan alignment in the presence of unstable or degenerate geometry (e.g., the flat surfaces of wall painting fragments).

- Algorithms for obtaining 2-D color scans and normals (used to document fine relief on the front surfaces of some wall paintings), and registering these to the 3-D geometry. All scanned data and metadata are stored in a database, permitting fragment queries according to a variety of criteria.

- An efficient algorithm for computing candidate 3-D matches between pairs of fragments. In contrast with descriptor-based methods, our algorithm incrementally computes the exact matching error at all possible orientations. Using a novel mesh parameterization, we obtain the performance of a 2-D contour matcher while computing matching error on full 3-D geometry.

Our approach builds on many previous collaborations between computer graphics, art history, and archaeology. The Digital Michelangelo [Levoy et al. 2000], *Forma Urbis Romae* [Koller et al. 2006], and Pietà [Bernardini et al. 2002] projects, among many others, have demonstrated the usefulness of high-resolution 3-D scanning: the David model was used to plan a cleaning [Bracci et al. 2004], several new matches between *Forma Urbis* fragments have been found using the 3-D models [Huang et al. 2006; Koller et al. 2006], and the Pietà model has been used to analyze the statue's damage and subsequent repair [Wasserman et al. 2002]. However, the scanners used in these projects were expensive and unwieldy, and operation and data processing required teams of trained experts.

Other projects have focused on archaeological excavations, especially on reassembling pottery (e.g. [Willis 2004]). The principal acquisition goal in these projects has been to recover profile curves — cross-sections of pottery through the axis of rotation — which are important for both matching and archaeological study. This requires less complete data than we acquire, but the challenges of scalability and usability remain. Karasik and Smilansky [2008] discuss these issues, and present a system with which a user can capture 3-D scans of 10–15 pottery fragments (sherds) per hour. (As with our system, multiple operators working together can achieve higher throughput using the same hardware.) However, this system mounts sherds in a frame that obscures parts of their edges. With a similar amount of manual labor, we capture *full* 3-D models, as well as high-resolution images and normals.

2   Design Decisions

Our goal is to design a system and interface for efficient acquisition of geometric and reflectance data for thousands of modest-sized artifacts (5–10 cm across). Because computers and acquisition hardware are inexpensive relative to people, our design decisions favor capturing the most complete representation of an object with the least amount of human supervision. The required time could be further reduced by capturing only the immediately necessary data types, deferring other acquisition until later. But retrieving items

Figure 2: *Projected 3-D color of the front surface (left) and the corresponding, aligned 2-D scan (right). The 2-D scan includes some edge information as well, which is distorted by the flatbed scanner's push-broom projection, directional lighting, and narrow depth-of-field. However the front surface is evenly lit, undistorted, and sharper than the projected 3-D color.*

from storage is itself time- and space-consuming, and becomes more difficult once fragments have been physically reassembled. We therefore prefer to capture as much data as possible at once.

2-D Images   High-resolution images of the fragments' front surfaces are the most basic type of acquired data. The benefits for archiving, analysis, digital restoration, and matching based on decoration are evident. A contour derived from the front surface image is also useful for matching.

Color images may be obtained with a digital camera, a flatbed scanner, or a 3-D scanner (in conjunction with geometry). We use a flatbed scanner, because it does not require careful control of ambient lighting or fragment positioning, achieves higher resolution than either a camera or 3-D scanner, and guarantees a fixed spatial resolution since the sensor's distance to the object surface is constant. Figure 2 compares the projected color obtained from our 3-D scans to a 2-D scan; we believe that the improved color fidelity and resolution justify the additional effort of performing 2-D scans.

Flatbed scanners do have drawbacks. Their depth-of-field is low (only a few millimeters for the hardware we have tested), and we cannot always extract contours for 2-D matching since it is difficult to distinguish pixels belonging to the front side from those coming from the fractured edges. Wall painting fragments, however, have flat front faces (with few exceptions), we use the scanned color for texture-mapping 3-D models rather than contour matching, and the availability of the 3-D geometry allows us to readily separate the front from the edges, as described in Section 4.2.

3-D Geometry   While iconographic study of the wall paintings relies primarily on images, geometry is the most important cue for matching. Many matches are found, and all are verified, by testing whether fragments "lock" together. Thickness is used to quickly discard incompatible pairs. The back may contain impressions of material from the underlying mud-brick walls (such as reeds, pebbles, and sand), which are useful cues in the absence of decoration. Finally, contours may be extracted from the intersection of the 3-D model and a plane, which is offset from the front and back surfaces to avoid eroded regions.

Most fragments are less than 10 cm in diameter, and vary from a few millimeters to a few centimeters in thickness, well within the viewing volume of commercial 3-D laser range scanners. Geometry acquisition with sub-millimeter accuracy is therefore technically feasible: the challenge is minimizing the necessary user interaction. While commercial scanning software such as TowerGraphics Easy3DScan and NextEngine ScanStudio can automatically align a *single* series of scans taken on a turntable, it is impossible to capture all sides of an object in a single scan sequence. Furthermore, wall painting fragments yield flat, degenerate geometry, which standard alignment algorithms cannot handle [Gelfand et al. 2003].

By positioning the scanner at a 45-degree angle to a turntable, it is possible to acquire all necessary range scans with *two* scan sequences (Figure 3). As shown in Section 4.1, this setup also

Figure 3: Using an off-the-shelf 3-D scanner (front), positioned at a 45-degree angle relative to a motorized turntable, we can acquire the front, back, and sides of a fragment with a single manual flip. A 2-D scanner (rear) simultaneously acquires high-resolution color and normals.

provides enough constraints to align all range scans robustly and automatically, despite degenerate geometry.

Normals In Akrotiri wall paintings, artists frequently used string impressions as an aid to demarcate long, straight color edges. More intricate designs such as spirals were sometimes prepared with narrow incisions [Vlachopoulos 2008]. The impressions are helpful in matching, analysis and restoration, because they survive even where the original pigment does not. Such narrow, shallow markings are difficult to capture with either an image or range scanner. However, their sharp changes in normal direction show up clearly, as seen in Figure 8.

Several systems have been proposed for capturing normals, using both calibrated [Woodham 1980; Bernardini et al. 2002; Nehab et al. 2005] and uncalibrated [Toler-Franklin et al. 2007] point light sources. Instead, we rely on the flatbed scanner's *linear* light source, which is placed at a fixed angle with respect to the moving CCD. Hence, we may compute normals from multiple scans with different object rotations. (We typically use four scans, rotating the object 90° between each.) Combining normal and image acquisition exploits the flatbed scanner's high resolution while keeping the number of discrete scanning steps low.

Our method of obtaining normals is similar to the *linear light source reflectometry* of Gardner et al. [2003]. However, because we obtain a single image per scan, instead of a full time-sequence of images, our setup requires different calibration and normal-computation techniques. These are described in Section 4.3.

3 Acquisition Workflow

From the user's perspective, there are three discrete acquisition steps: 2-D scanning (images and normals), 3-D scanning, and verification and correction of results. In addition, the system must be periodically calibrated. The manual labor required at each step is largely dependent on interface design. While verification and correction effort is affected by the robustness of our registration algorithms — fewer errors mean fewer corrections — presenting an interface in which corrections are simple and efficient is just as important. We have therefore organized the acquisition and verification workflow into a series of simple operations that can be performed by a single operator in sequence or by several in parallel. The user interaction with each stage of our system, captured in real time, is shown in the accompanying video.

Image and Normal Acquisition The user enters a fragment's ID into the 2-D scanning software, places the fragment face-down on the flatbed scanner, and clicks the "scan" button. The software automatically finds the fragment's location in a pre-scan using background subtraction, scans it, and stores it in the fragment database.

The user performs four scans, rotated by approximately 90 degrees, then turns the fragment over and repeats the process.

Acquiring four scans of the front and back surfaces, which is required to capture normals, a single user can scan approximately 20 fragments per hour. If normals are not required, only a single scan of each face is performed, for a 70 fragment-per-hour throughput.

Geometry Acquisition The user places the fragment face-down on the turntable, and enters the fragment ID and laser start and end positions. The software acquires a set number of range scans (six by default, but adjustable for large fragments), rotating the turntable between each scan. The user turns the fragment face-up, acquires a second sequence of scans, and moves on to the next fragment. The first fragment is automatically aligned in the background. While scanning is in progress, the user can enter any meta-information and notes (ancient and modern damage, erosion, stains from organic materials, etc., and the excavation finding unit that codes provenance, mode of destruction, and other group-based metadata), although notes can also be entered separately.

The 3-D scanners we use are capable of acquiring approximately five fragments per hour. However, because the process is almost completely automatic, a single user can operate four scanners simultaneously, thereby acquiring 20 fragments per hour.

Maximizing Efficiency We have observed that a number of "tricks" improve the efficiency of acquisition. Although each of these is seemingly trivial in itself, we nevertheless document them here as an aid to practitioners:

- When the flatbed scanner smudges, the operator simply places fragments on a different area of the glass, which reduces cleaning frequency. Our background subtraction algorithm for locating fragments is reliable enough to continue locating the fragment and ignore smudges;

- We cover the 3-D scanner's turntable with a black cloth, to exclude its geometry from the scans. Cloth is more diffuse than black paint and dust and debris can be easily shaken off. A small white dot placed at the turntable axis helps center objects;

- The operator uses small wedges, also covered in black cloth, to stabilize fragments that do not sit flat on the turntable;

- A target of concentric circles aids in selecting the laser start and end positions (Figure 4). The chart is calibrated before scanning by placing it on the turntable and noting the scanner position when the laser crosses each circle. During scanning, the operator places a fragment on this diagram and reads off the optimal limits for the scan. Selecting tight scanning limits, instead of performing full-range scans, reduces total scanning time (often by a factor of 3) for a typical assortment of fragment sizes;

- The time spent adjusting scan limits is further reduced if the fragments are scanned in approximate order of size;

- Both 2-D and 3-D scanning software play a sound whenever a scan sequence is complete and intervention is required. This allows for easy multi-tasking by operators.
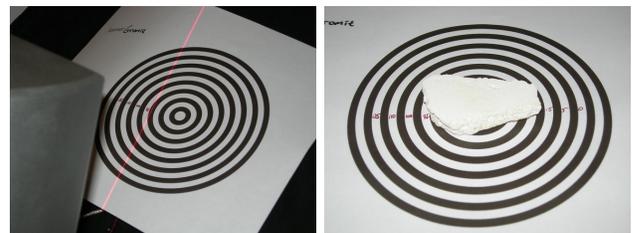

Figure 4: To reduce scanning time, a target records scanner settings for several laser positions (left). The operator uses this target to set the laser range for each fragment, so that only the relevant portion of the viewing volume is scanned (right).

Verification and Correction Once alignment and processing are complete, the user must verify their correctness. First he verifies that the geometry has been correctly acquired and aligned, then that the 2-D image is correctly registered to the geometry. When a processing stage fails, the simplest possible interface is presented to the user to correct *only* that error, after which automatic processing resumes. This minimizes the manual effort spent on each fragment.

There are three possible failure modes for geometry acquisition. If the turntable or scanner is bumped during acquisition, resulting in bad data, the fragment must be rescanned. If the fragment edge is large and flat, it may be incorrectly identified as the front surface; when that happens, the operator clicks on the front surface to automatically reorient the front scans. Finally, if the front scans are misaligned *horizontally* with respect to the back scans, the user repositions them with our constrained alignment interface. Because manipulation occurs within a plane, it is much simpler to control than a full 3-D manipulation interface. With this approach, correct alignments are verified in a second or two, while the occasional error is corrected in a few more seconds. In an experiment involving 150 fragments, verifying and correcting registration required 10 minutes, including correcting all 15 alignment errors.

Problems registering images to geometry are solved similarly to 3-D misalignment: the user provides an approximate initial alignment of the 2-D scan to a *projection* of the 3-D model's front surface, then the software re-optimizes the alignment. Although this optimization is slower than range-scan alignment, the error rate is also lower — only one fragment out of 150 failed to align automatically — so the overall verification time remains small.

Calibration Before acquisition starts, both the flatbed and 3-D scanners must be calibrated. The intrinsic parameters of the 3-D scanner are computed once only, and based on our experience do not change over time. We have also found that professional-quality, flatbed CCD scanners have very nearly square pixels. Any deviation can be compensated for when registering the image and geometry. However, we must calibrate for the 3-D scanner's turntable position (to enable automatic alignment), for the laser positions relative to the turntable (to speed up scanning), and for the flatbed scanner's lighting parameters (for normal acquisition). In practice, we calibrate the flatbed scanner only once, but recalibrate the turntable every day, even when it has not visibly moved.

The turntable position is determined by the plane of its surface, and its axis of rotation. A menu option in the 3-D scanner software guides the user through the calibration process. First, one scan is acquired of the empty turntable to determine its plane. Second, two scans are made of a single fragment, at a 30° rotation from each other. The two scans are automatically aligned (subject to user verification), and the axis of rotation is extracted from this alignment. Note that the axis does not have to be perfect: the multi-way ICP algorithm described below recovers the optimal axis of rotation for each scan sequence using this initial estimate.

Normal acquisition requires knowledge of how the scanner illuminates a diffuse target as a function of angle. We acquire this information by scanning diffuse white cards tilted at angles ranging from 0° to 25° towards the four sides of the scanner.

4 Automatic Processing Pipeline

The key to a practical acquisition system is automating 3-D scan alignment and registration of 2-D scans to 3-D models. Failures must be few and easy to correct. Because the classical iterative closest point (ICP) [Besl and McKay 1992; Chen and Medioni 1992; Rusinkiewicz and Levoy 2001] algorithm is unstable in degenerate cases such as a fragment's flat front surface [Gelfand et al. 2003], we extend it to operate simultaneously on all scans from a turntable series (i.e. all front scans or all back scans). Our extension, called *multi-way ICP*, is restricted to scans on a turntable, but it is faster

and more robust than full global registration. We also introduce a novel technique for estimating an initial alignment of the front and back scans to each other without any user input, resulting in a fully automatic alignment of all scans. To register images to the geometry, we optimize for a transformation rather than detecting feature points (e.g. [Lowe 2004]) because undecorated fragments do not contain enough meaningful features to compute a robust alignment.

4.1 3-D Alignment

Given a turntable's axis of rotation (obtained during calibration), we have good initial estimates of the relative alignments of all front scans and all back scans. We could refine these alignments with ICP, but the nearly-flat front scans often do not contain enough geometric constraints for the optimization to converge. Aligning fronts to backs is even more problematic: we have no initial alignment, since the flip was performed manually. We solve the geometric stability problems within the front and back scans using *multi-way ICP*, and handle the flip by explicitly detecting the front face and using a constrained ICP alignment to robustly align it to the back.

Multi-Way ICP For a typical sequence of six fragment scans we have a good initial alignment estimate because they were acquired on a calibrated turntable; any global registration algorithm should align them well. Indeed, even sequentially aligning each scan to the previous one [Chen and Medioni 1992], performs well when all alignments are stable. However, a typical front scan consists of a nearly-flat face, with minimal additional detail from the edge; in this *unstable* situation, each scan can slide parallel to the front surface without significantly increasing the alignment error.

In contrast to the general global registration case, we have important knowledge about the alignment of scans on a turntable. Specifically, for n scans, each scan is rotated $360/n$ degrees with respect to the previous one. This insight leads to the *multi-way ICP* algorithm, which aligns all scans simultaneously with *one* transformation (Figure 5).

In multi-way ICP, an initial rotation is estimated based on the number of scans and the turntable's approximate axis of rotation (the exact axis is recovered by the optimization). Corresponding point pairs are selected between each successive pair of scans under the assumption that each scan aligns to the previous one using the estimated rotation. The last scan is assumed to align to the first one, forming a closed loop. At each iteration, a new axis of rotation is computed that minimizes the alignment of all point pairs,
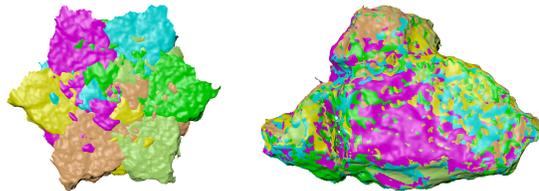

Figure 5: *Because all six scans were acquired on a turntable, each is rotated 60 degrees with respect to the previous one (left). Using multi-way ICP, we solve for the single rotation that simultaneously aligns all scans (right).*
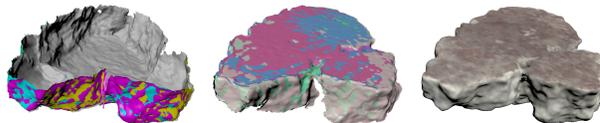

Figure 6: *Front and back scan sequences and the final model. Because the front and back scans overlap only on the vertical edges, alignment of front to back is unstable in the vertical direction. We compute the vertical alignment directly by detecting the front surface, eliminating the instability.*

as described in Appendix A. Even if each pair of scans is *individually* unstable, we can still extract a rotation from the information distributed among *all* scans because we simultaneously consider point pairs between all successive scans. Using multi-way ICP, we correctly aligned 149 out of 150 fragments; the remaining fragment failed because it moved during scanning.

Front-Back Alignment   Unlike alignment of just a front or back sequence, there is no good initial estimate of the relative alignment of fronts to backs. Moreover, there is little overlap between front and back scans — while vertical edges are visible in both sequences, angled edges are only visible in one. Therefore, instead of using a feature-based alignment technique such as spin images [Johnson and Hebert 1997], we directly extract the front surface.

The front surface's flatness, a hindrance for ICP, makes it easy to fit with a plane, using iterative outlier rejection. Of course, the front surface is not visible on the back scans because it is lying flat on the turntable. Since we know the plane of the turntable from calibration, we can rotate the front scans to also lie in this plane, which we define to be the $z = 0$ plane. Any remaining misalignment between front and back scans is an xy planar translation and z rotation. The translation is estimated by aligning centers of mass, and rotation is estimated by trying 40 rotations. From this estimate, we use ICP constrained to a planar transform to refine the alignment (Figure 6).

Out of 150 scans, the front surface was incorrectly detected only once. For 13 scans, front-to-back alignment failed. Using our verification interface, each error was corrected in seconds.

4.2 2-D to 3-D Registration

Our 3-D scanner provides color registered to each range scan, which we project onto the final model to obtain a colored result. However, the scanner's camera provides only limited color fidelity, lighting is uneven, and resolution is limited. High resolution color obtained under controlled illumination is important for archival purposes, for studying the wall painting's iconography, for matching based on color and texture, and for visually verifying matches.

As noted in Section 2, the front surface is difficult to isolate from edge data included in the scanned images. The scanned edges also do not align to a projected image of the 3-D model, making silhouette-based alignments [Lensch et al. 2000] unsuitable, because the scanner's depth-of-field is too limited to acquire the entire edge, and the linear light source is highly directional. Feature-based methods are also ill-adapted because the many solid-colored fragments lack stable image features.

Even alignment methods based on color difference or cross correlation are stymied by the extraneous edge data. While we could optimize for the scanner's projection and lighting model, there is a faster and more robust solution: register the scanned image to a projection of *only* the 3-D model's front surface. Recall that the model's front surface is aligned to the $z = 0$ plane, and can be extracted by rendering with the far clipping plane just behind it.

The projected image aligns to a subset of the scanned image (which contains no edges), so we do not penalize scanned image pixels overlapping the projected image's background. In the inverse case, in which a projected image pixel overlaps the scanned image's background, we do assess a penalty. Aside from this, we optimize the alignment using a standard, normalized cross-correlation. To accommodate minor misalignment or misorientation of the 3-D fragment, we solve for an affine transformation (with scale limited to $\pm 2\%$), rather than a rigid one. Figure 2 shows a sample result.

To obtain an initial alignment estimate, we use PCA for rotation and center-of-mass for translation. Because 2-D scans include edge data, we have found the estimate is more reliable if we use a projected image that also includes edges for this stage. We improve the initial guess further by trying rotations every $4°$ within a range of $20°$ around the estimated PCA axis, and translations every 4 mm up to 20 mm from the center of mass.

Using this procedure we were able to correctly align 2-D scans for 149 of our 150 test fragments automatically. The remaining fragment was corrected by roughly positioning the 2-D scan using our correction interface, and re-optimizing.

4.3 Normal Reconstruction

We obtain normal maps via shape from shading, beginning with several (typically four) scans of a fragment at different orientations. Our approach is conceptually straightforward: during a calibration phase we measure $I(\mathbf{n})$, the observed brightness as a function of the surface normal. Then, given a set of scans, we invert I to solve for the normal. As a side effect, we obtain the true color texture (diffuse reflectance) of the surface.

Calibration   Because the scanner's light source is linear, not a point, $I(\mathbf{n})$ will not be of the form $\mathbf{n} \cdot \mathbf{l}$. We therefore measure it in a calibration step, using a diffuse plane inclined at known angles with respect to the scanner platform. We measure the intensity at several angles of tilt and rotation to ensure good coverage of the space of normals. Because intensity varies with height off the scan bed, we perform all measurements at a constant height.

We fit a spherical-harmonic model to the diffuse plane measurements to obtain a parametric representation of $I(\mathbf{n})$ that averages out measurement noise and is more easily inverted. Using cross-validation, we have determined that second-order spherical harmonics do not substantially decrease the calibration error over a first-order representation containing only constant and linear terms in the normal components:

$$I(\mathbf{n}) = (a_0 \; a_1 \; a_2 \; a_3) \begin{pmatrix} n_x \\ n_y \\ n_z \\ 1 \end{pmatrix} = \mathbf{a}^T \mathbf{n}. \qquad (1)$$

Normal Computation   After capture, we register the scanned images using the algorithm of Section 4.2. The inverses of the resulting rotations, which we will call R_i, may be thought of as rotating
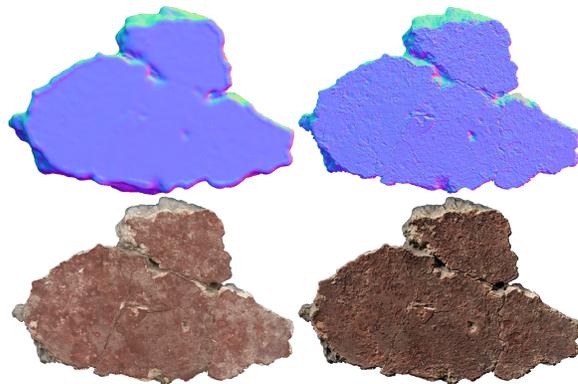

Figure 7: *Computed normals (top right) reveal more surface detail than those extracted from the geometry (top left). Combining the high-resolution normals with the extracted RGB color (bottom left) allows for enhanced visualizations [Toler-Franklin et al. 2007] of surface detail (bottom right).*
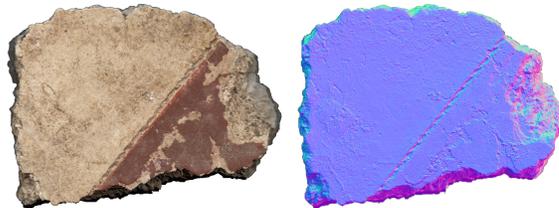

Figure 8: *String impressions, most clearly visible in the computed normals, mark boundaries of solid color; they are an important cue for reconstruction, restoration, and archaeological study.*
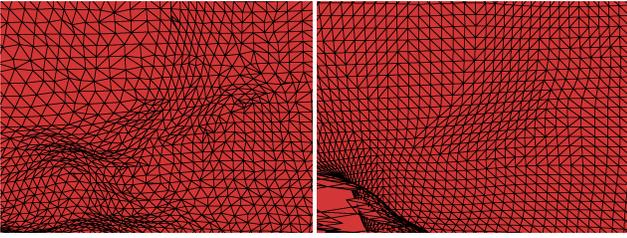
Figure 9: *Wireframe closeup of the original model (left), together with the corresponding regularly-sampled ribbon (right).*

Figure 10: *A strip of samples on each ribbon is used to compute a candidate alignment. Then, the overlap region is shifted one sample and a new alignment is computed incrementally in constant time.*

the light relative to each pixel; in other words, the R_i act on the calibration coefficient vector **a**. For each image i we may write

$$I_i = \rho (R_i \mathbf{a})^T \mathbf{n} , \qquad (2)$$

where ρ is the diffuse albedo. In matrix form, this expands to

$$\begin{pmatrix} - (R_1 \mathbf{a})^T - \\ - (R_2 \mathbf{a})^T - \\ \vdots \end{pmatrix} \begin{pmatrix} \rho n_x \\ \rho n_y \\ \rho n_z \\ \rho \end{pmatrix} = \begin{pmatrix} I_1 \\ I_2 \\ \vdots \end{pmatrix} . \qquad (3)$$

Note that it is not possible to solve this problem as stated using standard linear least squares, as the leftmost matrix in equation 3 (let us call it A) is singular: both its last and next-to-last columns will be constant across all rows, because all of the R_i represent planar rotations that leave the z component unchanged. Instead, we define the matrix $\tilde{A}$ to be A with its fourth column removed, giving

$$\tilde{A} \begin{pmatrix} \rho n_x \\ \rho n_y \\ \rho \tilde{n}_z \end{pmatrix} = \begin{pmatrix} I_1 \\ I_2 \\ \vdots \end{pmatrix} , \qquad (4)$$

where $\tilde{n}_z = n_z + a_3 \, \rho / a_2$. Equation 4 can be solved using linear least squares, and by adding the constraint that

$$n_x^2 + n_y^2 + n_z^2 = 1 , \qquad (5)$$

we are able to to solve for ρ, n_x, n_y, and n_z.

Results   The normal maps we recover have an order of magnitude higher resolution than our geometry, giving us the ability to document, visualize, and analyze fine surface detail (Figure 7), including hairline cracks, string impressions (Figure 8), plaster grain, and impressions of reeds or other materials.

5   Matching

Among the many applications for detailed fragment models, one of the most important is matching. Existing matching algorithms use a variety of different cues — color and texture [Fornasier and Toniolo 2005; Sağiroğlu and Erçil 2006], contours [Kong and Kimia 2001; Leitão and Stolfi 2002; Papaodysseus et al. 2002], and 3-D shape [Huang et al. 2006] are some of the most common. All of these algorithms can be performed on the data our system captures.

We have chosen to start with a 3-D shape-matching algorithm. The reasons are twofold. First, archaeologists and conservators can match fragments with decoration more easily than the many with no decoration; color- and texture-based matchers therefore provide less practical benefit. Second, matches are always verified by testing if the fragments "lock" together — i.e. the 3-D edge geometry is the most reliable matching cue. Matching fragments that lock together are not necessarily erosion-free; as long as *some* part of the matching interface is not eroded, there will be a snug fit.

Although we match based on 3-D edge geometry, we again take advantage of the fragments' flat front surfaces to limit our search space to planar transformations. A brute-force approach is to select a pair of points on the two fragments, compute an initial alignment based on the points and their normals, run ICP (constrained to a planar transform) to recover a candidate matching alignment and error, and repeat the process for points distributed around each fragment.
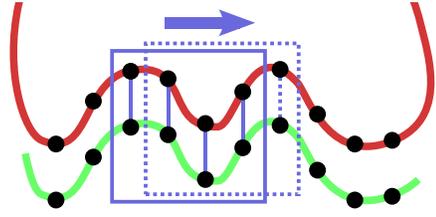
Repeated ICP alignments, however, are slow (requiring, on average, 45 seconds per fragment on our test data set), and the ill-constrained geometry produces instability in ICP. If many accurate point pairs could be determined quickly, there would be no need to iterate, and the process would be much faster. Eliminating iteration from the alignment also resolves the stability problems.

Suppose we had a mesh of only a fragment's edges, which we call a "ribbon," and these ribbons had regular sample spacing (Figure 9). Each ribbon vertex could be indexed by row and column, so we could select a strip of samples of a fixed width on each fragment, and ask whether they match. Now the correspondences are completely determined by the regular sample structure, and all that remains is to compute the alignment and associated error. The overlap region is then shifted by a single sample, and the process is repeated (Figure 10). This is similar to a 2-D convolution, and indeed each new alignment and error can be incrementally computed from the previous one in constant time (see Appendix B for details). The computational cost is therefore $O(nm)$ where n and m are the edge lengths of the two fragments. The ribbon matcher aligns a pair of fragments in an average of two seconds, irrespective of the fragments' thicknesses and of the width of the matching strip.

Ribbon Construction   We construct ribbons in two stages, as shown in Figure 11. First we extract a contour 2 mm from the front surface (to avoid erosion), placing samples every 0.25 mm in arclength. We account for noise by smoothing the extracted contour using a standard deviation of 2.5 mm, and reprojecting the smoothed points back onto the fragment edge along their normals. Next, we walk vertically from each contour point, adding samples every 0.25 mm along the z-axis. We stop when we reach a face pointing away from the contour point's normal. This procedure places isolated points on the fragment's front and back surfaces, which we prune by triangulating the ribbon and keeping only connected components containing a contour point.

We must walk vertically from each contour point within a particular plane. As the edge advances and recedes from the contour point, where we sample will depend on which plane we choose (Figure 12). If two fragments match, but we chose to sample the ribbons in different planes, we will not get good correspondences. As a result, we may compute a poor alignment, obtaining a higher associated error. To prevent this, we walk in the plane defined by the contour's *smoothed* normal and the z-axis.

Because each column of samples is constructed independently, the ribbons may self-intersect. In practice this occurs rarely, and in any case it does not affect matching: if one ribbon contains self-intersections, any corresponding ribbons will contain the same self-intersections, and all samples will be matched properly.

Erosion Detection   Many fragments exhibit some erosion, especially near the back surface. In eroded areas, correspondences assigned by the ribbon matcher will be wrong, leading to incorrect alignments (Figure 13). ICP eliminates poor correspondences using a normal compatibility constraint — any correspondence in which the normals are not nearly opposite each other is rejected. That constraint requires an approximate initial alignment, but we are able to use a weaker constraint that proves effective. Although a
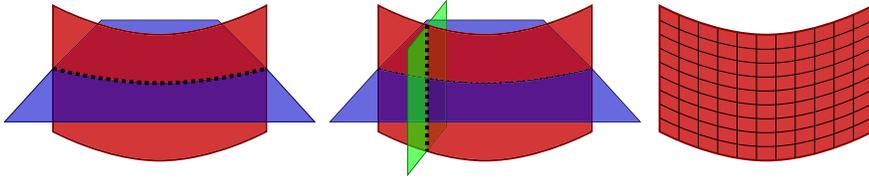
Figure 11: *To efficiently compute fragment matches, we regularly resample fragment edges into a "ribbon." A contour is extracted at a fixed offset from the front surface (left), then each sample is extruded vertically in a plane defined by the contour point's smoothed normal (center). Ribbon points are arranged in a grid (right), allowing efficient computation of correspondences.*
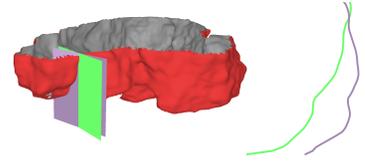

Figure 12: *A ribbon, with profiles extracted using smoothed and unsmoothed normals. We use smoothed normals to ensure consistency among fragments.*
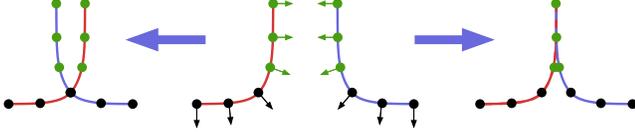
Figure 13: *Using all corresponding point pairs on the red and blue curves yields an incorrect alignment (left) because the black points are on eroded portions of the edge. Using only the green points, whose normals' z-components are opposing, yields a correct alignment (right).*

point's normal will be affected by the alignment, its z-component is constant because the transformation is always planar. Therefore the corresponding points' normals should have opposite z-components even before alignment. Specifically, we require that $|n_z + n'_z| \leq 0.5$. Significant erosion occurs mostly near the back surface, so normals in eroded areas tend to point down on both fragments. Since they are not opposing, they will be pruned.

Thickness Compatibility   In general, matching fragments are the same thickness. By counting the number of vertical positions along corresponding columns where one ribbon contains a sample but the other does not, we obtain an estimate of the difference in thickness on a per-column basis. We apply a fixed penalty to each unmatched sample beyond a fixed per-column threshold. Typically, we allow a maximum difference of 4 mm (16 samples) in height between fragments, and apply a penalty of 3 mm to each point exceeding the threshold (i.e. each such point is counted as having an additional 3 mm of error when determining the final alignment error).

5.1 Matching Results

We have applied our matching algorithm to three test sets of fragments acquired at the Akrotiri excavations from the wall painting of spirals, as well as a "ground truth" fresco (which we call the synthetic fresco), all prepared by the conservators to test our system. The test sets contain some known matches found by the conservators, with the possibility that additional matches might exist. To test the generality of our matching method, we also applied it to a set of fragments from the *Forma Urbis Romae* [Koller et al. 2006], an ancient marble map of Rome.

Data Set	Number of Fragments	Known Matches Found/Total
Red fragments	134	2 / 2
Spiral fragments	44	6 / 7
White fragments	105	1 / 3
Synthetic (6.25 mm)	129	112 / 253
Synthetic (12.5 mm)	129	101 / 253
Synthetic (25 mm)	129	138 / 253
Synthetic (50 mm)	129	101 / 253
Synthetic (Combined)	129	175 / 253
Synthetic (ICP)	129	124 / 253

Table 1: *Matches found on the wall painting data sets using the ribbon and ICP matchers. We also found two additional matches among the white fragments that were previously unknown.*

Wall Painting Matching Results   We ran the ribbon matcher on all pairs of fragments within each test set, producing a list of candidate matches ordered by alignment error. The results are summarized in Table 1. For the "real" fragments, we visually examined the top matches, and sent images of the ones we believed could be correct to the conservators for verification against the actual fragments. We used a strip width of 25 mm. For the synthetic fresco, we used strip widths of 6.25 mm, 12.5 mm, 25 mm, 50 mm, as well as the combined results of all four strip widths. We examined the proposed alignment for each matching pair to see if it was correct. Finally, we also tested the ICP matching approach described above on the synthetic fresco. The ribbon matcher not only finds more matches than the ICP matcher, it also ranks them significantly higher in the list of candidate matches than does the ICP matcher [Brown 2008], and is faster: an average of 2 seconds per pair of fragments vs. 45 seconds for ICP matching.

The matches we found in the red, spiral, and white data sets are shown in Figures 1, 15, and 16. These sets are drawn from the lower, monochrome zone of the composition, the main spiral motif, and the unpainted lime-plaster background, respectively. We also show the entire synthetic fresco (of which the numbered fragments have been scanned) in Figure 14; red lines indicate matches found with a 25 mm strip width, while blue and green lines indicate the *additional* matches found with 12.5 mm and 50 mm strip widths.

Forma Urbis Romae   Our *Forma Urbis* test set contains eight uninscribed fragments, for which edge geometry is the only definitive matching cue. This set contains three representative matches, all of which we found (Figure 17). We believe this indicates that our algorithm can be practical in a broad range of contexts.

Forma Urbis fragments are thicker than our wall painting fragments—about 5 cm thick—but because of erosion, matches usually occur along only a small portion of the edge's height. We therefore generated 15 mm-high ribbons at different depths along the fragment edge, and matched each set of ribbons independently. Also, because it is impossible to distinguish the front and back surfaces of uninscribed fragments (both sides are perfectly flat), we included each fragment twice, rotated 180° with respect to each other. We used a strip width of 25 mm.

6 Discussion and Future Work

Based on the results of our tests, we are planning a fixed installation of our system at the excavations, to scan and match large numbers of fragments. At the same time, there are several directions for improving our system and expanding its range of applications.

Efficiency and Automation   Relative to the effort of excavating, stabilizing, and conserving each fragment, we believe our acquisition speed of roughly 10 fragments per hour is a sufficiently small incremental cost to be practical for documenting *all* excavated items. We have therefore not pursued methods for further automating the scanning process, which would require custom hardware (at significant extra cost) and provide little net benefit. We have also retained the user in the processing loop, seeking only to make it suf-
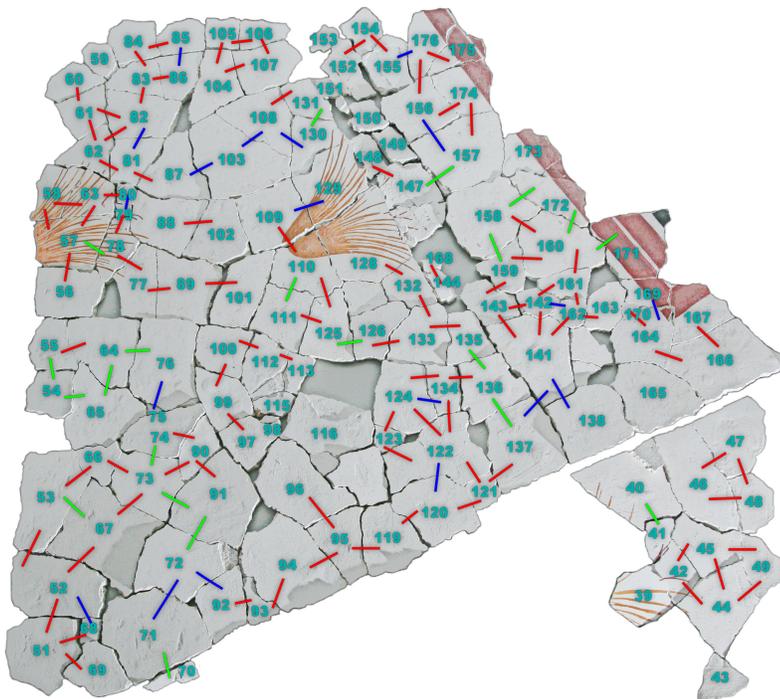
Figure 14: *The "ground truth" synthetic fresco. Red lines indicate matches found using the ribbon matcher with a 25 mm strip width. Blue links indicate additional matches found with a 12.5 mm strip width, and green links indicate further matches found with a 50 mm strip width. Only numbered fragments have been scanned.*

Figure 15: *The six matches from the spiral data set, found with the ribbon matcher. The matcher considers only edge geometry; the decoration's continuity just confirms success. (We did not find the match between fragments 170 and 174.)*

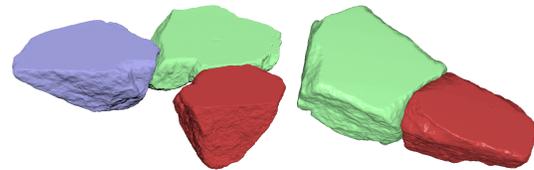**Figure 16:** *The three matches from the white data set.*

Figure 17: *Matches in the* Forma Urbis Romae *test set.*

ficiently robust that the required human time is, on average, seconds per fragment (most of the manual effort arises during acquisition).

Generalizability We believe most components of our hardware and processing setup are sufficiently general to be relevant in other applications. The acquisition setup, with a flatbed color scanner and a 3-D scanner oriented at 45 degrees relative to a motorized turntable, is applicable in many contexts, provided that object size is limited to a few tens of centimeters and high-resolution color is required only in flat regions. Our processing pipeline is also generic: multi-way ICP is applicable whenever a turntable is used, as is plane-fitting followed by constrained ICP whenever objects have known-flat regions. While our ribbon-based incremental matcher is most useful for reconstruction of fractured flat surfaces, wall paintings and mosaics are common finds at archaeological sites, as are incised and carved panels of wood, marble, or other stone material.

Nearly Flat Objects Our system relies on a flat front surface for both 3-D alignment and normal reconstruction; although flat surfaces are common, millimeter-scale deviations from the surface plane are also common. The plane-fitting required for 3-D alignment is robust to these deviations, but they result in small image misalignments, which show up as texture embossing when we compute normals. Given the detail present in our 2-D scans, we believe a non-rigid image registration or a shape-from-shading reconstruction should be possible, and is an interesting area for future work.

Multi-Cue Matching The data acquired by our system naturally lend themselves to multi-cue matching. Although we have begun with 3-D shape matching, conservators and archaeologists currently consider many cues including the excavation context; texture, quality, and thickness of the plaster; pigment texture and brightness; depicted motif and continuity of color, pattern and incisions between fragments; and the state of preservation of both front and back surfaces. We are investigating intuitive interfaces for specifying and evaluating multi-cue queries; we believe that this will allow us to propose more relevant matches.

Acknowledgments

References

ARUN, K. S., HUANG, T. S., AND BLOSTEIN, S. D. 1987. Least-Squares Fitting of Two 3-D Point Sets. *IEEE Trans. PAMI, Vol. 9, No. 5*, 698–700.

BERNARDINI, F., RUSHMEIER, H., MARTIN, I. M., MITTLEMAN, J., AND TAUBIN, G. 2002. Building a Digital Model of Michelangelo's Florentine Pietà. *IEEE Computer Graphics and Applications, Vol. 22, No. 1*, 59–67.

BESL, P. J., AND MCKAY, N. D. 1992. A Method for Registration of 3-D Shapes. *IEEE Trans. PAMI, Vol. 14, No. 2*, 239–256.

BRACCI, S., FALLETTI, F., MATTEINI, M., AND SCOPIGNO, R., Eds. 2004. *Exploring David: Diagnostic Tests and State of Conservation*. Giunti Press, Florence, Italy.

BROWN, B. J. 2008. *Registration and Matching of Large Geometric Datasets for Cultural Heritage Applications*. Ph. D. thesis, Princeton Univ.

CHEN, Y., AND MEDIONI, G. 1992. Object Modeling by Registration of Multiple Range Images. *Image and Vision Computing, Vol. 10, No. 3*, 145–155.

DOUMAS, C. 1992. *The Wall-Paintings of Thera*. Thera Foundation – P. M. Nomikos, Athens.

FORNASIER, M., AND TONIOLO, D. 2005. Fast, Robust and Efficient 2D Pattern Recognition for Re-Assembling Fragmented Images. *Pattern Recognition, Vol. 38, No. 11*, 2074–2087.

GARDNER, A., TCHOU, C., HAWKINS, T., AND DEBEVEC, P. 2003. Linear Light Source Reflectometry. *ACM Trans. Graphics (Proc. SIGGRAPH), Vol. 22, No. 3*, 749–758.

GELFAND, N., IKEMOTO, L., RUSINKIEWICZ, S., AND LEVOY, M. 2003. Geometrically Stable Sampling for the ICP Algorithm. In *Proc. 3DIM*, 260–267.

HUANG, Q.-X., FLÖRY, S., GELFAND, N., HOFER, M., AND POTTMANN, H. 2006. Reassembling Fractured Objects by Geometric Matching. *ACM Trans. Graphics (Proc. SIGGRAPH), Vol. 25, No. 3*, 569–578.

JOHNSON, A., AND HEBERT, M. 1997. Surface Registration by Matching Oriented Points. In *Proc. 3DIM*, 121–128.

KARASIK, A., AND SMILANSKY, U. 2008. 3D Scanning Technology as a Standard Archaeological Tool for Pottery Analysis: Practice and Theory. *Journal of Archaeological Science, Vol. 35*, 1148–1168.

KOLLER, D., TRIMBLE, J., NAJBJERG, T., GELFAND, N., AND LEVOY, M. 2006. Fragments of the City: Stanford's Digital Forma Urbis Romae Project. In *Proc. Third Williams Symposium on Classical Architecture, Journal of Roman Archaeology*, vol. Suppl. 61, 237–252.

KONG, W., AND KIMIA, B. 2001. On Solving 2D and 3D Puzzles under Curve Matching. In *Proc. CVPR*, vol. 2, 583–590.

LEITÃO, H. C. G., AND STOLFI, J. 2002. A Multiscale Method for the Reassembly of Two-Dimensional Fragmented Objects. *IEEE Trans. PAMI, Vol. 24, No. 9*, 1239–1251.

LENSCH, H. P., HEIDRICH, W., AND SEIDEL, H. 2000. Automated Texture Registration and Stitching for Real World Models. In *Proc. Pacific Graphics*, 317–326.

LEVOY, M., PULLI, K., CURLESS, B., RUSINKIEWICZ, S., KOLLER, D., PEREIRA, L., GINZTON, M., ANDERSON, S., DAVIS, J., GINSBERG, J., SHADE, J., AND FULK, D. 2000. The Digital Michelangelo Project: 3-D Scanning of Large Statues. In *Proc. SIGGRAPH*, 131–144.

LOWE, D. G. 2004. Distinctive Image Features from Scale-Invariant Keypoints. *IJCV, Vol. 60, No. 2*, 91–110.

NEHAB, D., RUSINKIEWICZ, S., DAVIS, J., AND RAMAMOORTHI, R. 2005. Efficiently Combining Positions and Normals for Precise 3D Geometry. *ACM Trans. Graphics (Proc. SIGGRAPH), Vol. 24, No. 3*, 536–543.

PAPAODYSSEUS, C., PANAGOPOULOS, T., EXARHOS, M., TRIANTAFILLOU, C., FRAGOULIS, D., AND DOUMAS, C. 2002. Contour-Shape Based Reconstruction of Fragmented, 1600 BC Wallpaintings. *IEEE Trans. on Signal Processing, Vol. 50, No. 6*, 1277–1288.

RUSINKIEWICZ, S., AND LEVOY, M. 2001. Efficient Variants of the ICP Algorithm. In *Proc. 3DIM*, 145–152.

SAĞIROĞLU, M. Ş., AND ERÇIL, A. 2006. A Texture Based Matching Approach for Automated Assembly of Puzzles. In *Proc. ICPR*, vol. 3, 1036–1041.

TOLER-FRANKLIN, C., FINKELSTEIN, A., AND RUSINKIEWICZ, S. 2007. Illustration of Complex Real-World Objects using Images with Normals. In *Proc. NPAR*, 111–119.

VLACHOPOULOS, A. 2008. The Wall Paintings from the Xeste 3 building at Akrotiri. Towards an Interpretation of the Iconographic Programme. In *Horizons: A colloquium on the prehistory of the Cyclades, Cambridge*, N. Brodie, J. Doole, G. Gavalas, and C. Renfrew, Eds., 451–465.

WASSERMAN, J., CAMIZ, F. T., VERDON, T., AND ROCKWELL, P. 2002. *Michelangelo's Florence Pietà*. Princeton Univ. Press.

WILLIS, A. 2004. *Stochastic 3D Geometric Models for Classification, Deformation, and Estimation*. Ph. D. thesis, Brown Univ.

WOODHAM, R. J. 1980. Photometric Method for Determining Surface Orientation from Multiple Images. *Optical Engineering, Vol. 19, No. 1*, 139–144.

A Multi-Way ICP Computation

The basic computation in multi-way ICP is solving for a rotation of $360/n$ degrees about *some* axis that minimizes point-to-plane distance between corresponding points. We first find an incremental rigid-body transformation, apply it to the current transformation, then coerce the result to be a rotation by $360/n$ degrees. We decompose the transformation into a rotation and translation:

$$\begin{pmatrix} a & d & g & j \\ b & e & h & k \\ c & f & i & l \\ 0 & 0 & 0 & 1 \end{pmatrix} \rightarrow \begin{cases} axis = (f - h, \; g - c, \; b - d) \\ angle = \text{atan2}(\|axis\|, \; a + e + i - 1) \\ \mathbf{t} = (j, \; k, \; l) \end{cases}.$$

$$(6)$$

We then set

$$angle' = 2\pi/n, \qquad \mathbf{t}' = \mathbf{t} - \frac{axis\,(\mathbf{t} \cdot axis)}{\|axis\|^2}, \qquad (7)$$

and reconstruct the transformation using Rodrigues's formula.

B Incremental Matching Details

Given two sets of corresponding points P and Q, we wish to find the sum of squared distances between them under the optimal rigid-body alignment TR. Defining the covariance matrix C as

$$C = \sum \tilde{\mathbf{p}}\tilde{\mathbf{q}}^T, \quad \text{where } \tilde{\mathbf{p}} = \mathbf{p} - \bar{\mathbf{p}}, \; \tilde{\mathbf{q}} = \mathbf{q} - \bar{\mathbf{q}}, \qquad (8)$$

$R = UV^T$ where USV^T is the singular value decomposition of C, and $T = \bar{\mathbf{p}} - R\bar{\mathbf{q}}$ [Arun et al. 1987].

In the ribbon matcher context, P and Q are strips from two ribbons A and B. Each time Q is shifted one sample around B, one column of points is discarded from the beginning of Q, and a new column is added to the end. For each corresponding column of A and B, we precompute $\sum \mathbf{pq}$, $\sum \mathbf{p}$, and $\sum \mathbf{q}$. Using the relation

$$\sum (x - \bar{x})(y - \bar{y}) = \sum xy - \bar{x}\sum y \qquad (9)$$

(the remaining terms in the expansion cancel, because $\sum x = \sum \bar{x}$), we generate the new covariance matrix and alignment.

The incremental error computation follows the same principle. The mean squared error between P and Q is

$$E = \sum_i \|\tilde{\mathbf{q}}_i - R\tilde{\mathbf{p}}_i\|^2 = \sum_i \left(\|\tilde{\mathbf{q}}_i\|^2 + \|R\tilde{\mathbf{p}}_i\|^2 + 2\tilde{\mathbf{q}}_i^T R\tilde{\mathbf{p}}_i \right). \quad (10)$$

Because R is a rotation matrix, $\|R\tilde{\mathbf{p}}_i\|^2 = \|\tilde{\mathbf{p}}_i\|^2$. This gives

$$E = \sum_i \|\tilde{\mathbf{p}}_i\|^2 + \sum_i \|\tilde{\mathbf{q}}_i\|^2 - 2\sum_i \tilde{\mathbf{q}}_i^T R\tilde{\mathbf{p}}_i, \qquad (11)$$

in which the first two terms are easily precomputed (the mean is subtracted from the sum using the identity from Equation 9). The third term expands to

$$R_{00}\tilde{p}_x\tilde{q}_x + R_{01}\tilde{p}_y\tilde{q}_x + R_{10}\tilde{p}_y\tilde{q}_x + R_{11}\tilde{p}_y\tilde{q}_y = \text{tr}(RC). \qquad (12)$$

Note that a similar incremental computation could be applied to any polynomial error metric, including point-to-plane.