The SHAPE Lab :

New Technology

and

Software for Archaeologists

April 2000

Authors:

1. Frederic F. Leymarie,

   PhD candidate, Division of Engineering, Brown University

   R&D Project Leader, The SHAPE Lab

   Box D, 182-4 Hope St., Providence, RI, 02912 USA

   Tel: +1.401.863.2760 Fax: +1.401.863.9039 email: leymarie@lems.brown.edu

2. David B. Cooper,

   Professor, Division of Engineering, Brown University, Box D

   cooper@lems.brown.edu

3. Martha Sharp Joukowsky,

   Professor, Center for Old World Archaeology and Art, Department of Anthropology,

   Director, Petra Great Temple Excavations

   Brown University, Box 1921

   Martha_Joukowsky@brown.edu

4. Benjamin B. Kimia,

   Associate Professor, Division of Engineering, Brown University, Box D

   kimia@lems.brown.edu

5. David H. Laidlaw,

   Assistant Professor, Dept. of Computer Science, Brown University, Box 1910

   dhl@cs.brown.edu

6. David Mumford,

   Professor, Division of Applied Mathematics, Brown University, Box F

   mumford@dam.brown.edu

7. Eileen L. Vote

   PhD candidate, Dept. of the History of Art and Architecture, Brown University,

   vote@lems.brown.edu

Abstract

The SHAPE Lab was recently established (1999), with a grant from the United States National Science Foundation, by Brown University Departments of Engineering, Applied Mathematics, Computer Science and The Center for Old World Archaeology and Art and Department of Anthropology. It is a significant interdisciplinary effort for scientific research with a direct application to important problems in the analysis of archaeological finds and artifacts. We present the concepts that will underlie a 3D shape language, and an interactive, mixed-initiative system, for the recovery of 3D free-form object and selected scene structure from one or more images and video. This work has impact by providing new practical tools. It also provides an effective testbed for 3D shape reconstruction and recognition, more descriptive local and global models for working with 3D shapes and performing free-form geometric modeling, and for extracting 3D geometry from one or more images and video, as well as associated computational complexity issues. As applied to the field of archaeology, this technology provides, specifically, new ways to analyze and reconstruct pottery, compare objects from different sites and reconstruct sculpture and architecture.

Key words: Shape language, 3D free-form modeling, ridges and valleys for shape, implicit polynomial models, medial axes, skeletal graphs, 3D object reconstruction and analysis, Great Temple of Petra, archaeological recordings.

Contents

1 Introduction

The SHAPE[1] Lab was recently established for the synergistic study of three dimensional (3D) free-form in the disciplines of mathematics, computer graphics, computer vision and archaeology. It answers the needs that arise in the domain of archaeology, but which are also generic across a range of applications. Specifically, we are investigating the following:

1. 3D free-form modeling for surface and volume representation, via the design of a shape language.
2. Geometric information extraction from either: (i) passive optical systems, such as obtained via a single image, many images, or a video stream, (ii) active data-capture systems, such as laser camera systems or structured light systems, and (iii) a combination of data obtained via (i) and (ii) together with auxiliary data when available (e.g. floorplans, survey data).
3. Human/computer interaction (HCI) for facilitating the model building and geometric information extraction, as well as to provide an interactive virtual system for archaeological analysis with site features, topography, architecture artifacts and special finds.
4. Decision-directed machine estimation for automatic model choice and geometric information extraction, with an emphasis on the important archaeological problem of stitching together fragments constitutive of an original object, such as obtained from pottery sherds.

[Figure 1 about here.]

These research topics are to be explored through an integrated effort because, on the one hand, the archaeology applications drive the 3D modeling and 3D from images research, and, on the other hand, the research on 3D free-form provides archaeologists with tools to be able to conduct research heretofore impossible or impractical. In addition, in light of the fact that the three disciplines of mathematics, computer graphics and computer vision view 3D free-form in various perspectives, ranging from theoretical to practical, this project provides a unique opportunity to conduct the study in a comprehensive way. Finally, we benefit from the direct involvement of our team of archaeologists at the site of the Great Temple in Petra, Jordan (Fig. 1) where on-going excavations have been conducted for many years, and from which a large database of artifacts is currently being built [29].

1.1 Archaeological Research Problems

We focus here on the development of a generic technology for the recovery of 3D models which is investigated in the context of archaeology. Archaeologists are typically faced with a series of bottlenecks, including the following ones, which this research aims to alleviate.

[Figure 2 about here.]

1. Excavators want to be able to register the location of artifacts in situ in order to maintain an accurate archaeological excavation record (see Fig. 2 and cf. § 2). Our proposed technology will allow archaeologists to use relatively inexpensive equipment to expedite excavations and maintain more comprehensive, accurate and accessible records of artifact *geometry* and find locations.
2. Currently artists assist in work on site by documenting the artifacts found and positing reconstructions of broken artifacts, thus leaving archaeologists out of the process with additional delays and much added cost. Our proposed technology allows the archaeologist to use shape models and computer graphics to document and interactively reconstruct artifacts.
3. A significant problem in archaeology is the inability to compare many artifacts stylistically, which requires substantial physical information (e.g., see Fig. 3). Relating one artifact to another, perhaps found in another site, is an integral part of discovering its role, age, responsible artisan or community, etc. The expression of artifacts in a *shape language* will advance possibilities for interactive or automatic quantitative and qualitative comparison.

[Figure 3 about here.]

We rely mainly on image analysis and photogrammetric methods in order to reconstruct and measure the 3D structure of objects. The use of a passive optical data acquisition technology, in contrast to active scanners, is of interest in order to:

- acquire data on-site at low cost, without imposing hard constraints on the size of objects or the ambient lighting, and without slowing down the excavation campaign;
- use existing image databases from previous excavation campaigns and from other sites.

[Figure 4 about here.]

However, we also make use of active data acquisition techniques, such as laser scanning, structured light and computerized tomography (CT, see Fig. 4) systems, in order to:

- provide "ground truth" measurements upon which we can gauge passive reconstruction techniques;
- rapidly acquire 3D data in order to conduct our other research objectives;
- maintain an expertise in using both types of systems, and keep track of their evolving differences.

The last point is emphasized by the fact that digital photogrammetry has yet to become automated, while laser camera remain relatively expensive, structured light systems have limited applications in the field (*i.e.*, constrained lighting conditions and limited field of view), and tomography is not a portable technique. See also the work of Pollefeys *et al.* [48] who advocate the use of both passive and active systems for different purposes in documenting archaeological sites.

1.2 Shape Modeling Research Problems

Our premise is that the use of more powerful 3D shape representations than the classical points, straight lines, planes, triangles or splines, can lead to practical solutions for the preceding problems, and markedly improve speed, accuracy, and user convenience over most of what is presently possible. The 3D representations we propose to study are hybrid constructions made of *ridges*, *implicit polynomial surfaces* (IPS, *i.e.*, algebraic surfaces) and *skeletal graphs*. They are studied for use individually and in concert, in order to understand their most effective synergy as new hybrid models and algorithms are developed.

Of course, the guiding principle in this work is discovering and understanding the fundamental issues in solving these complex problems in computationally fast, yet user-friendly ways. We seek to identify the most effective ways of handling and processing the huge amounts of data available in the acquired images and video streams, in particular, by identifying tradeoffs between accuracy and complexity. We emphasize that our program for the study of 3D free-form representations for shape:

1. Solves heretofore unsolved problems.
2. Improves speed and user convenience in handling complex problems.
3. Handles huge amounts of data in new and faster ways.

In the remaining of this paper we first describe, in § 2, the present day situation at our main site of excavation, at Petra. Then, in § 3, we give some early results in tackling the previously introduced research objectives. Finally, in § 4, we describe in some detail the basis for our shape language developed to tackle complex archaeological problems.

2 Archaeology at Petra

The Great Temple of Petra, Jordan, is a monolithic structure at the top of a three-leveled precinct the size of a football field (Fig. 1). In unearthing a site such as this, archaeologists want to use the most exact technology to register objects they excavate, and reconstructive technology to help them envision what the building and the objects within looked like [25, 26, 27, 28, 29, 30, 58]. Our proposed technology aims to help them do both.

The latest archaeological standard for gathering data about finds is to register each object as it is excavated with a costly laser transit station. This requires three people to digitally register the object with the survey equipment; one to shoot the point, another to hold the prism in order to register the location, and a third to label and bag the artifact. Even with this method there is no easy way to correlate the object with the survey. Some archaeologists register an object with one point, indicating the approximate centroid of an object; others take four or five points (or more) per object in an attempt to give additional information about the object's shape and orientation. In a typical excavation, relevant finds need to be registered with the survey station in different locations at once and excavators wait for the surveyors to register their objects before they can proceed with digging. Furthermore, all of the information regarding the find spot must be retrieved in the field. After the survey station has registered the object, it must go through other phases of registration. All artifacts are hand measured, drawn by a site artist, photographed and then put into a database of objects. All these steps must be completed at or near the site because artifacts cannot be taken home. Archaeologists require the ability to digitally register an object's orientation, detailed shape and other physical characteristics quickly and either on site, or a posteriori, when the data acquisition method allows it.

The database for the Great Temple excavation contains already more than 115 000 artifacts, recorded since 1993 [30]. Unfortunately, the full potential of archaeological databases is rarely realized. Most archaeologists are not able to analyze the geometric characteristics of artifacts and their spatial relationships with other elements of the site [15].

Our methodology encapsulates all of the above recording steps in one process. For example, 3D objects can be registered in the field via photogrammetric means [35, 36, 48, 58]. Our proposed technology will also permit archaeologists to do reconstructions of broken or eroded fragments. Once 3D information is gained about artifacts and architectural fragments while objects are being initially registered, it will be possible to better exploit reconstruction possibilities. A series of pot fragments (Fig. 4 & 10) can be interactively, and eventually automatically, reconstructed, eroded sculpture reconditioned to understand the original features and surface, a wall rebuilt without having to lift heavy fragments, and many elephant-head column capital trunks reconsolidated (Fig. 2). In many cases, archaeological artifacts go uncited as historically significant because they cannot be interpreted and referenced with other like examples. Our proposed technology allows archaeologists to understand and reference objects within a historic framework and also permits visualization that has, in the past, been unavailable or too costly.[2]

3 Early Results

[Figure 5 about here.]

3.1 Gestural and Verbal User Interfaces

User interfaces developed for production environments have not evolved significantly since the introduction of the windows, icons, menus, and point-and-click (WIMP) interface metaphor over two decades ago. Despite the advantages of WIMP interfaces (e.g., ease of use, short learning curve, and wide applicability), they greatly under-utilize the real-world capabilities and skills of users by limiting input and output to a keyboard, mouse, and monoscopic display. While effective for 2D desktop productivity applications, WIMP interfaces are not the ideal solution for intrinsically 3D applications. In this context, we have undertaken the study and evaluation of the next generation of post-WIMP interfaces that leverage application-specific knowledge and human skills to realize a more powerful, natural, and task-efficient user interface (e.g. see Fig 5).

Examples of our recent work are 3D widgets [14, 3], free-form deformations [24, 42] and gestural interfaces [60, 62, 61, 13]. 3D widgets demonstrate how parameters can effectively be represented by 3D geometry and embedded in a 3D dataspace. Our system *Sketch* [60] is a gestural interface for 3D geometric conceptual design which demonstrates that 2D drawn gestures can specify rich, context-sensitive commands to realize a powerful interface without relying on 2D WIMP user interface mechanisms.

Our most recent and on-going effort, the ARCHAVE[3] project, on the development of a multi-platform interactive virtual environment for archaeological analysis within the context of an accurate reconstruction of the site, both in space and time, is presented elsewhere [57].

3.2 3D Reconstruction from a Single Image

[Figure 6 about here.]

We have conducted preliminary work in order to extend our current *Sketch* system [60, 37, 55] to interactively generate and edit free-form 3D shape models in a sequence of images. Figure 6 illustrates our first generation system which makes use of single viewed perspective images [59] together with basic geometric primitives. Our approach in this stream is to maintain a functional system that is fully interactive using, in the early phases, our current knowledge of 3D shape and scene recovery, and incorporates novel shape models and automated shape recovery algorithms, as they become available in the later phases of this project.

3.3 3D Reconstruction from Multiple Images

[Figure 7 about here.]

In order to establish the geometry of a scene and its objects, a number of correspondences (e.g. feature points) need to be recovered between N images of a sequence [35, 16, 48]. In the context of archaeological scenery, corner detectors combined with a model-based approach (for position refinement), prove useful [7, 35]. For a video sequence, one can take advantage of the "continuity" of the sequence, by using robust tracking techniques [35, 55]. The correspondence problem is harder to solve for a set of photographic snapshots taken from a-priori unkown positions. This stage is user-driven in classical photogrammetry [36, 58]. To automate this task, one can make use of classical (window) correlation-based techniques combined with relaxation methods in an optimization stage. Given the intrinsic camera parameters, we can then establish the calibration of the sequence (also called "exterior orientation" in photogrammetry). Alternatively, one can recover the full set of parameters, via robust estimation techniques [17]. Once calibration is solved, more feature points can be acquired and matched to generate a cloud of 3D points. Finally, a triangulation can be obtained thanks to methods retrieving the connectivity (topology) of the bounding surface. Such methods permit to obtain realistic renditions of a statue (e.g. at roughly a ±5 mm accuracy in surface deviation). We expect our shape models, to be presented below, to constrain the reconstruction process and greatly simplify the final representation of such free-form objects while maintaining good accuracy.[4]

[Figure 8 about here.]

Similar techniques were applied to photographs taken at Petra and are illustrated in Figure 7 where we performed some detailed wall reconstruction under user supervision [58]. We have also experimented with a "dual" method to photogrammetry, where the camera and object positions are fixed, and, instead, the light source is moved to known positions. This is based on the work of Belhumeur *et al.* [23, 4, 5]. Note that, this technique bypasses the problem of calibration. However, such a setup provides for excellent accuracy to scan small objects in a constrained environment where lighting conditions can be controlled (see Fig. 8) and is, thus, comparable to structured light techniques such as used in [48].

3.4 Fragment Representation and Reassembly

The series of detailed head statuary in Figure 3 need to be reconstructed by filling-in missing sections, fusing related fragments, reconditioning eroded surfaces and, finally, comparing the shapes of the different heads with others found in the region of Petra [29]. Figure 9 illustrates a similar problem we have solved using IPS models (see § 4.2) by matching 3D fragments of an Egyptian bust [6]. The use of ridges and skeletal graphs for the same purposes represent on-going work, and more details about these methods are given in § 4.1 and § 4.3 below.

[Figure 9 about here.]

3.5 Site Content Discovery via 3D Geometric History

[Figure 10 about here.]

For analysis, it is essential to maintain the artifacts in their architectural and topographical context. Following what Forte proposed [18], we have started exploring how Geographical Information System (GIS) [32, 34] and Virtual Environments (VE) can be useful in helping archaeologists understand their data to develop new conclusions and hypotheses about the history and evolution of the Nabataean culture [57].

Figure 10 shows a GIS application with a 3D view of trenches. Color represents the concentration of pottery fragments found in each locus or layer of excavated material. Unfortunately, the "traditional" GIS cannot represent in 3D the in situ or find position of individual artifacts or allow reference to specific finds on site or in other sites around or outside of Petra. With the ability to reference the location and geometry of artifacts, archaeologists will have a more dynamic data set that can be used to reconstruct, link objects for analysis and maintain spatial information for future generations. This is explored in our ARCHAVE project which is described in detail elsewhere [57].

4 Three Dimensional Free-Form Shape Modeling

We have been investigating the use of 3D distinct representations for shape, *i.e.*, ridges, implicit polynomials and skeletal graphs. Our premise is that these representations are intimately connected (e.g. see Fig. 15.(b)), and we propose a joint, integrated, and comprehensive investigation of these, which shall lay the foundations to establish a complete and formal shape language for general use in archaeology and beyond. Briefly, *ridges* are a representative of curve loci on a surface, a one dimensional construct in space, e.g., a break curve where a sherd was broken from another piece. *Implicit polynomial surfaces* (IPS) are a representative of entire surface loci, a two dimensional construct in space, e.g., the outer and inner surfaces of a pottery sherd. *Skeletal graphs*, also called *Medial Axes* [9], are a representative of volumetric features, a three dimensional construct in space, e.g., the main axis of a pot and its symmetric relations with the pot surfaces. These three elements are constitutive of a vocabulary classification for shape. Their relationship via a hyper-graph structure, will define the equivalent of a syntax for 3D shape. In the remaining of this final section we detail each vocabulary class.[5]

4.1 Curvilinear Modeling through Ridges

What are ridges? It is simplest to define them in 3D by analogy with a 2D case. The boundary of a 2D shape is a simple closed curve which can be divided into convex and concave portions, separated by points of inflection where the curvature of the boundary vanishes. In addition, there are special points on the boundary where the curvature has a local maximum or minimum. The most important of these are the "vertices": local maxima in convex segments, and local minima in concave segments, which are analogs of the vertices of polygons. In particular, each endpoint of the medial axis or skeletal graph (see § 4.3) of the shape is the center of the osculating circle at a convex vertex [41] (see Fig. 14.(a) and (d)). The psychologist Attneave proposed that these were the most perceptually salient and informative points on the contour [2].

[Figure 11 about here.]

What happens in 3D? The situation is more complex. Instead of merely convex and concave pieces, the boundary of any 3D shape is divided into three kinds of pieces: (i) the convex parts with both principal curvatures positive, (ii) the parts where both principal curvatures are negative, *i.e.*, the surface is strictly concave, and (iii) the hyperbolic saddle-like parts where one principal curvature is positive, the other negative. Instead of local max and min points for the principal curvatures, one looks for curvilinear collections of points where the larger of the two principal curvatures has a local max on its corresponding *line* of curvature (Fig. 11), and points where the smaller curvature has a local minimum. The *ridges* in the *convex* parts of the surface are smooth analogs of the convex edges of a polyhedron and are perceptually salient as the prominent lines where the surface *protrudes*. Likewise, "ridges" in the *concave* parts of the surfaces look like the bottom of *valleys* where the surface is *creased* [12, 10, 44].

[Figure 12 about here.]

One goal is to use these features to describe 3D shape in an intuitive way. In Fig. 12.(a), ridge computations on a sherd surface data obtained via CT scanning (cf. Fig. 4) are depicted: "warm" colors (red and yellow) indicate positive extremal curvature, while "cooler" (*i.e.*, blue) colors indicate negative extremal curvature. We have developed an interactive algorithm to extract ridges and valleys based on this curvature map; this is illustrated in Fig. 12.(b). The user clicks a starting point and goal (which may be identical, to close a loop), decides whether a ridge or valley is needed, and then lets the computer rely upon an implementation of a 3D active contour to seek an optimal path [39, 43, 1]. Such an active contour model tends to minimize a cost function based on an integral of the curvature measures along a path as well as on measures of elastic tension along the contour. Because such features as ridges and valleys correspond well with (human) intuitive curvilinear descriptors for free-form shapes, we believe they will provide a very effective tool for manipulating shape for interactive modeling as well as for indexing and searching databases of shapes, and delimiting break surfaces of sherds.

The next stage in our research program is to explore the use of ridges/valleys on a variety of free-form shapes, bodies and a range of artifacts as well as faces, animals, humans, sculptures of various types, furniture and tools, etc. There has been psychophysics on the human perception of ridges [47] and an additional goal is to characterize how *stable* ridges are for shape modeling.

4.2 Surface Modeling through Implicit Polynomials

Multivariate Implicit Polynomials provide a powerful and rich representation for 2D and 3D curves and surfaces [8, 33, 54]. For example, a trivariate d^{th} degree Implicit Polynomial Surface (IPS) is the zero set of a d^{th} degree explicit polynomial, *i.e.*, the set of points (x, y, z) where the explicit polynomial is zero, $f(x, y, z) = \sum_{i+j+k \leq d} c_{ijk} x^i j^j z^k = 0$. These surfaces are generalizations, to more complicated shapes, of the conics, *e.g.*, a hyper-ellipsoid, a cylinder with hyperbolic cross section, etc.

[Figure 13 about here.]

For example, the set of points (x, y, z) for which $(x - x_0)^2 + (y - y_0)^2 + (z - z_0)^2 - R^2 = 0$ is the equation of a sphere of radius R having center $x = x_0$, $y = y_0$, $z = z_0$. These IPS are useful for representing blobby closed surfaces, open patches as in Figures 3, 7, 9, 10 and 13, surfaces attached to prominent ridges, generalized cylinders [46, 50], and other shapes, e.g., free-form shapes with holes. IPS can be used in at least two interesting ways: (i) as a coarse, but smooth, approximation or (ii) as a close fit to the data. As a low resolution approximation to a complex surface, an IPS can be used to extract coarse geometry useful for shape recognition, crude assembly of fragments into reconstructions, etc. On the other hand, a single high degree IPS or a number of patches made of IPS of more modest degree, can be used for a high resolution representation. These uses are illustrated in Figures 13 and 9, respectively.

A goal of our research program is to explore the use of ridges (§ 4.1) and skeletal graphs (§ 4.3) for the optimal placement of IPS patches so that low order fits, and thus fewer parameters, can be used. Fitting to data is fast, repeatable, and robust, since the fitting is linear least squares (thus resulting in an *explicit* expression for the estimated coefficient vector), it is regularized by our $3L$ fitting [6], and is further regularized by the use of *ridge regression* [6]. Note that, the principal computational cost is in computing only a scatter matrix of monomials based on the (x, y, z) data points. Once this is done, a refitting to subsets or unions of data point sets, or a modification of surfaces through human interaction, requires orders of magnitude less computation and is possible in real-time.

Our approach to human interaction with shape, when using IPS, is to modify the surface much as a sculptor or a designer might: by specifying a position, or a position and a tangent cut, or a position, a tangent cut and two bendings (e.g. via principal curvatures) that we want the deformed surface to satisfy approximately (soft constraint) or exactly (hard constraint), such that the surface is not modified much away from the position of interest. More generally, we can specify a number of points, or a curve in 3D, or a surface attached to a ridge that we want the deformed surface to approximate.

Our next step will be to investigate a hybrid model by interpolating with an IPS exactly (hard constraint) or approximately (soft constraint) by specifying some surface properties (e.g. curvatures, tangents, etc.) in-between the ridges, where these latter properties could be specified through stochastic processes or through probability distributions. For elongated surfaces like an arm, perhaps an upper torso, an elephant trunk, etc., a *generalized cylinder* [46, 50] can be realized by computing the skeletal axis of skeletal sheets (see below), and then sweeping a cross-sectional planar IP curve along the axis, where the plane is orthogonal to the axis and has a local coordinate system determined by the skeletal sheet.

4.3 Volumetric Modeling through Skeletal graphs

[Figure 14 about here.]

The skeletal graph of a 2D shape, or its medial axis, is the locus of centers of *maximally inscribed circles* (see Fig. 14). It is an intuitive and efficient representation for the recognition of 2D shapes, since variations in shape often leave the graph structure intact [9]. This medial axis graph structure, however, maps to a variety of shapes (*i.e.*, by varying its associated radius function) and thus is not sufficiently constrained to reveal qualitative shape. A dynamic view of the skeletal graph as the singularities (shocks) of wavefronts propagated from the initial boundary [38] defines a notion of velocity and direction of flow for each medial axis point and thus leads to a finer partitioning of the skeletal branches at points where the flow is reversed. The resulting shock graph, when stripped of radius information, is more discriminating in that it reveals qualitative shape [31].

Shape can be fully reconstructed from the medial axis and the corresponding radii as the envelope of circles centered on the axis. The local nature of this intimate connection between shape and skeletal graphs, however, is not explicit in the envelope reconstruction. In [19] the differential geometry of the boundary, *i.e.*, tangent and curvature, is derived as a function of the differential geometry of the medial axis and of the dynamics of shock propagation on the axis, *i.e.*, velocity and acceleration. It is shown that the shock graph, together with curvature and acceleration descriptions for each link, is a complete description of shape.

[Figure 15 about here.]

For 3D shape, the medial axis is the locus of *maximal bitangent spheres*. The wave propagation approach again leads to a dynamic view of shocks propagating on the skeletal locus [40]. The points of medial axis (and shock set) have been classified resulting in a hypergraph representation [20, 21] consisting of skeletal sheets with associated flow fields, which end either at a boundary corresponding to *ridges* or at curves shared by three medial sheets, much like the central axis of a *generalized cylinder* with a triangular base [46]. These curves interact only at special points, namely when they intersect each other at nodes. These are the only generic possibilities (Fig. 15). The skeletal hypergraph describes the connectivity among symmetries of each portion of the shape. An example of the computation of the trace of such skeletal loci is given in Figure 16 for a typical pottery sherd.

[Figure 16 about here.]

The next step in this research is to investigate how the skeletal hypergraph can be matched against other medial axis representations in a pre-stored database of similar objects, in analogy to 2D matches [49]. Also, partial matches of skeletal axial curves should prove useful to solve the difficult problem of automatically stitching together different sherds to recover a full pot [56].

5 Conclusion

The SHAPE Lab. has been created with the goals of: (i) introducing new geometric modeling and 3D surface and structure recovery from images; (ii) improving human/machine interaction tools for facilitating human input of geometric information to the machine and then visualizing the results in real time; (iii) developing new tools to facilitate reconstructing large geometric structures (e.g., walls of buildings) and smaller objects (e.g., columns and their capitals, and at more detailed levels, with statues and artifacts) from free-form fragments scattered about a site. These objectives require considerable domain-specific knowledge and are central in providing material for analysis in archaeology but also can be used extensively in architecture and architectural history, and ultimately in many other disciplines where the design and manipulation of free-form 3D shapes is required.

In order to fulfill this ambitious program, a key component is the development of a shape language for 3D free-form objects. We have reported in this paper on our early success in putting together a vocabulary based on three classes of elements: *ridges*, to model perceptually significant surface curves, *implicit polynomials*, to model surfaces of various complexity, and *skeletal graphs* to model volumetric features and, furthermore, provide the "glue" to relate together the three classes.

Difficult and interesting challenges still remain ahead of us. There is clearly a *continuum* from the ridges on polyhedra which are most precise as well as most salient and those in near planar or near spherical parts of the surface, and this "scale-space" for ridges needs to be studied [44]. A second question is how to *approximate* a 3D shape using ridge and skeletal data. In the plane, an old idea going back to Attneave is to approximate any 2D shape by the polygon joining its vertices. What analogs of this construction can we make in 3D? An essential step in the HCI part of this research (§ 3.1), is to be able to estimate an entire shape roughly based on the user marking approximate ridges and local planes of symmetry, and then let the computer position and select implicit polynomial models of the surface patches bounded by such ridges. Another question concerns the location of ridges using reflectance data gathered from one or more images of an object. The basic idea is that since the tangent plane is changing rapidly at ridge points, images of the surface will have rapid changes in intensity along ridges. In addition, specularities "cling" to ridges and with elongated light sources, may even make the whole ridge shine. We want to make these ideas precise and integrate them in the reconstruction of 3D shape from multiple images with varying illumination as a constraint used in the recovery of shape (Fig. 8).

In addition to developing this approach to shape representation, we want to apply it to object recognition based on shape. It is broadly recognized that one of the most effective techniques for object recognition is the use of Bayesian statistical methods [11, 53, 52, 45]. In order to apply this method to free-form shapes, we need priors of the space of such shapes [45]. For example, the shapes we find are often built out of parts which may be generalized cylinders or rectangular parallelepipeds; or they may have limbs like a statue, a human or a tree in winter, etc. The approach we want to take is to model stochastically the generic features of shapes, their skeletal graphs and ridges and decomposition into parts. In 2D, Zhu and Yuille [63] have constructed stochastic models of shapes based on the medial axis. In 3D the development of such priors, involving the explicit representation of ridges and skeletal graphs, is needed.

Notes

[1] SHAPE: SHape, Archaeology, Photogrammetry, Entropy; a multi-disciplinary project established in the Fall of 1999; visit our website at: `www.lems.brown.edu/vision/extra/SHAPE/`.

[2] N.B., we also plan to relate and compare objects and aspects of our site with other sites within Petra and other Nabataean sites like Medain Saleh.

[3] ARCHAVE: ARCHAeology with Virtual Environment systems; see [57].

[4] Such a model-based constraint paradigm is similar to Debevec *et al.*'s, earlier DARPA community's and others' approach to 3D reconstruction from images. However, in these other approaches, much simpler set of models are used. These simpler models are typically made of regular primitives to represent simple architectural shapes [51, 22, 16].

[5] The syntaxic properties of our shape language will be reported elsewhere; see [19, 20, 21] for early theoretical investigations.

Acknowledgments

Major funding for this project comes from the USA NSF/KDI grant #BCS-9980091. Frederic Leymarie was partially supported by a PhD fellowship from IBM. Besides the co-authors, a number of students and researchers have been involved in the first year of the SHAPE Lab and have had a major impact on the work reported in this paper. These include graduate students Daniel Acevedo, Stuart Andrews and Jill Baker, Dr. Peter Giblin (Liverpool) and Dr. Jean-Philippe Tarel (Paris).

References

[1] Stuart Andrews. Interactive generation of feature curves on surfaces. Master's thesis, University of Toronto, Ontario, Canada, January 2000.

[2] F. Attneave. Some informational aspects of visual perception. *Psychological Review*, 61(3):183–193, 1954.

[3] M.R. Ayers and Zeleznik R.C. The lego interface toolkit. In *Proceedings of the ACM Symposium on User Interface and Software Technology (UIST'96)*, pages 97–98. ACM, 1996.

[4] P. Belhumeur and D. Kriegman. What is the set of images of an object under all possible illumination conditions? In *Proceedings of CVPR*, pages 270–277, San Francisco, CA, June 1996. IEEE CS Press.

[5] P. Belhumeur, D. Kriegman, and A. Yuille. The bas-relief ambiguity. In *Proceedings of CVPR*, pages 1060–1066, San Juan, PR, June 1997. IEEE CS Press.

[6] M. M. Blane, Z. Lei, and D. B. Cooper. The 3L algorithm for fitting implicit polynomial curves and surfaces to data. *IEEE Trans. on PAMI*, 22(3):298–313, March 2000.

[7] Thierry Blaszka and Rachid Deriche. Recovering and characterizing image features using an efficient model based approach. Research Report RR-2422, INRIA, Sophia Antipolis, France, November 1994. 59 pages.

[8] J. Bloomenthal, C. Bajaj, J. Blinn, M.-P. Cani-Gascauel, B. Wyvill, and G. Wyvill. *Introduction to Implicit Surfaces.* Morgan Kaufmann Pub., San Francisco, CA, USA, 1997.

[9] H. Blum. Biological shape and visual science. *Journal of Theoretical Biology*, 38:205–287, 1973.

[10] J.W. Bruce, P.J. Giblin, and F. Tari. Ridges, crests and sub-parabolic lines of evolving surfaces. *International Journal of Computer Vision*, 19:195–210, 1996.

[11] B. Cernuschi-Frias, D. B. Cooper, and Y. P. Hung. Toward a model-based Bayesian theory for estimating and recognizing parametrized 3-D objects using two or more images taken from different positions. *IEEE Trans. on PAMI*, 11(10):1028–1052, 1989.

[12] R. Cipolla, G. Fletcher, and P.J. Giblin. Surface geometry from cusps of apparent contours. In *Proc. Int. Conf. on Computer Vision (ICCV'95)*, pages 858–863. IEEE Computer Society Press, June 1995.

[13] J.M. Cohen, L. Markosian, R.C. Zeleznik, J.H. Hughes, and R. Barzel. An interface for sketching 3D curves. In *Proceedings of 1999 Symposium on Interactive 3D Graphics*, Atlanta, April 1999. ACM.

[14] D. Brook Conner, Scott S. Snibbe, Kenneth P. Herndon, Daniel C. Robbins, Robert C. Zeleznik, and Andries van Dam. Three-dimensional widgets. *Computer Graphics (1992 Symposium on Interactive 3D Graphics)*, 25(2):183–188, March 1992.

[15] M. Crescioli and F. Niccolucci. P.E.T.R.A.-Data: An integrated environment for archaeological data processing. In *New Techniques for Old Times: Computer Applications and Quantitative Methods in Archaeology. Proc. of the 26th CAA Conf.*, number 757 in BAR International Series, pages 133–134, Barcelona, Spain, March 1998.

[16] Paul E. Debevec, Camillo J. Taylor, and Jitendra Malik. Modeling and rendering architecture from photographs. In *Computer Graphics Proceedings, SIGGRAPH'96.* ACM, August 1996.

[17] O. Faugeras et al. 3D reconstruction of urban scenes from sequences of images. Research Report RR-2572, INRIA, Sophia Antipolis, France, June 1995.

[18] Maurizio Forte. About virtual archaeology: Disorders, cognitive interactions and virtuality. In J.A.Barcelo, M.Forte, and D.H.Sanders, editors, *Virtual Reality in Archaeology*, number 843 in BAR International Series, pages 247–259, Oxford, UK, 2000. Archaeopress.

[19] P. J. Giblin and B. B. Kimia. On the intrinsic reconstruction of shape from its symmetries. In *Proc. of CVPR*, pages 79–84, Colorado, June 1999.

[20] P. J. Giblin and B. B. Kimia. On the local form and transitions of symmetry sets, and medial axes, and shocks in 2D. In *Proc. of ICCV*, pages 385–391, Greece, September 1999.

[21] P. J. Giblin and B. B. Kimia. On the local form of symmetry sets, and medial axes, and shocks in 3D. In *Proc. of CVPR*, pages –. IEEE Computer Society, 2000.

[22] G.Schmitt et al. Toward virtual reality in architecture: Concepts and scenarios from the architectural space laboratory. *Presence*, 4(3):267–285, Summer 1995.

[23] G. Hager and P. Belhumeur. Real time tracking of image regions with changes in geometry and illumination. In *Proceedings of CVPR*, San Francisco, CA, June 1996. IEEE CS Press.

[24] W. Hsu and J.F. Hughes. Direct manipulation of free-form deformations. In *Computer Graphics Proceedings, SIGGRAPH'92*, pages 177–184. ACM, July 1992.

[25] Martha C. S. Joukowsky. *A Complete Manual of Field Archaeology: Tools and Techniques of Field Work for Archaeologists.* Prentice-Hall, Englewood-Cliffs, New Jersey, 1980. (Eight printing).

[26] Martha C. S. Joukowsky. *Early Turkey: An Introduction to the Archaeology of Anatolia from Prehistory through the Lydian Period.* Kundall-Hunt, 1996.

[27] Martha C. S. Joukowsky. Brown University 1997 excavations at the Petra Great Temple. *Annal of the Department of Antiquities of Jordan*, XLII:293–318, 1997.

[28] Martha C. S. Joukowsky. The Petra southern temple: The fourth season, 1996. *American Journal of Archaeology*, 101:339, 1997.

[29] Martha C. S. Joukowsky. The Great Temple at Petra. *American Journal of Archaeology*, 102(3):593–596, July 1998.

[30] Martha Sharp Joukowsky. *Petra Great Temple - Volume 1: Brown University Excavations 1993-1997*. Petra Exploration Fund, Providence, RI, USA, 1999.

[31] Benjamin B. Kimia, Allen R. Tannenbaum, and Steven W. Zucker. Shapes, shocks, and deformations, I: The components of shape and the reaction-diffusion space. *International Journal of Computer Vision*, 15:189–224, 1995.

[32] M. Kofler, H. Rehatschek, and M. Gruber. A database for a 3D GIS for urban environments supporting photo-realistic visualisation. In *International Archives of Photogrammetry and Remote Sensing (XVIII ISPRS Congress)*, volume XXXI, pages 198–202, Vienna, Austria, July 1996.

[33] Z. Lei, T. Tasdizen, and D.B. Cooper. PIMs and invariant parts for shape recognition. In *Proc. of International Conference on Computer Vision (ICCV'98)*, pages 827–832, Bombay, India, January 1998. IEEE CS Press.

[34] F. Leymarie. Exploitation of 3D georeferenced datasets in a GIS. In *Proc. International Workshop on "High Performance Computing Networking Exploitation of Multimedia Databases*, number EUR 17349 EN, pages 10–29, Ispra, Italy, March 1997. European Commission.

[35] F. Leymarie et al. REALISE: Reconstruction of reality from image sequences. In P.Delogne, editor, *Proc. of IEEE International Conference on Image Processing (ICIP'96)*, volume III, pages 651–654, Lausanne, Switzerland, September 1996.

[36] F. Leymarie and M. Gruber. Applications of virtual reality: CyberMonument and CyberCity. In *Proc. of the* 3^{rd} *European Digital Cities Conference*, Berlin, Germany, December 1997. European Commission.

[37] F. Leymarie, B.B. Kimia, and M. Gruber. Direct camera pose evaluation from a single view and a map. Technical Report LEMS-174, Brown University, Eng. Div., Providence, RI, USA, August 1998.

[38] F. Leymarie and M. D. Levine. Simulating the grassfire transform using an active contour model. *IEEE Trans. on PAMI*, 14(1):56–75, January 1992.

[39] F. Leymarie and M. D. Levine. Tracking deformable objects in the plane using an active contour model. *IEEE Trans. on PAMI*, 15(6):617–634, June 1993.

[40] F. F. Leymarie and B. B. Kimia. Discrete 3D wave propagation for computing morphological operations from surface patches and unorganized points. In J. Goutsias, L. Vincent, and D. Bloomberg, editors, *Math. Morphology and its Applications to Image and Signal Processing*, volume 18 of *Computational Imaging and Vision*, pages 351–360, Palo Alto, CA, USA, June 2000. Kluwer Academic.

[41] M. Leyton. *Symmetry, Causality, Mind.* MIT Press, Cambridge, MA, 1992.

[42] Lee Markosian, Jonathan Cohen, Thomas Crulli, and John Hughes. Skin: A constructive approach to modeling free-form shapes. In *Proceedings of SIGGRAPH'99*, Los Angeles, CA, August 1999. ACM.

[43] Eric N. Mortensen and William A. Barrett. Interactive segmentation with intelligent scissors. *Graphical Models and Image Processing*, 60(5):349–384, September 1998.

[44] D. Mumford, P.Giblin, G.Gordon, P.Hallinan, and A.Yuille. *Two and Three dimensional Patterns in Faces.* A.K. Peters, 1999.

[45] David Mumford. *The Statistical Description of Visual Signals.* Akademie Verlag, 1996.

[46] Ambjorn Naeve and Jan-Olaf Eklundh. Generalized cylinders - what are they? In C. Arcelli, L.P. Cordella, and G. S. di Baja, editors, *Aspects of Visual Form Processing*, 2nd International Workshop on Visual Form, pages 384–409, Capri, Italy, May-June 1994. World Scientific.

[47] Flip Phillips. *Geometric structure, frames of reference, and their implication in the localization of features on smoothly curved surfaces.* PhD thesis, Ohio State University, 1997.

[48] M. Pollefeys, M. Proesmans, R. Koch, M. Vergauwen, and L. Van Gool. Acquisition of detailed models for virtual reality. In J.A.Barcelo, M.Forte, and D.H.Sanders, editors, *Virtual Reality in Archaeology*, number 843 in BAR International Series, pages 71–77, Oxford, UK, 2000. Archaeopress.

[49] Daniel Sharvit, Jacky Chan, Hüseyin Tek, and Benjamin B. Kimia. Symmetry-based indexing of image databases. *Journal of Visual Communication and Image Representation*, 9(4):366–380, December 1998.

[50] C. Steckner. Form and fabric, the real and the virtual. In J.A.Barcelo, M.Forte, and D.H.Sanders, editors, *Virtual Reality in Archaeology*, number 843 in BAR International Series, pages 121–128, Oxford, UK, 2000. Archaeopress.

[51] A. Streilhen. Towards automation in architectural photogrammetry: CAD-based 3D-feature extraction. *ISPR Journal of Photogrammetry and Remote Sensing*, 49(5):4–15, 1994.

[52] J. Subrahmonia, D. B. Cooper, and D. Keren. Practical reliable Bayesian recognition of 2D and 3D objects using implicit polynomials and algebraic invariants. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 18(5):505–519, May 1996.

[53] J. Subrahmonia, D. Keren, and D. B. Cooper. An algebraic and Bayesian technology for object recognition. In K.V. Mardia, editor, *Advances in Applied Statistics*, pages 165–205. Carfax Publishing Co., Abingdon, Oxfordshire, UK, 1994.

[54] J.-P. Tarel, H. Civi, and D. B. Cooper. Pose estimation of free-form 3D objects without point matching using algebraic surface models. In *Workshop on 3D Modeling, ICCV98, Bombay, India*, pages 13–21, January 1998.

[55] J.-P. Tarel, F. Guichard, and D. Aubert. Tracking occluded lane-markings for lateral vehicle guidance. In *Proc. of IEEE International Conference on Circuits, Systems, Communications and Computers (CSCC'99)*, Athens, Greece, July 1999.

[56] Gokturk Ucoluk and Ismail Hakki Toroslu. Automatic reconstruction of broken 3-D surface. *Computers and Graphics*, 23(4):573–582, August 1999.

[57] E. Vote, D. Acevedo Feliz, D. Laidlaw, and M. S. Joukowsky. ARCHAVE: A virtual environment for archaeological research. In *Proc. Int. Conf on Computer Applications in Archaeology (CAA'00)*, Lubljana, Slovenia, April 2000. Submitted.

[58] Eileen Vote. Using desktop photogrammetry to document archaeological remains: The Great Temple at Petra, Jordan. In G. Cooney and B. Masterson, editors, *Proc. Int. Conf on Computer Applications in Archaeology (CAA'99)*, University College, Dublin, Ireland, April 1999. To be published.

[59] James R. Williamson and Michael H. Brill. *Dimensional Analysis through Perspective - A Reference Manual.* Kendal/Hunt Publishing co., Iowa, USA, 1990.

[60] R. C. Zeleznik, K. Herndon, and J. F. Hughes. SKETCH: An interface for sketching 3D scenes. In *Computer Graphics Proceedings, SIGGRAPH'96*, Annual Conference Series, pages 163–170, New Orleans, Louisiana, August 1996. ACM.

[61] R.C. Zeleznik and A.S. Forsberg. Unicam - 2D gestural camera controls for 3D environments. In *Proceedings of 1999 Symposium on Interactive 3D Graphics*, Atlanta, April 1999. ACM.

[62] R.C. Zeleznik, A.S. Forsberg, and P.S. Strauss. Two pointer input for 3D interaction. In *Proceedings of 1997 Symposium on Interactive 3D Graphics*, pages 27–30, Providence, RI, April 1997. ACM.

[63] Song Chun Zhu and A. L. Yuille. FORMS: A flexible object recognition and modeling system. *International Journal of Computer Vision*, 20(3):187–212, December 1996.

List of Figures

Figure 1: An aerial view of the Great Temple at Petra, Jordan, and part of the floorplan.

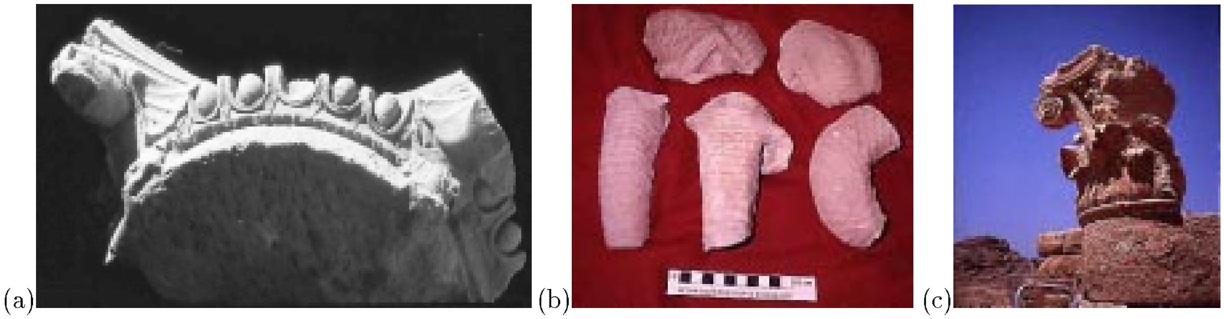


Figure 2: (a) Piece of column head with elephant-headed capitals (missing trunk on left). (b) Some of the hundreds of elephant trunk and head fragments found at Petra. (c) Complex column capital typical of the Nabataean style.

Figure 3: Typical fragments of head sculptures from Petra, to be analyzed. These fragments have considerable detail although some parts are missing and the surface is badly eroded.

Figure 4: Figure made from density measurements of a box of sherds. Densities within the box were imaged volumetrically using a hospital-based CT scanner. Iso-density surfaces were then created and rendered using marching cubes [43], a computer graphics technique.

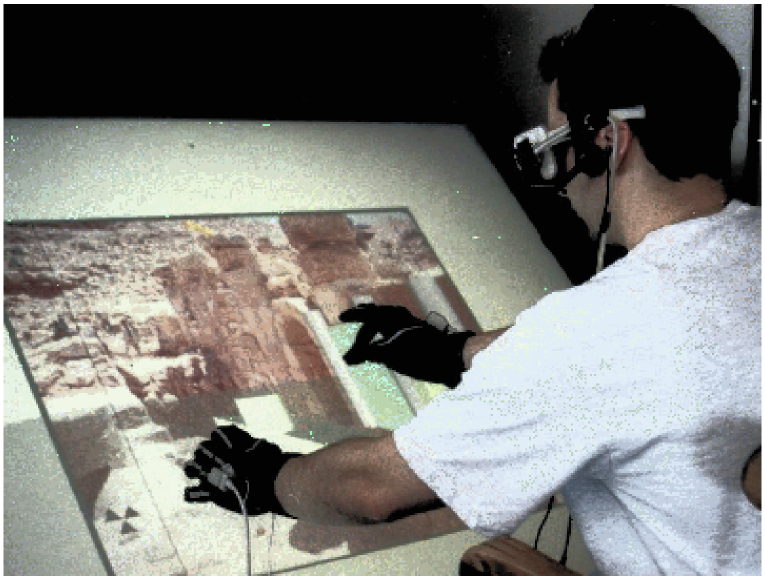


Figure 5: Illustration of a gestural and voice-driven user interface for interactively modeling 3D objects in a virtual environment. Simple geometric models here are overlaid on a photograph from the site at Petra. Stereo shutter glasses show each eye a different image, a microphone picks up speech, and gloves detect gestures and hand location, permitting natural 3D object manipulations. The large-sized screen is a BARCO Baron projection table providing a semi-immersive environment.

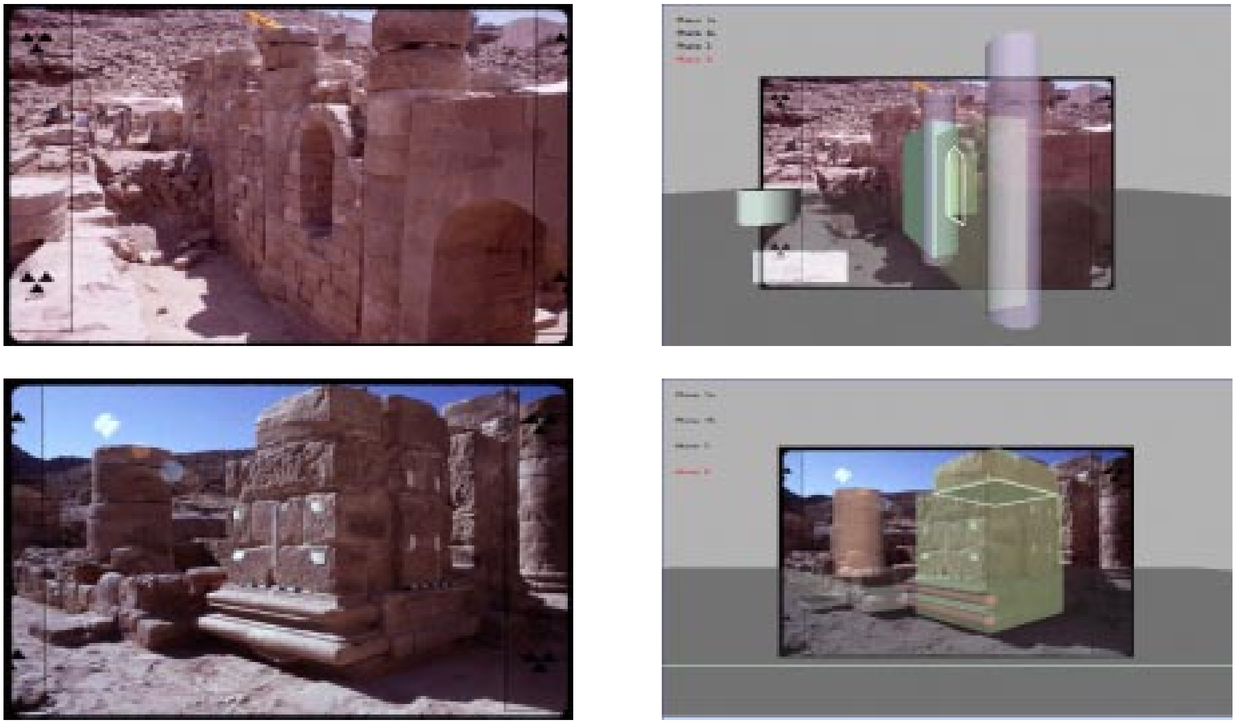


Figure 6: Examples of our perspective-based reconstruction using simple geometric primitives (images from Petra). On the left are shown single snapshots taken with an ordinary camera with the addition of fiducials. On the right are shown the results of drawing CAD-like primitives as overlays on the basis of perspective cues.

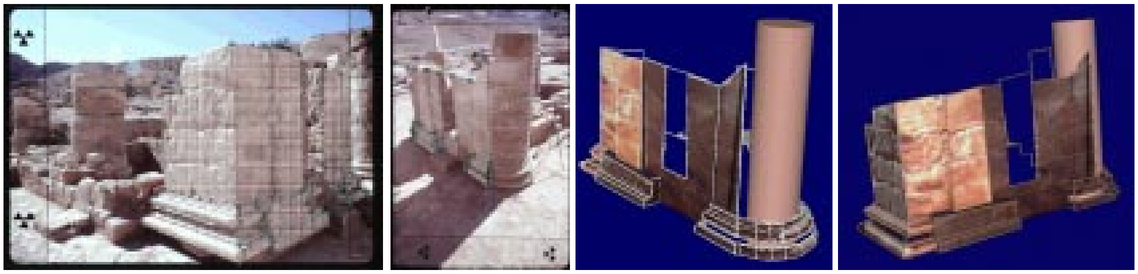
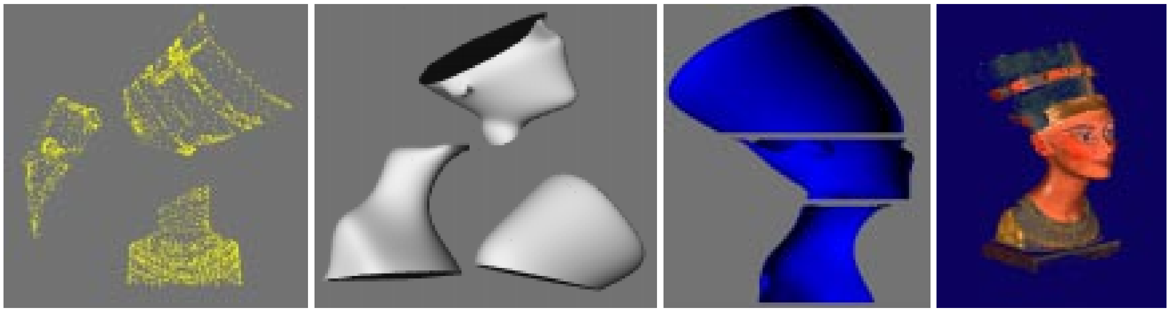


Figure 7: Illustration of photogrammetric 3D reconstruction from multiple images for architectural forms; more details in [58].

Figure 8: Six images (4 of which are shown) of the face of one of the co-author, D. Mumford, are used to determine its shape. The pose is constant, while the lighting varies. The resulting photometric variation reveals the shape, as shown in the depth map and wire-frame (5^{th} & 6^{th} image).



(a) 3D sample points (b) Fitted IPs. (c) Reassembly (d) Original bust

Figure 9: Bust of Nefertiti reassembly - (a) Three subsets of (x, y, z) data points, each constituting 3D surface measurements of one of the fragments into which a bust in (d) has broken. (b) Shaded images of algebraic surfaces (IPS) that have been fitted to each of the data subsets in (a). (c) Resulting IPS aligned to represent the unbroken bust. The 3D fragment of the face stops near the bottom of the eyes, the one of the hat includes top of eyes and upper part of one ear, and the one of the neck and shoulders includes lower surface of the chin. Degrees of IPS used were chosen arbitrarily and are single 4^{th} degree for hat, single 6^{th} degree for the large potion of face and head, and single 8^{th} degree for neck and shoulders. See [6] for more technical details.

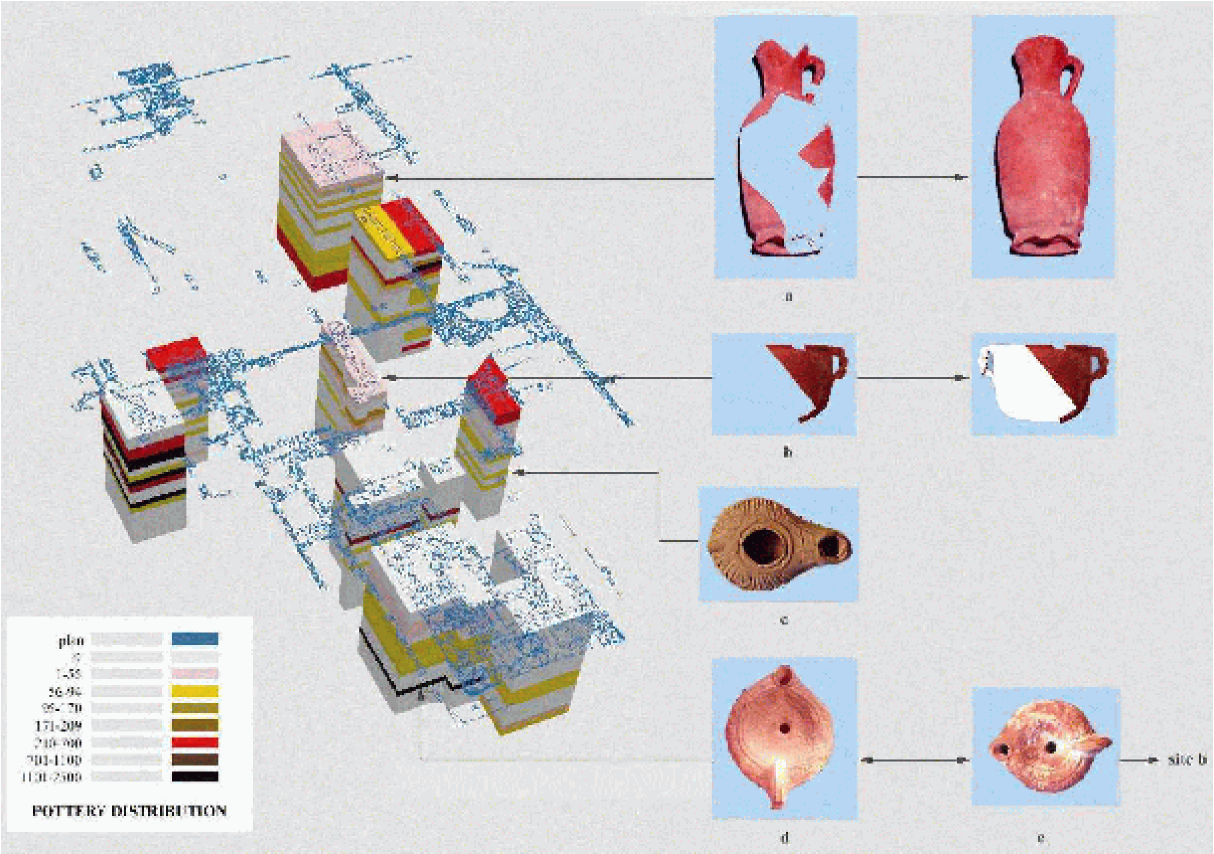


Figure 10: GIS application at Petra's Great Temple. Artifacts referenced and defined in 3D will allow archaeologists to reconstruct objects (*a* and *b*), do advanced spatial analysis, link artifacts between sites (For example, tracing lamps between sites in the region will allow us to trace trade routes.) (*c*, *d* and *e*), and maintain more comprehensive, accurate and accessible records of artifact geometry and find locations.

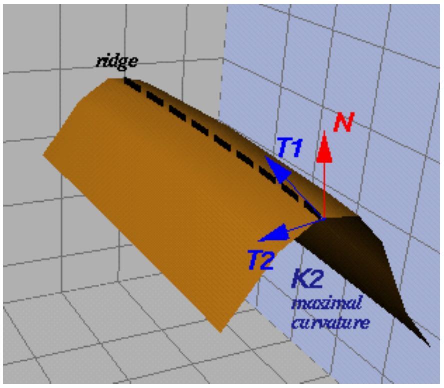


Figure 11: Concept of a ridge as a line of maximal principal curvature on a surface.

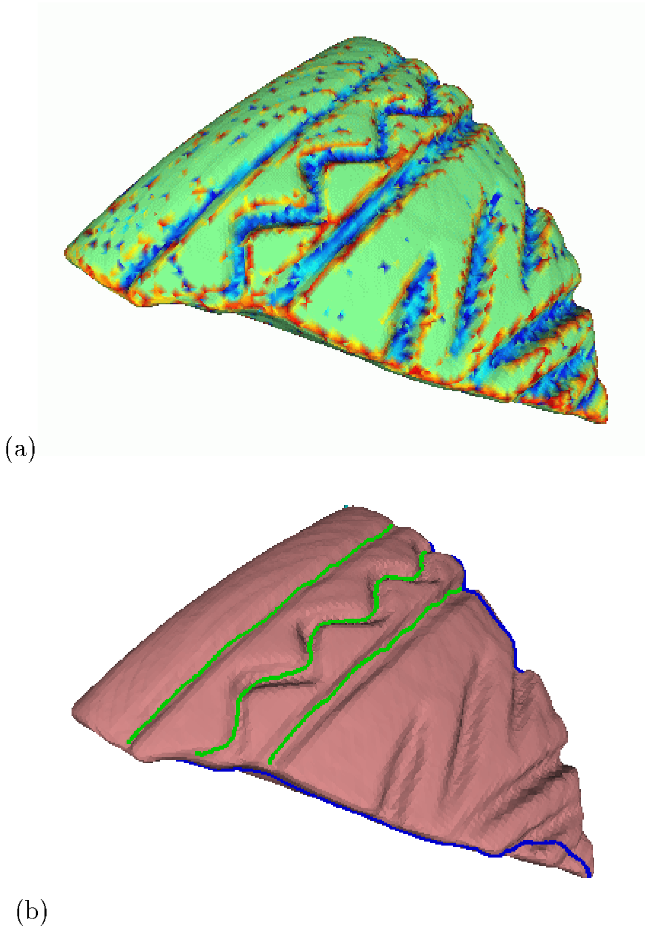


Figure 12: (a) Local curvature computations on a sherd from its recovered 3D surface. (b) Result of an interactive ridge and valley computation using a 3D active contour model (see [1] for more technical details).

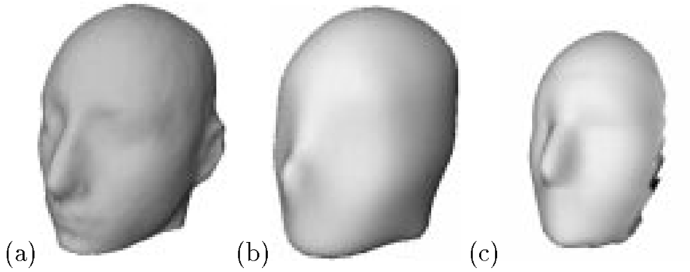
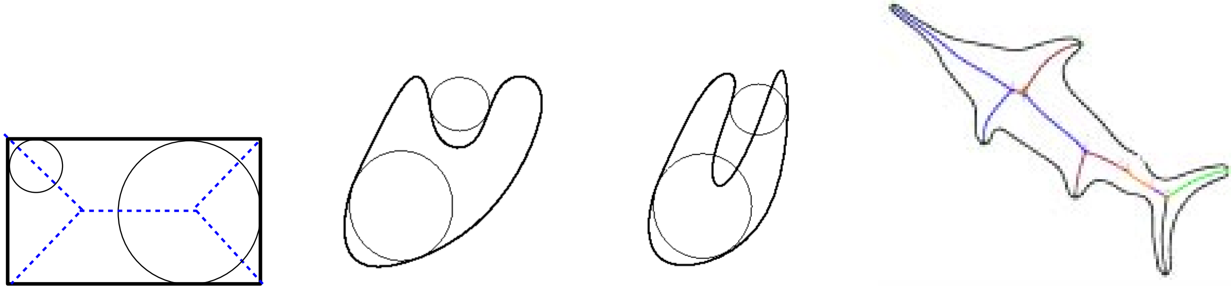
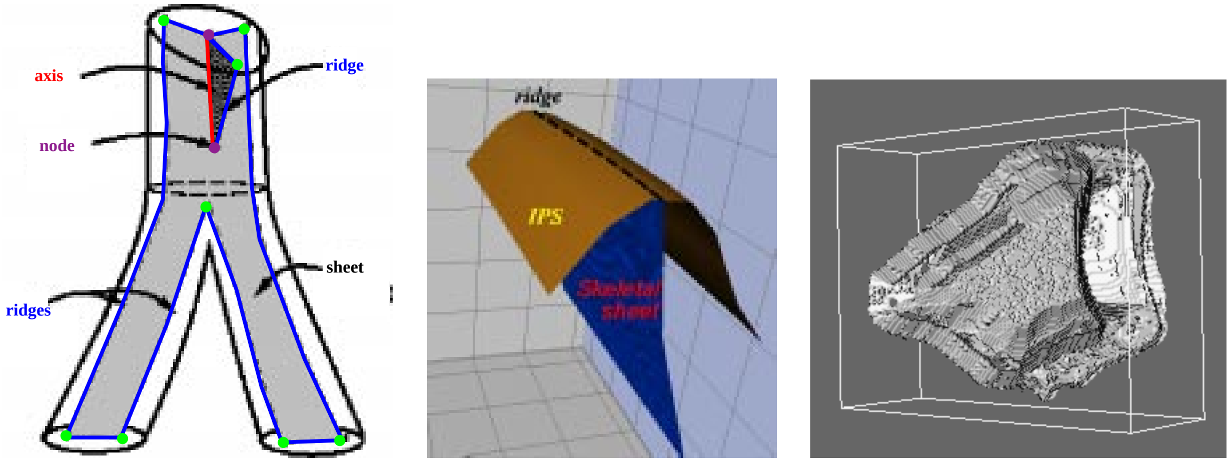


Figure 13: (a) Original triangulated head data set; (b) 10^{th} degree fit (ears discarded); (c) reconstruction via 12 patches, using 4^{th} degree IPS for each patch.



(a) (b) (c) (d)

Figure 14: 2D skeletons: (a) skeleton of a rectangle with two examples of maximally inscribed circles; (b) example of two bitangent circles which are maximal and (c) two which are not, since they cross the boundary; (d) more complex skeleton for a swordfish outline - note that at each curvature extrema of the boundary corresponds the end of a skeleton branch.



(a) (b) (c)

Figure 15: (a) Sketch of the 3D skeleton of a branching shape: the skeleton is made of sheets, axial and ridge curves, and nodes. (b) Concept of skeletal sheet and its relation to IPS and ridges. (c) Computed skeletal sheets of a carpal bone.

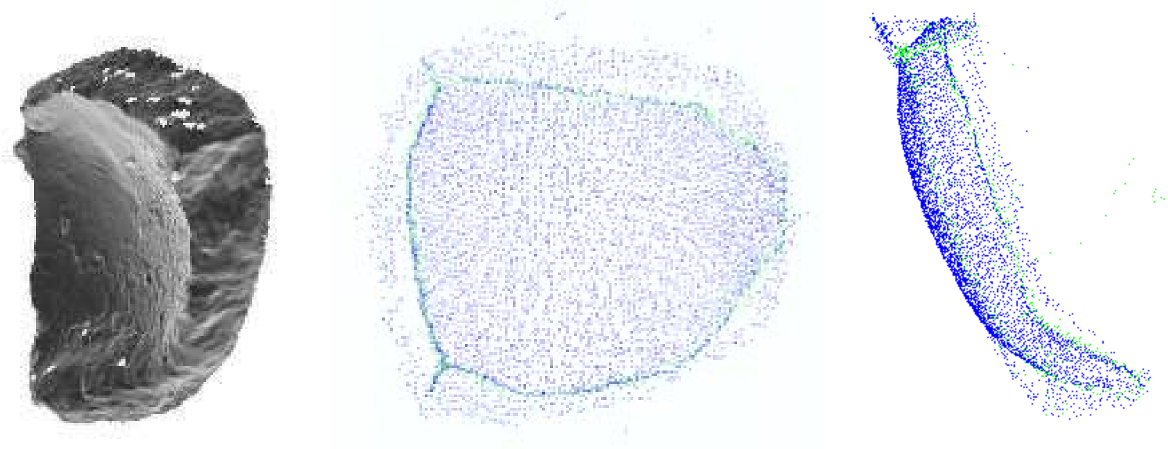


Figure 16: Example of the computation of the fine (sub-voxel) trace of skeletal loci for a 3D scanned pottery sherd shown on the left. Two views of the computed skeletal loci are given to the right. In blue are indicated surface loci (skeletal sheets), while green loci correspond to skeletal (axial) curves (see [40] for more details).