The SHAPE Lab: New Technology and Software for Archaeologists

Frederic F. Leymarie

Division of Engineering, Brown University, The SHAPE Lab

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

e-mail: leymarie@lems.brown.edu

David B. Cooper

Division of Engineering, Brown University, Box D

e-mail: cooper@lems.brown.edu

Martha Sharp Joukowsky

Center for Old World Archaeology and Art, Dept. of Anthropology, Brown University, Box 1921

e-mail: martha_joukowsky@brown.edu

Benjamin B. Kimia

Division of Engineering, Brown University, Box D

e-mail: kimia@lems.brown.edu

David H. Laidlaw

Dept. of Computer Science, Brown University, Box 1910

e-mail: dhl@cs.brown.edu

David Mumford

Division of Applied Mathematics, Brown University, Box F

e-mail: mumford@dam.brown.edu

Eileen L. Vote

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

e-mail: 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 Centre 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 artefacts. 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 modelling, 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 analyse and reconstruct pottery, compare objects from different sites and reconstruct sculpture and architecture.

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

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 modelling 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 archaeo-

Figure 1: Typical fragments of head sculptures from Petra, to be analysed. These fragments have considerable detail although some parts are missing and the surface is badly eroded.

logical analysis with site features, topography, architecture artefacts 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.

These research topics are to be explored through an integrated effort because, on the one hand, the archaeology applications drive the 3D modelling 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 hitherto 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 (see an aerial view of the Great Temple in Vote et al. this volume, figure 1) where on-going excavations have been conducted for many years, and from which a large database of artefacts is currently being built (Joukowsky 1998).

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.

1. Excavators want to be able to register the location of artefacts in situ in order to maintain an accurate archaeological excavation record (see section 2). Our proposed technology will allow archaeologists to use relatively inexpensive equipment to expedite excavations and maintain more comprehensive, accurate and accessible records of artefact *geometry* and find locations.

2. Currently artists assist in work on site by documenting the artefacts found and positing reconstructions of broken artefacts, 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 artefacts.

3. A significant problem in archaeology is the inability to compare many artefacts stylistically, which requires substantial physical information (e.g., see figure 1). Relating one artefact to another, perhaps found in another site, is an integral part of discovering its role, age, responsible artisan or community, etc. The expression of artefacts in a *shape language* will advance possibilities for interactive or automatic quantitative and qualitative comparison.

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.

However, we also make use of active data acquisition techniques, such as laser scanning, structured light and computerized tomography (CT, see figure 2) 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. (2000) who advocate the use of both passive and active systems for different purposes in documenting archaeological sites.

1.2. Shape modelling 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 section 2, the present day situation at our main site of excavation, at Petra. Then, in section 3, we give some early results in tackling the previously introduced research objectives. Finally, in section 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-levelled precinct measuring 35 meters east-west, and 42.5 meters in length. 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 (Joukowsky 1980). 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 artefact. 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

Figure 2: Figure made from density measurements of a box of sherds. Densities within the box were imaged volumetrically using a Rhode Island Hospital CT scanner. Iso-density surfaces were then created and rendered using marching cubes (Mortensen and Barrett 1998), a computer graphics technique.

of registration. All artefacts 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 artefacts 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 artefacts, recorded since 1993 (Joukowsky 1999). Unfortunately, the full potential of archaeological databases is rarely realized. Most archaeologists are not able to analyse the geometric characteristics of artefacts and their spatial relationships with other elements of the site (Crescioli and Niccolucci 1998).

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 (Leymarie et al. 1996). Our proposed technology will also permit archaeologists to reconstruct of broken or eroded fragments. Once 3D information is gained about artefacts and architectural fragments while objects are being initially registered, it will be possible to better exploit reconstruction possibilities. A series of pot fragments (figure 2 and 6) 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. In many cases, archaeological artefacts 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]

Figure 3: 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.

3. Early results

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. Gestural and voice-driven user interface is being used for interactively modelling 3D objects in a virtual environment. In figure 3, simple geometric models here are overlaid on a photograph from the site at Petra.

Examples of our recent work are 3D widgets (Brook Conner et al. 1992), free-form deformations (Hsu and Hughes 1992) and gestural interfaces (Zeleznik et al. 1996). 3D widgets demonstrate how parameters can effectively be represented by 3D geometry and embedded in a 3D dataspace. Our system *Sketch* (Zeleznik et al. 1996) 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 (Vote et al., this volume).

3.2. Three dimensional reconstruction from a single image

We have conducted preliminary work in order to extend our current *Sketch* system (Zeleznik et al. 1996) to interactively generate and edit free-form 3D shape models in a sequence of images. Figure 3 illustrates our first generation system which makes use of single viewed perspective images (Williamson and Brill 1990) 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.

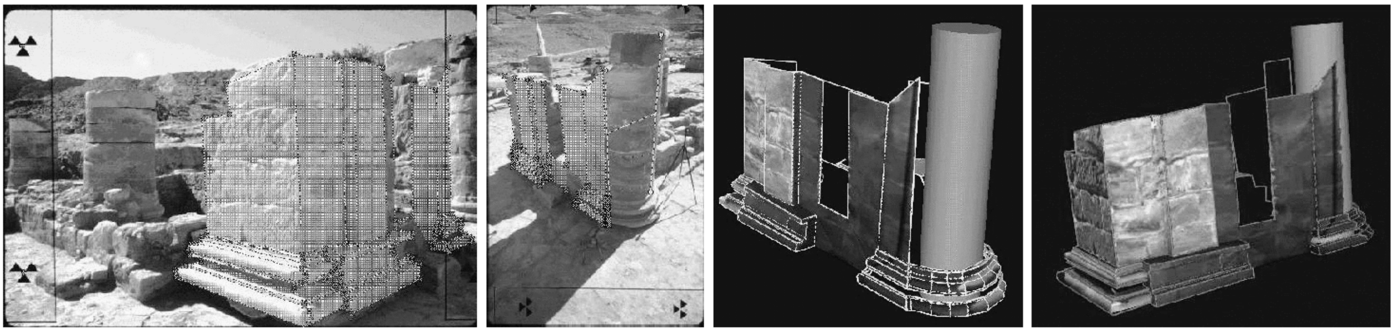

Figure 4: Illustration of photogrammetric 3D reconstruction from multiple images for architectural forms; more details in Vote (1999, in press).

Figure 5: 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 (fifth and sixth pictures).

3.3. Three dimensional reconstruction from multiple images

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 (Leymarie et al. 1996). In the context of archaeological scenery, corner detectors combined with a model-based approach (for position refinement), prove useful (Blaszka and Deriche 1994). For a video sequence, one can take advantage of the "continuity" of the sequence, by using robust tracking techniques (Leymarie et al. 1996). The correspondence problem is harder to solve for a set of photographic snapshots taken from a-priori unknown positions. This stage is user-driven in classical photogrammetry (Leymarie and Gruber 1997). To automate this task, one can make use of classical (window) correlation-based techniques combined with relaxation methods in an optimisation 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 (Faugeras et al. 1995). 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]

Similar techniques were applied to photographs taken at Petra and are illustrated in figure 4 where we performed some detailed wall reconstruction under user supervision (Vote, in press). 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. (1996). 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 figure 5) and is, thus, comparable to structured light techniques such as used in (Pollefeys et al. 2000).

3.4. Fragment representation and reassembly

The series of detailed head statuary in figure 1 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 (Joukowsky 1998). A similar problem we have solved using IPS models (see section 4.2) by matching 3D fragments of an Egyptian bust (Blane et al. 2000). The use of ridges and skeletal graphs for the same purposes represent on-going work, and more details about these methods are given in section 4.1 and section 4.3 below.

3.5. Site content discovery via 3D geometric history

For analysis, it is essential to maintain the artefacts in their architectural and topographical context. Following what Forte (2000) proposed, we have started exploring how Geographical Information System (GIS) (Kofler et al. 1996) 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 (Vote et al., this volume).

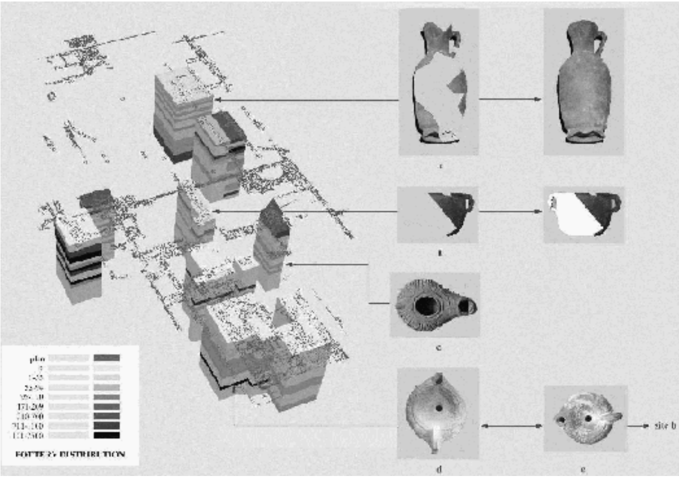

Figure 6: GIS application at Petra's Great Temple. Artefacts referenced and defined in 3D will allow archaeologists to reconstruct objects (a and b), do advanced spatial analysis, link artefacts between sites (For example, tracing lamps between sites in the region will allow us to trace trade routes.) (c, d and e), and maintain more comprehensive, accurate and accessible records of artefact geometry and find locations.

Figure 6 shows a GIS application with a 3D view of trenches. Colour 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 artefacts 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 artefacts, 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 (Vote et al., this volume).

4. Three dimensional free-form shape modelling

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 figure 10b), 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 (*Blum 1973), 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 modelling 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 section 4.3) of the shape is the centre of the osculating circle at a convex vertex (Leyton 1992) (see figure 9a and 9d). The psychologist Attneave proposed that these were the most perceptually salient and informative points on the contour (Attneave 1954).

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 (figure 7a), 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* (Cipolla et al. 1995).

One goal is to use these features to describe 3D shape in an intuitive way. In figure 7b, ridge computations on a sherd surface data obtained via CT scanning (cf. figure 2) are depicted as different shades of grey corresponding to different measures of extremal curvature. We have developed an interactive algorithm to extract ridges and valleys based on this curvature map; this is illustrated in figure 7c. 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 (Leymarie and Levine 1993). 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 modelling 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 artefacts as well as faces, animals, humans, sculptures of various types, furniture and tools, etc. There has been psychophysics on the human perception of ridges (Phillips 1997) and an additional goal is to characterize how *stable* ridges are for shape modelling.

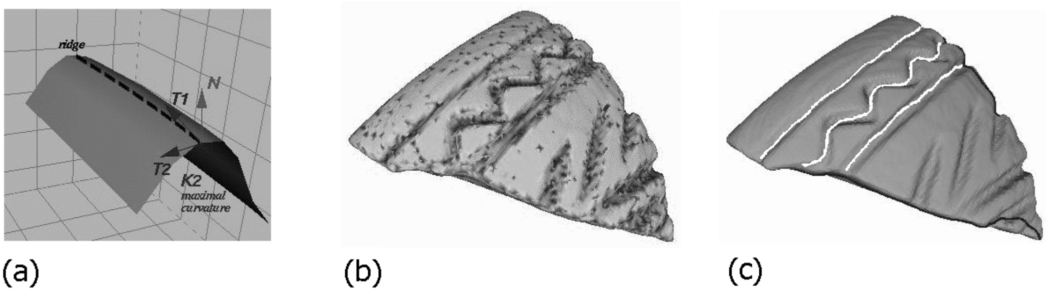

Figure 7: (a) Concept of a ridge as a line of maximal principal curvature on a surface. (b)Local curvature computations on a sherd from its recovered 3D surface. (c) Result of an interactive ridge and valley computation using a 3D active contour model (see Andrews 2000 for more technical details).

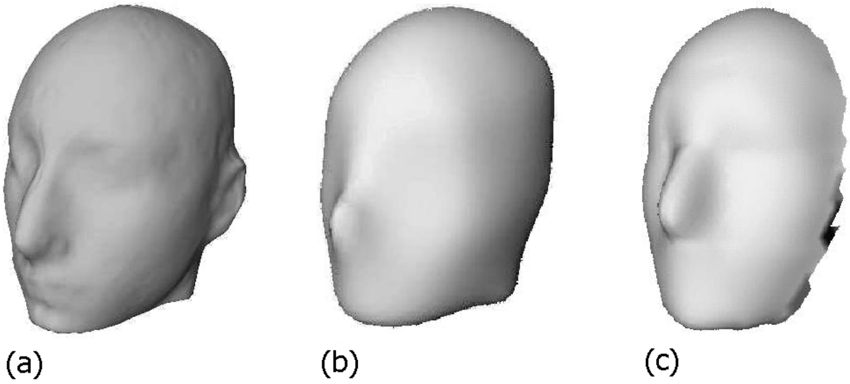

Figure 8: (a) Original triangulated head data set. (b) 10th degree fit (ears discarded). (c)Reconstruction via 12 patches, using 4th degree IPS for each patch.

4.2. Surface modelling through implicit polynomials

Multivariate Implicit Polynomials provide a powerful and rich representation for 2D and 3D curves and surfaces (Bloomenthal et al. 1997). 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, $f(x,y,z) = \sum_{i+j+k \le d} c_{ijk} x^i j^j z^k = 0$. These surfaces are generalizations, to more complicated shapes, of the conics, e.g., a hyper-ellipsoid, a cylinder with hyperbolic cross section, etc.

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 centre $x = x_0$, $y = y_0$, $z = z_0$. These IPS are useful for representing blobby closed surfaces, open patches as in figures 1, 4, 6 and 8, surfaces attached to prominent ridges, generalized cylinders (Naeve and Eklundh 1994), 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. Some of these uses are illustrated in figure 8.

A goal of our research program is to explore the use of ridges (section 4.1) and skeletal graphs (section 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 (Blane et al. 2000), and is further regularized by the use of *ridge regression* (Blane et al. 2000). 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* (Naeve and Eklundh 1994) 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

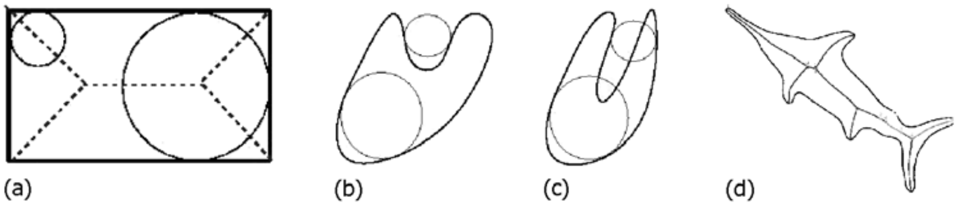

Figure 9: 2D skeletons: (a) Skeleton of a rectangle with two examples of maximally inscribed circles; (b) example of two bitangent circles which are maximal and (c) two which are not, since they cross the boundary; (d) More complex skeleton for a swordfish outline - note that at each curvature extremum of the boundary corresponds the end of a skeleton branch.

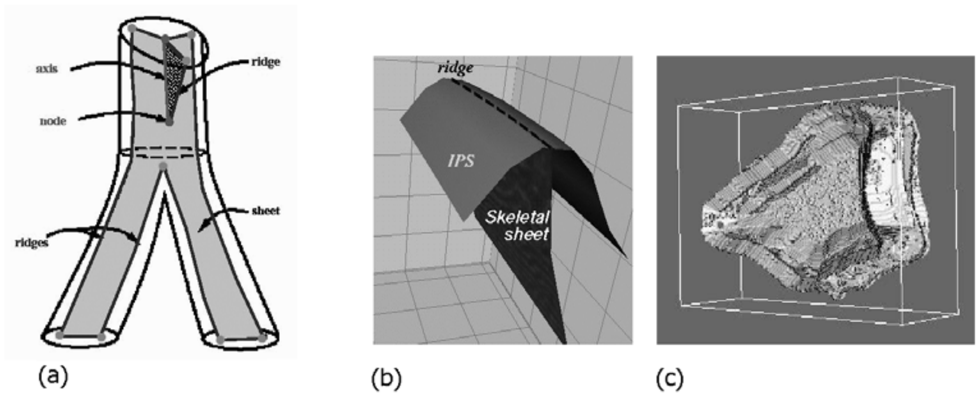

Figure 10: (a) Sketch of the 3D skeleton of a branching shape: the skeleton is made of sheets, axial and ridge curves, and nodes. (b) Concept of skeletal sheet and its relation to IPS and ridges. (c) Computed skeletal sheets of a carpal bone.

orthogonal to the axis and has a local coordinate system determined by the skeletal sheet.

4.3. Volumetric modelling through skeletal graphs

The skeletal graph of a 2D shape, or its medial axis, is the locus of centres of *maximally inscribed circles* (see figure 9). It is an intuitive and efficient representation for the recognition of 2D shapes, since variations in shape often leave the graph structure intact (Blum 1973). 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 (Leymarie and Levine 1992) 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 (Kimia et al. 1995).

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 (Giblin and Kimia 1999) 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.

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 (Leymarie and Kimia 2000). The points of medial axis (and shock set) have been classified resulting in a hypergraph representation (Giblin and Kimia 1999) 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 (Naeve and Eklundh 1994). These curves interact only at special points, namely when they intersect each other at nodes. These are the only generic possibilities (figure 10). The skeletal hypergraph describes the connectivity among symmetries of each portion of the shape (see Leymarie and Kimia (2000) for more details).

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 (Sharvit et al. 1998). 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 (Ucoluk and Toroslu 1999).

5. Conclusions

The SHAPE Lab has been created with the goals of: (i) introducing new geometric modelling 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 artefacts) 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 fulfil 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 (Mumford et al. 1999). 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 (section 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 (figure 5).

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 (Cernuschi-Frias et al. 1989). In order to apply this method to free-form shapes, we need priors of the space of such shapes (Mumford 1996). 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 (1996) 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.

Acknowledgments

Major funding for this project comes from the USA NSF/KDI grant #BCS-9980091. Frederic Leymarie was partially supported by a Ph.D. 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

ANDREWS, S., 2000. *Interactive generation of feature curves on surfaces*. Master's thesis, University of Toronto, Ontario, Canada.

ATTNEAVE, F., 1954. Some informational aspects of visual perception. *Psychological Review*, 61(3):183-193.

AYERS, M.R. and ZELEZNIK, R.C., 1996. The lego interface toolkit. In *Proceedings of the ACM Symposium on User Interface and Software Technology (UIST'96)*:97-98.

BELHUMEUR, P. and KRIEGMAN, D., 1996. What is the set of images of an object under all possible illumination conditions? In *Proceedings of CVPR*: 270-277. San Francisco, CA, IEEE CS Press.

BELHUMEUR, P., KRIEGMAN, D. and YUILLE, A., 1997. The bas-relief ambiguity. In *Proceedings of CVPR*:1060-1066. San Juan, PR, IEEE CS Press.

BLANE, M.M., LEI, Z. and COOPER, D.B., 2000. The 3L algorithm for fitting implicit polynomial curves and surfaces to data. *IEEE Trans. on PAMI*, 22(3):298-313.

BLASZKA, T. and DERICHE, R., 1994. Recovering and characterizing image features using an efficient model based approach. Research Report RR-2422, INRIA, Sophia Antipolis, France.

BLOOMENTHAL, J., BAJAJ, C., BLINN, J., CANI-GASCAUEL, M.P., WYVILL, B. and WYVILL, G., 1997. *Introduction to Implicit Surfaces*. Morgan Kaufmann Pub., San Francisco, CA, USA.

BLUM, H., 1973. Biological shape and visual science. *Journal of Theoretical Biology*, 38:205-287.

BROOK CONNER, D., SNIBBE, S.S., HERNDON, K.P., ROBBINS, D.C., ZELEZNIK, R.C. and Van DAM, A., 1992. Three-dimensional widgets. *Computer Graphics (1992 Symposium on Interactive 3D Graphics)*, 25(2):183-188.

BRUCE, J.W., GIBLIN, P.J. and TARI, F., 1996. Ridges, crests and sub-parabolic lines of evolving surfaces. *International Journal of Computer Vision* 19:195-21.

CERNUSCHI-FRIAS, B., COOPER, D.B. and HUNG, Y.P., 1989. 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.

CIPOLLA, R., FLETCHER, G. and GIBLIN, P.J., 1995. Surface geometry from cusps of apparent contours. In *Proc. Int. Conf. on Computer Vision (ICCV'95)*:858-863. IEEE Computer Society Press.

COHEN, J.M., MARKOSIAN, L., ZELEZNIK, R.C., HUGHES, J.H. and BARZEL, R., 1999. An interface for sketching 3D curves. In *Proceedings of 1999 Symposium on Interactive 3D Graphics*, Atlanta, ACM.

CRESCIOLI, M. and NICCOLUCCI, F., 1998. 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.*, BAR International Series 757:133-134.

DEBEVEC, P.E., TAYLOR, C.J. and MALIK, J., 1996. Modeling and rendering architecture from photographs. In *Computer Graphics Proceedings, SIGGRAPH'96*. ACM.

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

FORTE, M., 2000. About virtual archaeology: Disorders, cognitive interactions and virtuality. In Barceló, J.A., Forte, M. and Sanders, D.H. (eds.), *Virtual Reality in Archaeology*. BAR International Series 843:247-259, Oxford, Archaeopress.

GIBLIN, P.J. and KIMIA, B.B., 1999. On the intrinsic reconstruction of shape from its symmetries. In *Proc. of CVPR*:79-84.

GIBLIN, P.J. and KIMIA, B.B., 1999. On the local form and transitions of symmetry sets, and medial axes, and shocks in 2D. In *Proc. of ICCV*:385-391.

GIBLIN, P.J. and KIMIA, B.B., 2000. On the local form of symmetry sets, and medial axes, and shocks in 3D. In *Proc. of CVPR.* IEEE Computer Society 2000.

HAGER, G. and BELHUMEUR, P., 1996. Real time tracking of image regions with changes in geometry and illumination. In *Proceedings of CVPR*. IEEE CS Press.

HSU, W. and HUGHES, J.F., 1992. Direct manipulation of free-form deformations. In *Computer Graphics Proceedings, SIGGRAPH'92*:177-184.

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

JOUKOWSKY, M.C.S., 1996. *Early Turkey: An Introduction to the Archaeology of Anatolia from Prehistory through the Lydian Period*. Kundall-Hunt.

JOUKOWSKY, M.C.S., 1997. Brown University 1997 excavations at the Petra Great Temple. *Annal of the Department of Antiquities of Jordan*, XLII:293-318.

JOUKOWSKY, M.C.S., 1997. The Petra southern temple: The fourth season 1996. *American Journal of Archaeology*:101:339.

JOUKOWSKY, M.C.S., 1998. The Great Temple at Petra. *American Journal of Archaeology*, 102(3):593-596.

JOUKOWSKY, M.C.S., 1999. *Petra Great Temple - Volume 1: Brown University Excavations 1993-1997*.

KIMIA, B.B., TANNENBAUM, A.R. and ZUCKER, S.W., 1995. Shapes, shocks, and deformations, I: The components of shape and the reaction-diffusion space. *International Journal of Computer Vision*, 15:189-224.

KOFLER, M., REHATSCHEK, H. and GRUBER, M., 1996. 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:198-202.

LEI, Z., TASDIZEN, T. and COOPER, D.B., 1998. PIMs and invariant parts for shape recognition. In *Proc. of International Conference on Computer Vision (ICCV'98)*:827-832. IEEE CS Press.

LEYMARIE, F. and GRUBER, M., 1997. Applications of virtual reality: CyberMonument and CyberCity. In *Proc. of the 3rd European Digital Cities Conference*, Berlin, Germany. European Commission.

LEYMARIE, F. and KIMIA, B.B., 2000. Discrete 3D wave propagation for computing morphological operations from surface patches and unorganized points. In Goutsias, J., Vincent, L. and Bloomberg, D. (eds.), *Math. Morphology and its Applications to Image and Signal Processing*, volume 18 of *Computational Imaging and Vision*:351-360.

LEYMARIE, F. and LEVINE, M.D., 1992. Simulating the grassfire transform using an active contour model. *IEEE Trans. on PAMI*, 14(1):56-75.

LEYMARIE, F. and LEVINE, M.D., 1993. Tracking deformable objects in the plane using an active contour model. *IEEE Trans. on PAMI*, 15(6):617-634.

LEYMARIE, F. et al., 1996. REALISE: Reconstruction of reality from image sequences. In Delogne, P. (ed.), *Proc. of IEEE International Conference on Image Processing (ICIP'96)*, volume III:651-654.

LEYMARIE, F., 1997. 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:10-29. European Commission.

LEYMARIE, F., KIMIA, B.B. and GRUBER, M., 1998. *Direct camera pose evaluation from a single view and a map*. Technical Report LEMS-174, Brown University.

LEYTON, M., 1992. *Symmetry, Causality, Mind*. MIT Press, Cambridge.

MARKOSIAN, L., COHEN, J., CRULLI, T. and HUGHES, J., 1999. Skin: A constructive approach to modeling free-form shapes. In *Proceedings of SIGGRAPH'99*.

MORTENSEN, E.N. and BARRETT, W.A., 1998. Interactive segmentation with intelligent scissors. *Graphical Models and Image Processing*, 60(5):349-384.

MUMFORD, D., 1996. *The Statistical Description of Visual Signals*. Akademie Verlag.

MUMFORD, D., GIBLIN, P., GORDON, G., HALLINAN, P. and YUILLE, A., 1999. *Two and Three dimensional Patterns in Faces*. A.K. Peters.

NAEVE, A. and EKLUNDH, J.O., 1994. Generalized cylinders - what are they? In Arcelli, C., Cordella, L.P. and Di Baja, G.S. (eds.), *Aspects of Visual Form Processing*, 2nd International Workshop on Visual Form:384-409. World Scientific.

PHILLIPS, F., 1997. *Geometric structure, frames of reference, and their implication in the localization of features on smoothly curved surfaces*. PhD thesis, Ohio State University.

POLLEFEYS, M., PROESMANS, M., KOCH, R., VERGAUWEN, M. and Van GOOL, L., 2000. Acquisition of detailed models for virtual reality. In Barceló, J.A., Forte, M. and Sanders, D.H. (eds.), *Virtual Reality in Archaeology*. BAR International Series 843:71-77. Oxford, Archaeopress.

SCHMITT, G. et al., 1995. Toward virtual reality in architecture: Concepts and scenarios from the architectural space laboratory. *Presence*, 4(3):267-285.

SHARVIT, D., CHAN, J., TEK, H. and KIMIA, B.B., 1998. Symmetry-based indexing of image databases. *Journal of Visual Communication and Image Representation*, 9(4):366-380.

STECKNER, C., 2000. Form and fabric, the real and the virtual. In Barceló, J.A., Forte, M. and Sanders, D.H. (eds.), *Virtual Reality in Archaeology*. BAR International Series 843:121-128. Oxford, Archaeopress.

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

SUBRAHMONIA, J., COOPER, D.B. and KEREN, D., 1996. 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.

SUBRAHMONIA, J., KEREN, D. and COOPER, D.B., 1994. An algebraic and Bayesian technology for object recognition. In Mardia, K.V. (ed.), *Advances in Applied Statistics*:165-205. Carfax Publishing Co., Abingdon, Oxfordshire.

TAREL, J.P., CIVI, H. and COOPER, D.B., 1998. Pose estimation of free-form 3D objects without point matching using algebraic surface models. In *Workshop on 3D Modeling, ICCV98:*13-21.

TAREL, J.P., GUICHARD, F. and AUBERT, D., 1999. Tracking occluded lane-markings for lateral vehicle guidance. In *Proc. of IEEE International Conference on Circuits, Systems, Communications and Computers (CSCC'99)*.

UCOLUK, G. and TOROSLU, I.H., 1999. Automatic reconstruction of broken 3-D surface. *Computers and Graphics*, 23(4):573-582.

VOTE, E., in press. Using desktop photogrammetry to document archaeological remains: The Great Temple at Petra, Jordan. In Cooney, G. and Masterson, B. (eds.), *Proc. Int. Conf on Computer Applications in Archaeology (CAA'99)*.

VOTE, E., FELIZ, D.A., LAIDLAW, D. and JOUKOWSKY, M.S., 2000. ARCHAVE: A virtual environment for archaeological research. In *Proc. Int. Conf on Computer Applications in Archaeology (CAA2000)*. This volume.

WILLIAMSON, J.R. and BRILL, M.H., 1990. *Dimensional Analysis through Perspective - A Reference Manual*. Kendal/Hunt Publishing co.

ZELEZNIK, R.C. and FORSBERG, A.S. and STRAUSS, P.S., 1997. Two pointer input for 3D interaction. In *Proceedings of 1997 Symposium on Interactive 3D Graphics*:27-30.

ZELEZNIK, R.C. and FORSBERG, A.S., 1999. Unicam - 2D gestural camera controls for 3D environments. In *Proceedings of 1999 Symposium on Interactive 3D Graphics*.

ZELEZNIK, R.C., HERNDON, K. and HUGHES, J.F., 1996. SKETCH: An interface for sketching 3D scenes. In *Computer Graphics Proceedings, SIGGRAPH'96*, Annual Conference Series:163-170.

ZHU, S.C. and YUILLE, A.L., 1996. FORMS: A flexible object recognition and modeling system. *International Journal of Computer Vision* 20(3):187-212.

Notes

[1] SHAPE: SHape, Archaeology, Photogrammetry, Entropy; a multidisciplinary 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 (Vote et al., this volume).

[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 is used. These simpler models are typically made of regular primitives to represent simple architectural shapes (Streilhen 1994).

[5] The syntaxic properties of our shape language will be reported elsewhere; see (Giblin and Kimia 1999) for early theoretical investigations.