Modeling and Analysis of Empirical Data in Collaborative Environments

Many tasks in such diverse fields as medicine, oil exploration, and mechanical design depend on the collection and interpretation of empirical data. Our long-term vision is a collaborative scientific visualization environment, where scientists, engineers, and physicians work together on modeling and analyzing empirical data using an integrated set of tools and techniques from computer graphics, computer vision, and image processing. This environment would also take advantage of interaction techniques that include sound, spoken commands, and gestures.

To illustrate how such an environment could be used, let us look at some hypothetical scenarios. In the first scenario, a radiologist gives an "interactive report" on findings in a CT scan to a surgeon at a remote location. This report contains three-dimensional data that is manipulated by the radiologist to illustrate features of interest. The surgeon can see the radiologist as the report is delivered. The surgeon can also interject questions and manipulate the data to clarify the questions. Alternatively, this report can be left in the surgeon's electronic mail, to be accessed at a later time. Similarly, the surgeon may want to consult the radiologist from the operating room and discuss a preoperative or intraoperative scan. The surgeon and the radiologist can carry on a conversation between the operating room and the radiologist's office, while the radiologist monitors the operation on a workstation. In addition, both doctors can simultaneously manipulate, point at, and analyze the three-dimensional data.
A second scenario from the field of medicine is preoperative facial analysis and postoperative quantitative evaluation. Patients expecting surgery to improve their facial paralysis are preoperatively scanned to construct a synthetic facial model. In this model, synthetic muscles are activated to simulate and improve the patient's appearance. After the surgery, the surgeon and the referring neurologist can examine any improvement by quantitatively comparing pre- and postoperative image sequences of the patient making facial expressions.

In many research domains, the ability to share massive amounts of empirical data is crucial to significant progress. One example is neuroscience, where researchers work only on a very small part of the bigger problem of mapping the brain and its functions. Published papers may contain only one or two images of a 2,000-image dataset, which is in general inaccessible to other researchers. Currently, there is a nationwide effort to establish a central database of neuroscience data [11], and a collaborative visualization environment could be used to access such a database.

In another domain, engineers at different locations collaborate on the design of a new product. Starting with a geometric model of last year's version, they interactively modify and refine the model through local changes to its shape, color, and texture. Alternatively, a manufactured object, such as a competitor's product or a clay mock-up, may be the starting point from which a computer-aided design (CAD) model is automatically created and subsequently modified. As in the biomedical examples, the engineers can consult colleagues and collaboratively modify the design over the computer network.

To support such scenarios, we envision a scientific visualization environment incorporating video, sound, and user interfaces that understand spoken requests and gestures. This vision suggests a seamless environment in which the user can consult a colleague, while the conversation and respective faces are transmitted and displayed. Furthermore, the computer is no longer passive, but an active participant understanding limited speech, responding with an expressive face and a synthesized voice.

The sources of data in the preceding examples, namely CT scanners, transmission electron microscopes, and video cameras, are different and the different types of data require considerable domain knowledge in their analysis. Nevertheless, there is a great deal of commonality in the techniques used to visualize, quantify, and interact with the data and the models. These common techniques range from data enhancement, feature extraction, and reconstruction, to interactive visualization, computation, and simulation on the models.

We are capitalizing on these similarities by building a common scientific visualization environment which supports such diverse applications as biomedical scientific visualization and reverse engineering applications. We discuss three projects we are pursuing toward our goal of such an environment: (1) interactive modeling and visualization of medical and biological data, (2) three-dimensional shape acquisition, modeling, manipulation, and culminating in three-dimensional faxes, (3) teleconferencing with personal computers. These projects are demonstrated at the SIGGRAPH'92 Showcase.

**Interactive Modeling and Visualization of Medical and Biological Data**

Interactive modeling and visualization of medical and biological data in a collaborative environment has the potential for improving patient care and reducing medical costs. Visualization environments for these domains must address all stages of data analysis, including registration, segmentation, three-dimensional reconstruction, rendering, analysis, and simulation.

We discuss some fundamental methods in each area and demonstrate examples from neuroscience, embryology, radiology, and surgical planning. The components of the visualization environment are implemented in a collaborative environment, suggesting that biomedical imaging, medical diagnosis, and surgical planning will soon be feasible over high-speed networks, allowing electronic intra- and inter-hospital collaboration among physicians and researchers.

**Registration**

Registration refers to the alignment of data from the same or different modalities, or sensors, such as the alignment of slices of neural tissue, or the alignment of MR and CT data. We use two registration techniques: interactive registration with a digital blink comparator, and automatic registration through minimization. The first, the digital blink comparator, is a manual technique using visual motion for the alignment of images. Holding one image stationary, the user translates and rotates the other image while the stationary and moving images are alternately shown on a graphics screen. The images are aligned when the sum of the visual motion is minimized [5]. The blink comparator is used to register sections of a neuronal dendrite from transmission electron microscopy, as well as pre- and postcontrast midsagittal MR head scans.

The automatic method attempts to minimize the difference between the two images to be registered. One image is kept stationary and the other is rotated and translated, in a systematic manner, with respect to the stationary image. The transformation parameters which yield the minimum difference between the images also yield the optimum alignment. This registration technique is used to register serial sections of an embryo from light microscopy.

**Segmentation and Reconstruction**

Segmentation refers to the process
of extracting meaningful regions from images or volumes. We employ a user-assisted segmentation method, snakes, or interactive deformable contours, which is a model-based image feature localization and tracking technique [5,7]. Although consistent with manual tracing methods, snakes are considerably faster and more powerful. The user quickly traces a contour which approximates the desired boundary, then starts a dynamic simulation that enables the contour to locate and conform to the true boundary. Where necessary, the user may guide the contour using a mouse to apply simulated forces. With some guidance, snakes exploit the coherence between adjoining images to quickly extract a sequence of regions. Snakes are used to extract a neuronal dendrite from transmission electron micrographs (Figure 1), and an embryo heart and its substructures from light micrographs, and to track lung and liver vessels in spiral CT scans. After extracting all the regions of an object, we stack them to form three-dimensional volumetric models of a neuronal dendrite (Figure 2a), an embryo heart, and the heart and liver vessels.

**Rendering**

The three-dimensional volume data is rendered using a novel parallel volume ray-tracing algorithm. This ray-tracing algorithm differs from previous methods by holding the data stationary while accumulating the opacity along the rays in parallel. The three-dimensional volumetric models can be interactively rotated, cut open (Figure 2b), and viewed transparently. The three-dimensional shapes of the models are displayed, using both shading and motion parallax. Emphasis is placed on data fidelity, or loss-less rendering [4], using accurate interpolation filters [3].

**Simulation**

In a different application, we simulate facial tissue dynamics for surgery planning. Two iso-valued surfaces of the facial skeletal bone and epidermal skin tissue are extracted from CT data using the marching cubes algorithm [10] (Figure 3). These geometric surfaces provide a foundation for a discrete layered spring lattice tissue model [21]. Three layers are constructed for the numerical simulation: the skeletal bone, the muscle, and the epidermis. One end of each synthetic muscle attaches to the muscle layer and the other end attaches rigidly to bone. When a muscle is articulated, it deforms the lattice, which causes forces to propagate outward until an equilibrium is established. The simulation is implemented on a massively parallel computer to achieve a rapid response.

**Three-Dimensional Shape Acquisition, Modeling, and Manipulation**

Three-dimensional modeling is required in applications such as computer-aided design and manufacturing (CAD/CAM), architectural design, and computer animation. Traditionally, three-dimensional model entry involves a laborious manual process based on two-dimensional input devices such as tablets. The recent advent of interactive three-dimensional input techniques [12] and automatic shape acquisition using special-purpose three-dimensional active range finders [6] could alleviate some of these problems. However, these approaches rely on expensive special-purpose input devices that may not be available to the general public.

In this section we describe an alternative shape acquisition technique based on regular video images, which we expect to become widely applicable as video technology becomes embedded in workstations. In our system, the object to be scanned rotates on a turntable in front of an ordinary, stationary video camera (Figure 4a). We have developed two algorithms to reconstruct the object from the resulting image sequence: shape from silhouettes, and shape from image flow. On acquiring the object's shape, we can interactively modify it using a variety of three-dimensional manipulation techniques.

**Three-Dimensional Shape from Silhouettes**

Our first algorithm for shape acquisition constructs a bounding volume for the object from the sequence of silhouettes [15]—the binary classification of images into object and background (Figure 4b). The three-dimensional intersection of these silhouettes defines a bounding volume within which the object must lie. We represent this volume using an octree [13] (Figure 4c). For each silhouette, the octree cubes are projected into the image plane and classified as either wholly or partially inside or outside the object. After one complete revolution at a given resolution, cubes which cannot be unambiguously classified are subdivided into eight subcubes, and the process is repeated [15]. Once the complete shape of the object has been reconstructed, we associate each octree cube with a set of pixels in the input images, producing a "texture-mapped" three-dimensional object. The complete procedure can be performed in a few minutes on a workstation, in time proportional to the desired resolution.

**Three-Dimensional Shape from Image Flow**

Our second algorithm computes the optic flow (two-dimensional motion) at each pixel to estimate the three-dimensional location of points on the surface of the object [16]. The optic flow is computed from successive pairs of images by minimizing a correlation measure at each pixel. This produces both a dense estimate of the local motion (Figure 5b), and a confidence measure which depends on the amount of variation in the local texture in the image (Figure 5c). Flow measurements are converted into three-dimensional points on the surface using the known turntable motion (Figure 5d). These three-
dimensional positions are then refined by merging them with new measurements from successive images using a statistical framework [16]. The final model is a collection of points on the object's surface tagged with colors and intensities derived from the set of images. Since the computation of flow and the merging points are computationally intensive, we have implemented these algorithms on a massively parallel processor.

Three-Dimensional Surface Interpolation and Manipulation
Because the flow-based reconstruction algorithm only produces estimates at locations with sufficient texture, there may be gaps in the surface. To solve this problem, we have developed a flexible surface modeling technique based on interacting particles [17]. Our method interpolates across gaps in the surface by placing particles at the initial surface measurements and then adding new particles automatically. The surface is automatically smoothed, based on specially tailored interaction potentials. Each particle is then colored with an appropriate intensity derived from the initial image sequence, producing a texture-mapped surface model of the object. To refine or reshape the object, we can use traditional techniques such as global or local free-form deformations [2,14], as well as our particle-based surface modeler. This latter approach is capable of extending or topologically altering surfaces by cutting or merging particle sheets [17].

Applications
The primary application we demonstrate at the SIGGRAPH'92 Showcase is a three-dimensional fax with collaborative design revision. By three-dimensional fax, we mean the ability to transmit to a remote site the full three-dimensional description of a real object. Once the

![Figure 1. Cell segmentation using a deformable contour](image1)

![Figure 2. This reconstructed dendrite was built using an earlier version of our system which was implemented on a Silicon Graphics 4D/220GTX. Voxel-View was used for volume rendering.](image2)

![Figure 3. a) Iso-value surface of bone b) Iso-value surface of skin tissue](image3)

![Figure 4. Shape reconstruction from Image silhouettes: a) one of the original images, b) a thresholded silhouette image and c) an octree representation of the internal model](image4)
three-dimensional object model has been entered, it can be interactively modified or refined at each site, both through local changes to its shape, and through manipulation of the texture map (painting). The object could also be reproduced in three-dimensions using NC milling or stereolithography.

These techniques also have applications in reverse engineering (deriving CAD descriptions from manufactured objects), in the creation of virtual reality environments, and in the creation of object models for computer graphics animation. We expect the addition of simple automatic shape acquisition capabilities, as well as the wider availability of networked display and manipulation capabilities, to greatly enhance the usefulness and power of existing three-dimensional modeling systems.

Teleconferencing with Personable Computers

The face is a powerful medium of communication, and humans are highly tuned to comprehend subtle and complex facial signals. An articulating synthetic face suggests novel scenarios for presenting information. By adding such a face to a synthetic speech, we can increase the bandwidth and the expressiveness of the spoken word. If the synthetic speech/face generator were combined with a system that performs basic facial analysis of the user, tracks the focus, decodes emotional states from facial expression, and analyzes the user's speech, we would transform the computer from an inert box into a personable computer.

How does this technology fit into a collaborative scientific visualization environment? Perhaps, it could be the front end of an expert system for empirical data analysis. A synthetic character could present some technical knowledge and, on being asked a question, explain how a piece of information was derived from the existing data. The advantage of the synthetic character over a textual interface will, we conjecture, be similar to the advantage of interacting with a person rather than reading explanations to understand a technical problem.

Within the bounds of today's technology we can demonstrate teleconferencing of today and of the future. Users can talk not only to people but also to a remote computer answering simple questions. The computer will have a personable character with expressive faces synchronized with synthetic speech.

Modeling Three-Dimensional Facial Structure

Three-dimensional face data captured from high-speed active scanners can be used to model facial structure [6,19]. These datasets are mapped onto a discrete deformable mesh with a known topology, enabling the use of a standard articulation model, and also reducing the size of the data. This deformable mesh behaves as an elastic mask, with some parts fixed on key areas of the face (e.g., eyes), and others free to move over the facial structure. Once the mesh establishes an equilibrium, it is frozen and used as the foundation for articulation (Figure 6).

Articulation and Control

We have developed techniques for articulating facial geometries through the use of synthetic muscle actuators [20]. This approach, augmented by facial tissue models, yields realistic facial expressions [21]. Two primary muscle types are used in the facial model: linear muscles which pull in an angular direction, and sphincter muscles which squeeze, like the drawing together of a string bag. Only the most significant facial muscles are used. These muscles can operate in isolation or as small functional groups to generate facial expressions and articulate the mouth for speech. Facial expressions such as happiness, sadness, anger, fear, disgust, and surprise are accomplished by grouped muscle activities (Figure 7).

Speech Synchronization

The facial model is synchronized to an automated speech synthesizer [8]. The synthesizer converts regular text into a phonetic transcription annotated with timing, intonation, and stress information as well as audible sound. The phonetic transcription is coordinated with the muscle activation of the lips, resulting in a synthetic character that appears to speak.

Demonstrations

We demonstrate two scenarios at the SIGGRAPH92 Showcase. First, video teleconferencing of today, using color images of real people interacting and talking while their mutual images are displayed in live video windows. Second, teleconferencing of tomorrow, using synthetically generated facial images. The user is able to communicate with a remote personable computer and obtain responses to a limited set of questions.

Hardware and Software

The primary platform for the demonstrations is the DECmpp 12000, a massively parallel processor, with several DECstations networked over FDDI. The bidirectional person-to-person video teleconferencing uses an experimental video JPEG compression/decompression board which provides real-time compression/decompression of video into the DECstation's main memory, as well as audio capability using DECaudio. The video approach is unusual in that video is treated as a normal data type by using unextended X servers for display rather than back-door paths into the frame buffers, as has generally been the case with "video in a window" in the past. For our inter-

Figure 5. Shape reconstruction from image flow: a) one of the original input images, b) local flow estimates (blue to the right, red to the left), c) certainty in flow estimates (red is highest certainty), and d) final 3D point cloud computed from all views.
active visualization and modeling, we use a version of AVS [18] running on top of shared X [1], extended with special-purpose modules for registration, segmentation, model building, and rendering.

Future Research

Based on the demonstrations described and other experiences from our research, we have identified a number of basic capabilities which must be added to current scientific visualization environments to make them more applicable to a wide range of data analysis problems (Also see [4,9]):

- Three-dimensional modeling and visualization environments must support a wide range of representations. For example, in the three-dimensional object reconstruction system, both octrees and distributed surface representations based on particles have been used, and additional representations based on more traditional CSG primitives and NURBS could be added. In the medical applications, two-dimensional images, two-dimensional and three-dimensional contour and surface models, and three-dimensional volumetric representations are all important.
- A large number of data presentation modalities must be available, including images, height fields, contours, surfaces with texture mapping, and volumes. The data presentation should not be determined by the internal data representation. For example, while an image is usually displayed as a grayscale or color image, it could also be displayed as a height field or as an array of numbers. Similarly, a volume can either be ray-traced or displayed as a sequence of images using visual motion to convey the internal shape.

- The scientific visualization environment should support concurrent analysis of multiple datasets from the same or different modalities and from the same or different subjects. The user should be able to visually correlate parts of these datasets in many different ways, through overlays or pointing in multiple windows.
- The rapid display and three-dimensional manipulation of data facilitates the use of motion parallax to discern three-dimensional shape. Accurate, or loss-less, rendering is also important, as is the quantification of any loss of accuracy in the data due to rendering.
- The modeling and visualization system should include automatic and semiautomatic image and data reconstruction techniques. We expect such capabilities, which do not exist in most current visualization systems, to play an increasingly important role. Existing registration and segmentation techniques do not perform well on broad classes of data. A wide variety of special-purpose registration and segmentation techniques must be developed and integrated into the environment, while research into more generally applicable techniques must continue. Similarly, the recovery of three-dimensional shape and color distributions is becoming possible under controlled conditions, and provides a powerful tool both for model acquisition and for more general image analysis. The solution of the general problems of shape recovery and segmentation is difficult, however, and is the focus of much research in computer vision.

Finally, many problems in scientific visualization are complex and require the cooperation of many different experts. In the future, these will include computers as active collaborators. A collaborative modeling and visualization environment would facilitate this task, provided that we can extend interaction paradigms for complex data analysis to multiple collaborating users.

While many of these components involve open research questions, we believe that rapid progress is being made. Including a richer set of existing analysis, modeling, and presentation techniques should greatly enhance the power and usefulness of today's modeling and visualization environments. In the future, we expect continued advances in these components, as well as increases in the performance of computers and communication networks, resulting in a new generation of visualization environments that will have a profound impact on collaborative work.

References


CR Categories and Subject Descriptors: D.2.2 [Software Engineering]: Tools and Techniques—User interfaces; H.5.1 [Information Interfaces and Presentation]: Multimedia Information Systems—Audio input/output; I.2.10 [Artificial Intelligence]: Vision and Scene Understanding—Shape; I.3.5 [Computer Graphics]: Computational Geometry and Object Modeling—Physically based modeling; I.3.7 [Computer Graphics]: Three-Dimensional Graphics and Realism—Reconstruction; I.4.3 [Image Processing]: Enhancement—registration; I.4.6 [Image Processing]: Segmentation— Edge and feature detection; I.6.8 [Simulation and Modeling]: Types of simulation—animation

General Terms: Algorithms, Human Factors

Additional Key Words and Phrases: Facial animation, medical and biological imaging, octrees

About the Authors:

INGRID CARL BOM is manager of visualization research at the Cambridge Research Lab of Digital Equipment Corporation. Her research interests include scientific visualization, medical and biological imaging, geometric modeling, and computer graphics architecture; email: carl bom@cr l.dec.com

WILLIAM M H SU is a member of the research support staff at the Cambridge Research Lab of Digital Equipment Corporation. His research interests include scientific visualization, computer graphics, geometric modeling, and parallel algorithms; email: hsu@cr l.dec.com

GUDRUN KLINKER is a member of the research staff at the Cambridge Research Lab of Digital Equipment Corporation. Her research interests include color computer vision and visualization environments for semiautomatic data interpretation. email: gudrun@cr l.dec.com

RICHARD Szeliski is a member of the research staff at the Cambridge Research Lab of Digital Equipment Corporation. His research interests include 3D computer vision, computer graphics, and parallel processing; email: szeliski@cr l.dec.com

KEITH WATERS is a member of the research staff at the Cambridge Research Lab of Digital Equipment Corporation. His research interests include computer graphics, physically based modeling, volume visualization, medical facial applications, and facial synthesis. email: waters@cr l.dec.com

Authors' Present Address:

Carl Bom, Hsu, Klinker, Szeliski, Waters, Gettys, and Levergood: Digital Equipment Corporation, Cambridge Research Lab, 1 Kendall Square, Building 700, Cambridge, MA 02139; tel. (617) 621-6650; email: last name@cr l.dec.com

Doyle: Biomedical Visualization Lab (M/C 527), College of Associated Health Professions, Univ. of Ill. at Chicago, 1919 West Taylor, RM 211, Chicago, I L 60612; tel. (312) 996-7337; fax: 312-996-8342; email: n52838@ui cvm. cc. uic.edu

Harris: Neurology Research Dept., Children's Hospital, 300 Longwood Ave., Boston MA 02115; tel. (617) 735-6373; fax: (617) 730-0636

Palmer R. and Palmer L. and Wallace: Digital Equip. Corp. 146 Main St., Maynard, MA 01754-2571; tel. (617) 493-511

Picart: Dept. of Electrical Engineering, Univ. of Mass. Lowell, 1 University Ave., Lowell, MA 01854

Terzopoulos and Tonnees: Computer Science Dept., Univ. of Toronto, 10 Kings College Rd., Toronto, Ontario, Canada M5S IA4; tel. (416) 978-7777; email: dt@ai.toronto.edu

Vannier: Mallinckrodt Institute of Radiology, Washington Univ. School of Medicine; 512 S. Kinghighway Blvd., St. Louis, MO 63110; tel. (314) 362-8467; fax 314 362-8491; email: mmvannier@dtavey.wustl.edu

VoxelView is a trademark of VitalImages, Inc. It was reproduced from [5] with the publisher's permission.

© ACM 0002-0782/92/0600-0479 $1.50

Acknowledgment

John C. Hart, guest editor of the special section on graphics extends his appreciation to the following: Tom DeFanti and Maxine Brown of the Electronic Visualization Laboratory (EVL) at the University of Chicago, and Jim George of Mesa Graphics, Inc. in Los Alamos, N. Mex., were very helpful in putting this issue together. Thanks also to Charlie Catlett of the National Center for Supercomputing Applications, University of Illinois, Champaign Urbana, for information on showcases proceedings at Supercomputing conferences. Many of the color images were submitted electronically and reproduced on film in Chicago thanks to Gary Undahl of the EVL and Mary Rasmussen of the Biomedical Visualization Laboratory at the University of Chicago. Katry Grieve of the EVL produced the images of the virtual "Showcase" exhibition.