SURFACE RECONSTRUCTION FROM IMAGE SEQUENCES

H. Harlyn Baker
Artificial Intelligence Center
SRI International
333 Ravenswood Avenue
Menlo Park, CA 94025.

Abstract

This article presents a facility developed in support of image sequence processing for an autonomous vehicle. As in man, the success of the sequence analysis depends on the ability to operate explicitly in both space and time, and on exploiting the massive redundancy present in the hundreds of views that can be obtained when moving through a scene. The mechanism in our analysis that integrates these space-time factors is a 3-D surface-building process called the *Weaving Wall* – this paper focuses on the construction and use of surfaces from this process. In robotic navigation work the surfaces represent the space-time *evolution* of scene images, and this representation, in conjunction with geometric constraints, allows us to determine the three-dimensional structure of a scene. In other domains where there is a similar gradual evolution of data over a third dimension, in medical tomography or scale-space studies, for example, the surfaces constructed by the Weaving Wall are immediately of value for their topographic structure. The designs of both the surface-building and scene-reconstructing processes make them well-suited for real-time operation given appropriate hardware. The principle behind the Weaving Wall may have a dramatic impact on vision, medical imaging, and other fields.

1. SURFACES FOR EPIPOLAR-PLANE IMAGE ANALYSIS

The principal theme of this paper is the construction and use of a 3-D surface-building process. This process was developed to support our research in autonomous navigation, and was then found to have useful applications well beyond this area. The discussion begins by introducing the requirements of our autonomous navigation work driving the development of the surface-building process, and then explains what this process provides and how this was achieved. The following sections present examples of the application of the surface building to other domains – medicine, biology, scale-space studies, and higher dimension visualization.

1.1 Earlier Approach

Bolles, Baker and Marimont [1] describe a sequence analysis technique developed for use in obtaining depth estimates for points in a static scene. The approach bridged the usual dichotomy of depth sensing in that its large number of images led to a large baseline and thus high accuracy, while rapid image sampling gave minimal change from frame to frame, eliminating the correspondence problem. Rather than choosing quite disparate views and putting features into correspondence by stereo matching, with this technique we chose to process massive amounts of very similar data, but with much simpler and more robust techniques. The technique capitalized on the fact that the camera was moving along a linear path, viewing at right angles to this path, and acquiring images at equal spacing. These constraints enabled us to extract scan lines from successive frames and form space-time images where the motion of scene features would appear as linear tracks. Some simple calculations on these tracks gave us the positions of the features in the scene. We termed the space-time images *epipolar-plane images*, or EPIs, and the process *Epipolar-Plane Image Analysis*.

This research was supported by DARPA Contracts MDA 903-86-C-0084 and DACA 76-85-C-0004.

The restrictions on this straight-path orthogonal-viewing arrangement limited its applicability in autonomous navigation tasks: One might need to look where one is going, or pan and tilt to track some particular feature. It was essential that we generalize the analysis to allow this flexibility. In the generalization, we wanted to enable the process to:

1. Work for any camera attitude: allowing the vehicle to be looking along its direction of motion, or panning to track some particular feature as it moves across the scene.
2. Acquire images without dependence on rate of image capture or vehicle velocity.
3. Operate sequentially as the imagery is acquired, not demanding that all data be obtained before processing could be begun.
4. Directly provide spatially coherent results – three-dimensional object contours, not sets of isolated points.

These goals necessitated some major changes in our approach. Baker [2] presents the new approach in detail, and we summarize some of the elements here as they relate to the surface-reconstruction process.

1.2 Generalizing the Approach

In common with our earlier work, the new approach we developed involves the processing of a very large number of images acquired by a moving camera. The analysis is based on three constraints:

1. The camera's movement is restricted to a linear path.
2. The camera's position and attitude at each imaging site are known.
3. Image capture is rapid enough with respect to camera movement and scene scale to ensure that the data is, in general, temporally continuous.

The first and third are constraints shared with the earlier approach; the second is a significant relaxing of the earlier system's insistence on lateral viewing at fixed intervals.

Within this framework, we generalized from the traditional notion of epipolar *lines* to that of epipolar *planes* – a set of epipolar lines sharing the property of transitivity. These epipolar planes can have various attitudes in space-time. We then formulated a tracking process that exploited the above constraints in determining the position of features in the scene.

Visualizing this space-time approach may be difficult without an understanding of the geometry of the sensing situation: The basis of this was first presented by Gibson [3] nearly forty years ago. Figure 1 shows the geometry, indicating several imaging positions and attitudes along a straight path. We will refer often to this figure in our discussion. Notice that there are three camera positions, V_1, V_2 and V_3, along the camera path, a pencil of planes (a few representatives are indicated and labelled θ_1 through θ_{n+3}), and three rectangular imaging surfaces scribed with *epipolar lines.* The epipolar lines are formed by intersecting the imaging surfaces with the pencil of planes.

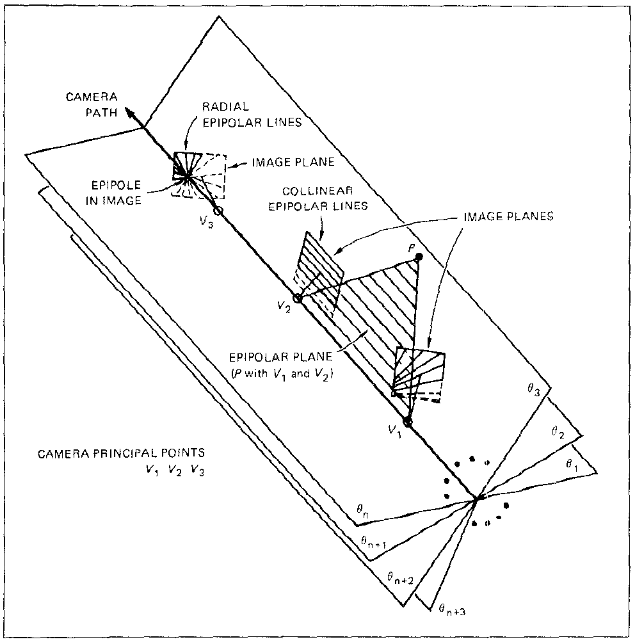

Figure 1. General Epipolar Configuration

In the case at V_2, the simplest case and that handled by our earlier analysis, we could form EPIs by stacking these lines. Feature paths in the EPIs were linear so estimation was simple. Unfortunately, in the situations demonstrated at V_1 and V_3, the structure of the intersections is radial, and not scan-line based. Even worse, feature paths in the planes will not be straight lines – they will be hyperbolic. The most complex situation occurs when the camera is allowed to change its direction while moving, varying say from V_1 to V_3 and positions between. Here, the structure of the planes depends on the direction of the camera at each point along its trajectory, and feature paths are neither linear nor hyperbolic but are arbitrary curves. These complications made our earlier approach – partitioning the data by EPI slices and tracking linear paths – inapplicable for more general motion. Our new approach is based on similar but more general principles.

1.3 Using the Constraints

At the outset, our goal was to determine the 3-D position of observed features, and we do this by tracking them through the spatiotemporal data. The estimation itself was facilitated by working in a space of line-of-sight vectors [4] – this made the analysis linear regardless of viewing direction or vehicle velocities. The tracking part of the job could be fairly simple, as temporal continuity generally ensures that observations of a single feature will form a continuous path in space-time. Still, we needed a way to constrain and follow the path. The epipolar-plane structure of Figure 1 would do this if we could restructure the data to that form. Using the geometry to resample to our earlier EPIs was a consideration, but the combination of singularities in the mapping, the computation required, the aliasing that would result, and the disruption of pixel variance measures, all made this a very unattractive option. Instead, we produce space-time elements with a 3-D generalization of 2-D edge contouring, transform these elements to the appropriate epipolar space of Figure 1, and do our observation grouping, tracking, and estimation in this new space. Our way of carrying this out brought us, at the same time, attainment of our other two goals: processing the images sequentially, and producing spatially coherent results.

This solution has been achieved through the design and implementation of a 3-D surface constructor, the *Weaving Wall.* It builds decriptions of the evolution of scene features over time. Using these surfaces and the constraints provided by knowledge of the camera path, we have been able to achieve our goal of robust tracking of scene contours for arbitrary viewing directions. This surface description is a crucial tool in the generalization of our approach, and we describe elsewhere how it will allow us to carry the analysis over to nonlinear paths and to solve for the path [2]. The discussion of the tracking work here, however, is limited to known straight-line camera motions.

1.4 Structuring the Data – Spatiotemporal Connectivity

The need for maintaining spatial connectivity can be observed by viewing our earlier results [1]. There, in processing the EPIs independently, we obtained separate planes of results. Wishing to exploit the fact that there should be some spatial coherence between these sets of points, we used proximity of the resulting estimates on adjacent planes to filter outliers. Features not within the error (covariance) ellipses of those above or below them were discarded. The remaining points, although reliable, formed a sparse and fragmented set. The problem, however, did not lie with this filtering, but with the loss of spatial connectivity in the first place. Our separation of the data into EPIs and then subsequent independent processing of these lost the spatial connectivity clearly apparent in the original images. We maintained instead the temporal connectivity that was critical to the feature tracking. For spatial connectivity in scene reconstruction, spatial connectivity in the imagery must be preserved. With both temporal and spatial connectivity vital to our analysis, we needed a process that could obtain these three-dimensional descriptions from the sequence data. The next section turns to the development of this 3-D surface-building process. Section 3 returns to its use in tracking and scene reconstruction. Finally, Section 4 presents results from other areas in which we have applied the surface weaver.

2. SURFACES IN VOLUMETRIC DATA

Our data are similar to tomographic medical data, and we looked to the medical imaging field for a process to build the needed three-dimensional surfaces. However, the work we found there had not been done in a manner that could be effective for our needs. Artzy *et al.* [5] describe what has become the principal approach to surface reconstruction from sensed tomographic-style data; it was, in fact, the *only* surface constructor of note. In their approach, which was developed in the context of surface building from CT data, surfaces are built consecutively, with each constructed by a sequential process that begins at a selected seed voxel and traverses the isocontour by means of a connected-component search. This is called the *cuberille model* of surface reconstruction. Wyvill [6] improved on the search efficiency by characterizing local surface structure, but maintained the technique's inherent sequential nature.

2.1 Design Objectives

Although the description we seek is closely related to this, our requirements demanded a very different approach:

1. *All* the surfaces in the sequence are of interest to our tracker, not just selected ones, and we need to follow them all.
2. Our aim of autonomous navigation through sequential processing means that we would never expect or desire to have all the data available in advance – our data (images from a camera) are obtained as we move, and must be processed as they are acquired.
3. A recursive traversal of individual surfaces is not appropriate, nor is processing the surfaces separately in sequence – all surfaces must be built incrementally and in parallel as each new image is acquired.
4. Integral positioning of facets is insufficiently precise – for accurate mapping we need sub-pixel resolution in surface definition.
5. Simple density or intensity themselves are inappropriate for surface definition – we track the evolution of image features over time, and define our features, and hence our surfaces, as being located at the zeros of Laplacians of a Gaussian.
6. We will have need of various sorts of computation on the local surface as it is being constructed, and the organization of the process must support this.

Each of these design objectives differs in important ways from Artzy *et al.*, but the one most distinguishing of our approach is point 3. Both sequential processing and a capability for parallel implementation are crucial for a realistic tracking system.

2.2 Constructing Surfaces

The surface builder we have developed operates on images sequentially as they are acquired, knitting together a connected representation of the spatial and temporal evolution of the sequence over time. It maintains the continuity of feature paths irrespective of changes in viewing angle, or a varying rate of image acquisition. It's sole requirement of the data is that between frames, surfaces move in some direction no more than their width.[1] The processor acts as a *loom* during surface construction, with a wall of accumulators meeting each image in sequence and weaving its elements into the mesh of surfaces prepared behind. From this action we give it its name, the Weaving Wall. Although currently implemented to work sequentially within the first two dimensions (as it must on a sequential machine), the process could be recoded to operate in parallel over the spatial images, and run in real time.

Since the principal reason for developing the Weaving Wall was for use in motion sequence analysis, we will develop its operation in the context of Laplacians. For additional applications that we will touch on later, other measures may be more appropriate – for example, we use Hounsfield density when we process CT data, and track not density zeros, but zeros with respect to a bias, the chosen surface density value. But the distinction is only incidental to the development, and we will work for now with the Laplacians.

2.3 Local Surface Elements from the Volumetric Data

We form a three-dimensional data set by treating the collection of images as a set of arrays of voxels each one pixel deep. In three dimensions, zero-crossings will form surfaces enclosing *volumes* composed of voxels having Laplacian values of the same sign. The zero-crossing surface is composed of three types of voxel facets: U and V facets separate voxels horizontally and vertically, while T facets separate voxels temporally – they occur between voxels at identical positions in adjacent images. Interpolation allows us to position the surface elements with sub-voxel precision. This defines our surface elements; we now demonstrate that we can obtain them from the sequence data.

Given the local nature of our facet definition – facets are positioned between adjacent voxels based solely on their Laplacian values – it would seem that their detection should be attainable with very simple tests applied to only local parts of the data set. In fact, everything necessary for determining the local surface structure is available in each $2 \times 2 \times 2$ window of the data. Consider a $2 \times 2 \times 2$ set of voxels containing Laplacian values. There are 2^8 different combinations of positive and nonpositive Laplacian values in these 8 voxels. If some of these voxels are of different sign, then we have a surface (or several surfaces) passing through this $2 \times 2 \times 2$ part of the data set.[2] The eight-bit signature formed by the eight *on/off* states of the eight voxels specifies exactly the local surface structure.[3] The weaving process computes these local surface structures – connected arrangements of U, V, and T facets – and connects them with all of their appropriate neighbors. Doing this $2 \times 2 \times 2$ operation throughout the entire volume gives us a description of all the surfaces in the data set. Baker [9] details the principal elements of this construction.

Figure 2 shows a wire-frame description of a surface formed by this process. The surface facets are positioned at the mesh intersections – these are determined by the interpolated zeros of the three-dimensional Laplacian of the Gaussian. Their three orientation components are determined by the values of the three-dimensional gradients of the Gaussian.

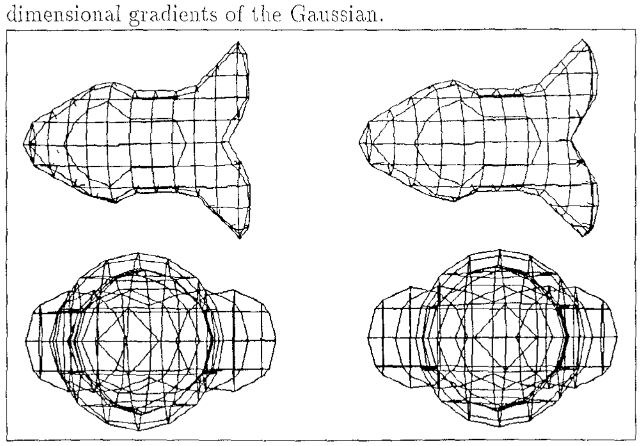

Figure 2. Surface Mesh: top view above, front view below

3. TRACKING ON THE SPATIOTEMPORAL SURFACE

The features we track are zeros of a 3-D Laplacian, and bear (u, v, t) coordinates; they are spatiotemporal *voxel facets*. Figure 5 shows a mesh description of the facets for the spatiotemporal surfaces associated with the forward-viewing sequence whose first and last images are depicted in Figure 3. In the interest of clarity, the surface representations with which we currently work are based on a simplified version of this imagery – one eighth the linear resolution of the originals. Figure 4 shows these two frames at the reduced resolution. It is important to understand that these surfaces are determined in (u, v, t) image space – it requires the geometric parameters of the camera and its path to enable motion-dependent interpretation of these surfaces.

[1] Otherwise they would be disjoint, and form two volumes, not one.

[2] This approach is an extension of a 2-D contour finder having a 4-bit index developed by Marimont [7].

[3] Lorensen *et al.* [8] describe an independently developed algorithm for surface rendering based on the same use of voxel sign signatures.

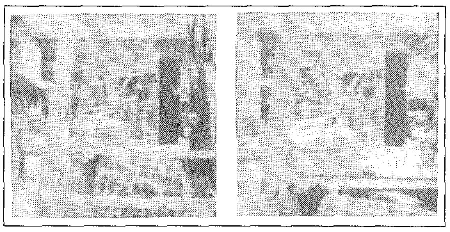

Figure 3. Sequence 1^{st} and 128^{th} Images

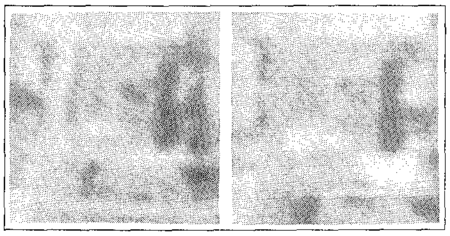

Figure 4. 1^{st} and 128^{th} Images at $\frac{1}{8}$ Resolution

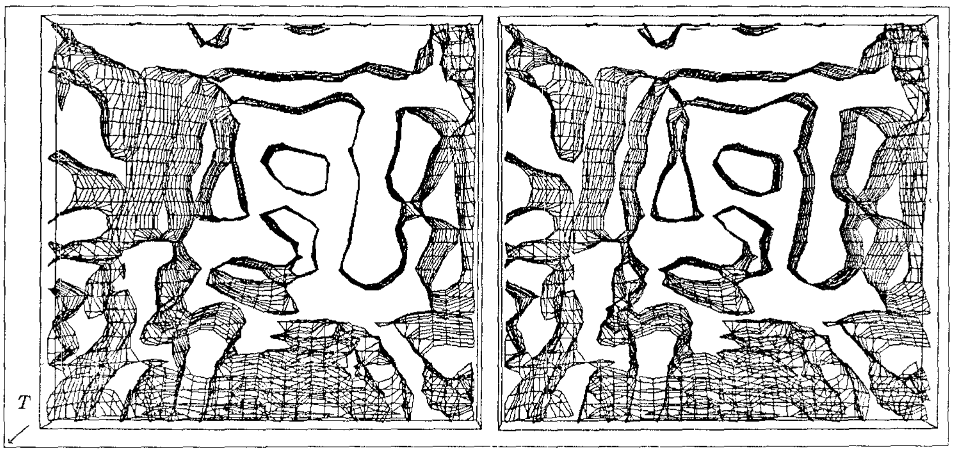

Figure 5. Spatiotemporal Surface Representation, First 10 Frames

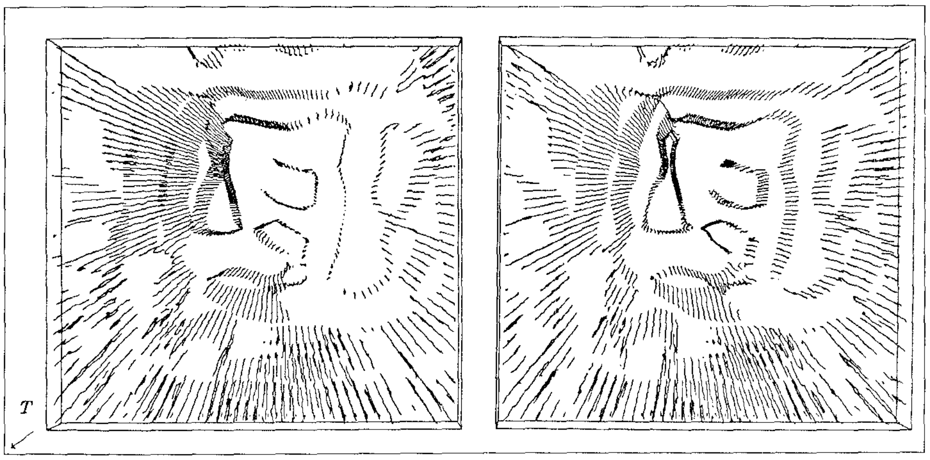

Figure 6. Epipolar-Plane Surface Representation

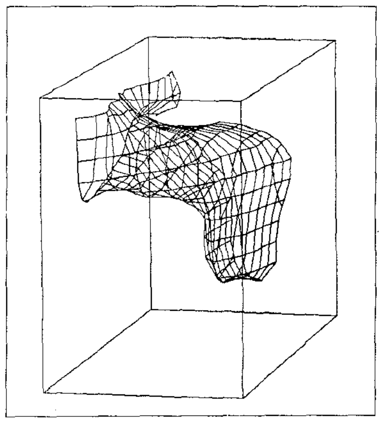

Figure 8. Spatiotemporal Surface

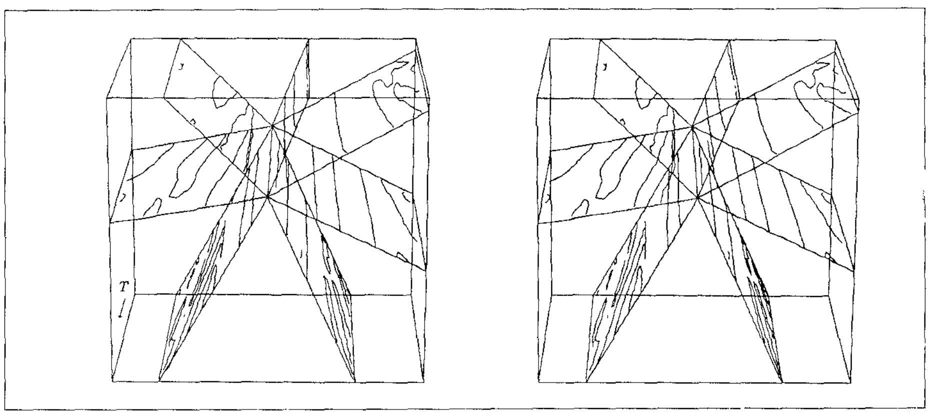

Figure 7. Intersection of 7 Epipolar Planes with Spatiotemporal Surfaces (30 frames)

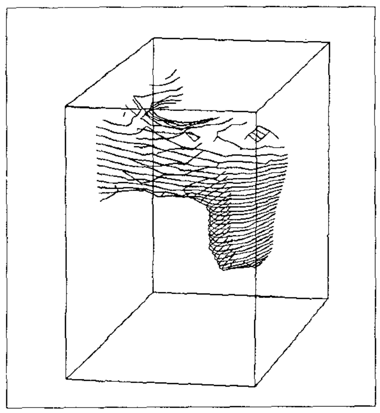

Figure 9. Epipolar-Plane Representation

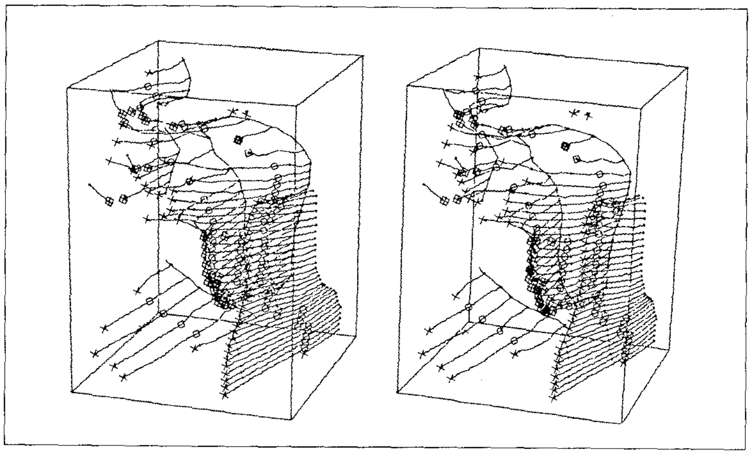

Figure 10. Sequential Feature Tracking on the Spatiotemporal Surface

3.1 Epipolar Structuring, Tracking and Estimation

For nonorthogonal viewing directions, epipolar lines are not distinguished by the spatial v scan-line coordinate. To obtain this necessary structuring we develop within our spatiotemporal description an *embedded* representation that makes the epipolar organization explicit. Over each of the sequential images, we transform the (u, v, t) coordinates of our spatiotemporal zeros to (r, h, θ) *cylindrical coordinates* (θ indicates the epipolar-plane angle ($\theta \in [0, 2\pi]$), the quantized resolution in θ is a supplied parameter, and the transform for each image is determined by the particular camera parameters). In this new coordinate system, we build a structure similar to our earlier EPI edge contours, but dynamically organized by epipolar plane. This is done by intersecting the spatiotemporal surfaces with the pencil of appropriate epipolar planes. We weave the epipolar connectivity through the spatiotemporal volume, following the known camera viewing direction changes. Figure 6 shows a sampling of the spatiotemporal surfaces as they intersect the pencil of epipolar planes (every fifth plane is depicted). You will notice the obvious radial flow pattern away from the epipole (FOE). Figure 7 shows seven of these surface/plane intersections, along with the associated bounding planes (refer to Figure 1). The edge that all share is the camera path (the locus of the epipole). Figure 8 isolates a single surface from the top left of Figure 5, and shows it's spatiotemporal structure. Figure 9 shows the same surface structured by its epipolar-plane components. Baker [9] presents details of this intersection operation on the spatiotemporal surface.

Figure 10 shows the tracking of scene features on the spatiotemporal surfaces in the vicinity of the surface of Figure 8. The coding is as follows: the leading observation of a feature (active front) is shown as an $\times$; lines join feature observations; 5 observations must be acquired before an estimate is made of the feature's position – at that point an initial batch estimate is made, and a Kalman filter [10] is turned on and associated with the feature – this initiation of a Kalman filter is coded by a square; where two observations merge or observations begin to deviate from the linear model, the tracking is stopped and the features are entered into the database – this is coded by a diamond. Individual trackers estimate the positions of individual scene features: the spatial part of the space-time surface enables composing these feature estimates into estimates of scene contours. Baker [2] presents details of this tracking and reconstruction process.

4. WEAVING WALL APPLICATIONS

Some exciting secondary applications of this research arose from the development of the Weaving Wall spatiotemporal surface-building process. This process was designed to satisfy our needs for spatially-coherent incremental tracking over an image sequence, and these characteristics make it particularly useful for other applications where coherent description of nearly continuous three-dimensional data is sought. Most obvious among these is the construction of surface models from CT (computed tomography) and other tomographic (*eg.*, magnetic resonance, ultrasound, electron microscope) medical data. We have also been applying it to visualization problems in studying the behaviour both of images with respect to spatial resolution and of analytic functions of dimension higher than three.

4.1 Medical Computer-Tomography Data

In applying the surface-building algorithm to surface reconstruction from CT slice data – data whose components are all spatial – we currently use tissue density rather than Laplacian values.[4] Figure 11 shows the evolution of surfaces judged to be 'bone' (greater than 240 units) in a 70×30 window of a 52-image CT data set. You can see the incremental nature of the surface development. Figure 12 shows two rendered views of this spine. Figure 13 shows a stereo pair of the bone from the full data set. Since these surfaces are distinct objects, they may be manipulated for many purposes – for example, simulation of kinematics and dynamics as shown in Figure 14, and structural analysis.

Figure 15 presents another CT case study, with this showing stereo views of the skin, skull, and cranial vault. The patient has been undergoing cranial reconstruction. These are from 46 slices of roughly 120×120 CT images. Interslice spacing is five times the intraslice resolution. Sampling density was increased for a section of the data set in the area about the ear, and it would appear from the horizontal crease there that the patient moved slightly as the scan adjustments were being made. There was motion also in the area of the jaw, and X-ray reflections from metal teeth fillings.

[4] We do, however, calculate the 3-D Gaussian gradients for determining local surface orientation.

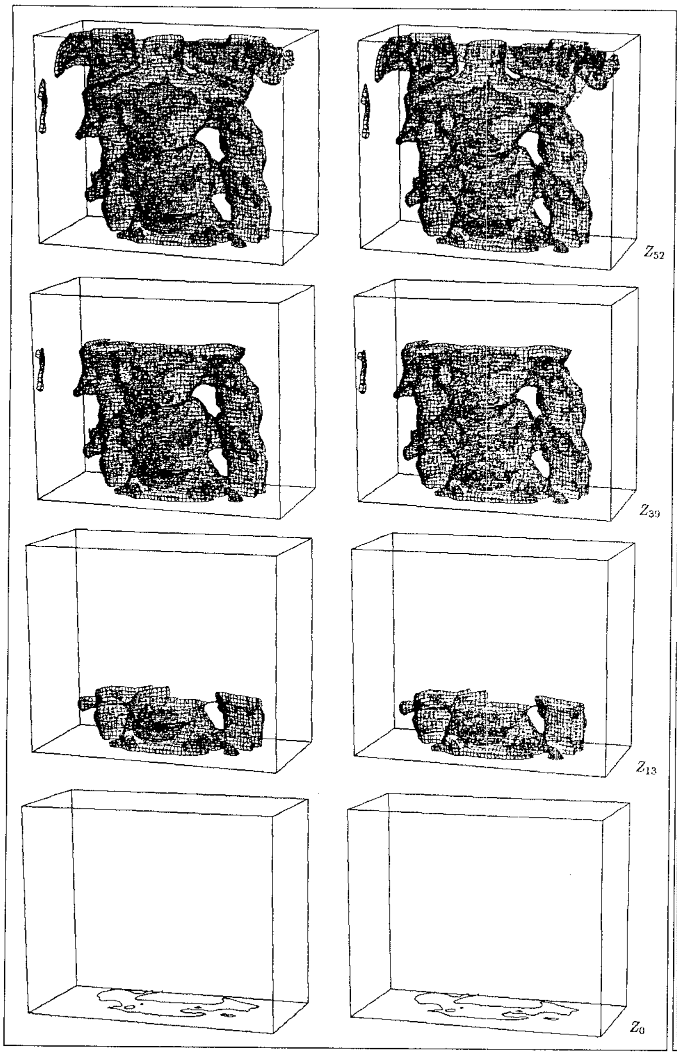

Figure 11. Slice Evolution of Spine

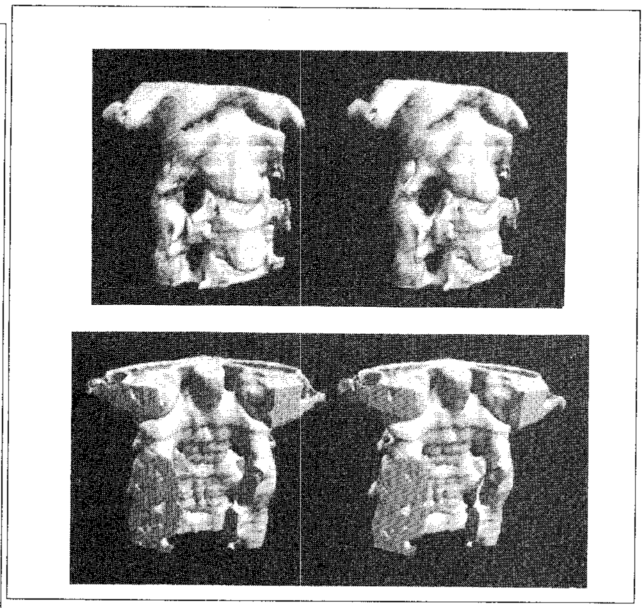
Figure 12. Rendered Spine Views

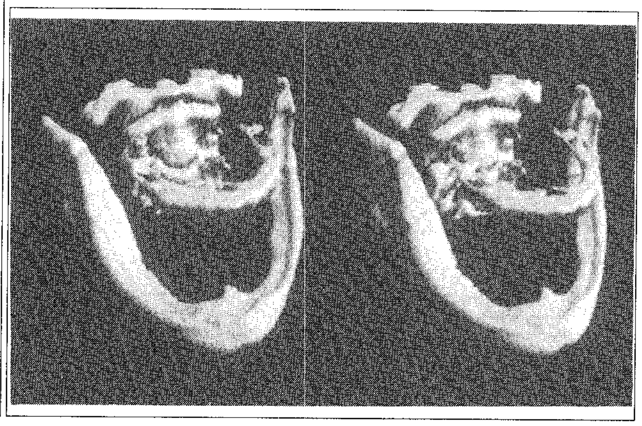
Fig. 13. Display of Three Major Bone Surfaces

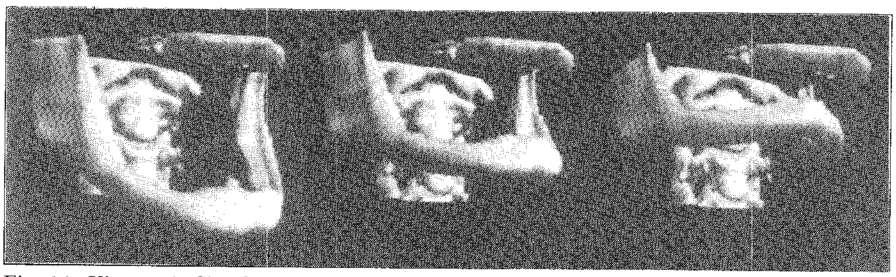
Fig. 14. Kinematic Simulation

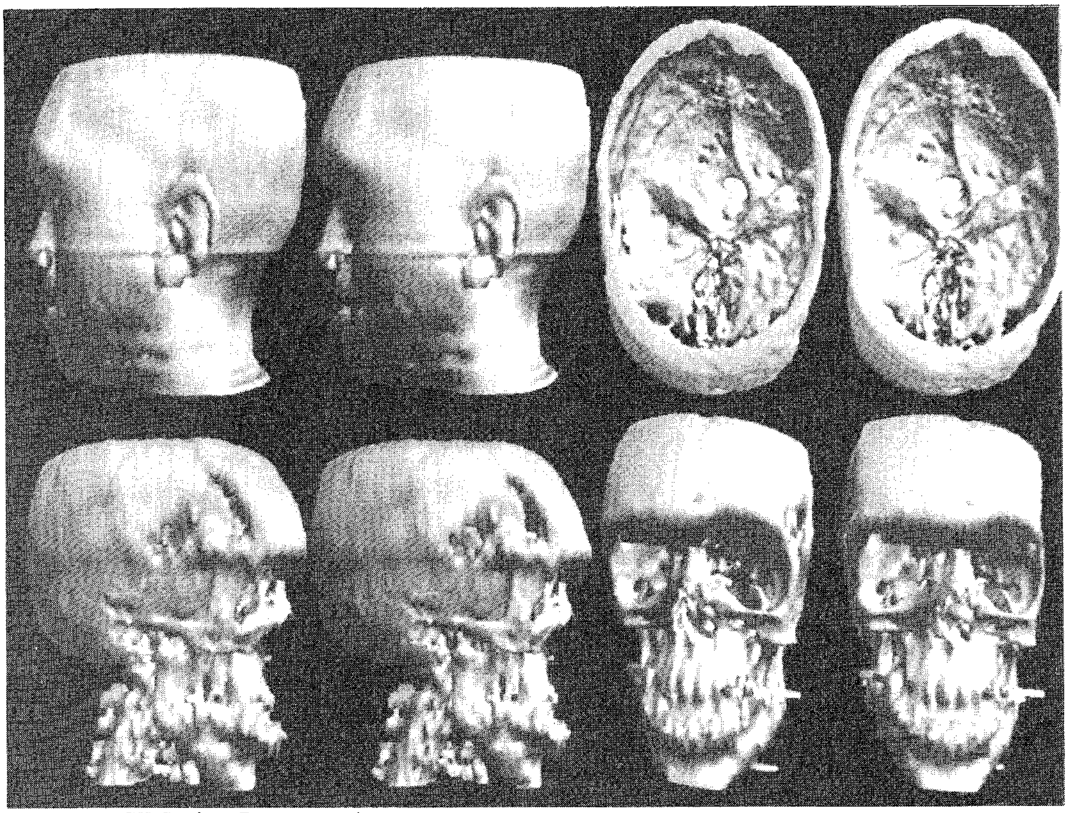
Figure 15. CT Surface Reconstructions

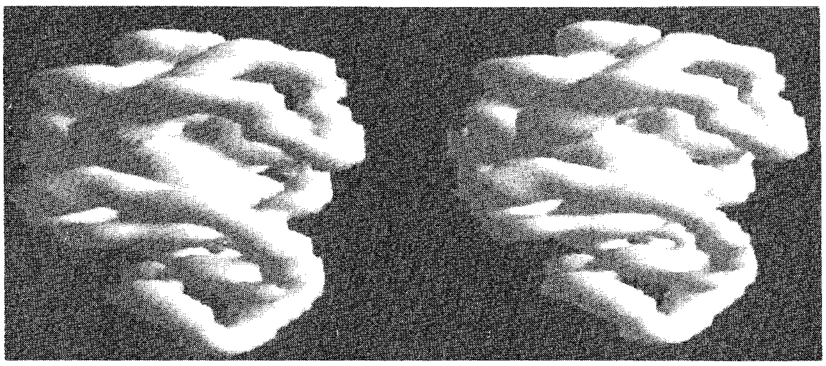
Figure 16. Prophase Chromosomes.

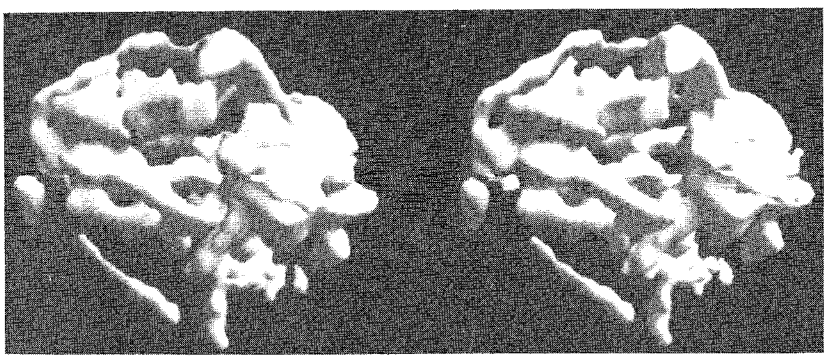
Figure 17. Early Anaphase Mitotic Chromosomes

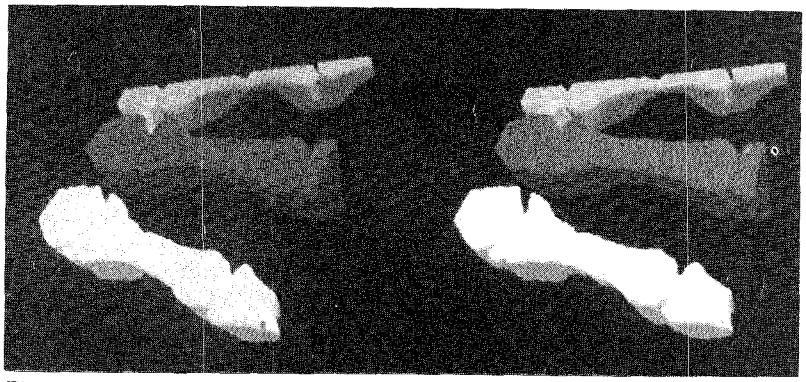

Figure 18. Hand Bone from Ultrasound Sequence

4.2 Other Tomographic Data

Coupling electron microscopy with image-enhancement techniques, Agard and his colleagues at UCSF have been investigating cell three-dimensional structure [11]. Their data collection process enables them to form tomographic images through cell specimens. To aid them in this study, we have used the Weaving Wall to reconstruct chromosomes' 3-D form. Figure 16 shows a stereo pair of *Drosophila melanogaster* chromosomes at prophase, before cell division. Notice the coiling. Figure 17 shows polytene chromosomes at early anaphase while undergoing mitotic division. These surface descriptions represent the beginnings of a new tool for structural investigation in molecular biology. We will be working with UCSF in further development of this. Ultrasound data is similarly tomographic in nature. The Bioengineering Laboratory here at SRI has been a leader in the development of this clinical tool. Image-enhancement techniques are involved in producing tomographic imagery here as well. Figure 18 shows a reconstruction of the bones in a hand from a ten-image ultrasound data set.

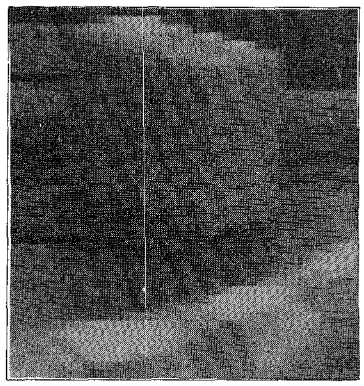

Figure 19. First Image in Scale-Space Hierarchy

Although the display aspects of this surface-building technique are evidently quite worthwhile, bear in mind that the primary representation is a surface model, with all the connectivity appropriate for full model-based computation (*eg.*, finite element analysis, symmetry mappings, elastic deformation operations, chromosome unravelling, et cetera).

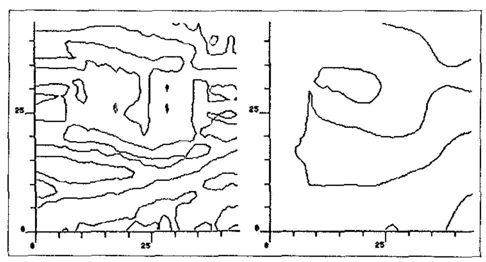

Figure 20. 1^{st} and 8^{th} Zeros

4.3 Images in Scale-Space

One of the unresolved issues in the motion sequence analysis work, and in computer vision in general, is the selection of a Gaussian to be the basis for detection operations; in effect, selecting the scale of analysis. In researching this issue, we have done some limited experimentation with surface building where the third dimension is Gaussian scale (σ). Building a surface of the evolution of an image as its resolution varies provides a vivid picture of how Witkin's 1-D scale-space studies [12] can extend to images. Figure 21 shows a surface constructed at the 3-D Laplacian zeros obtained by applying a battery of increasingly larger Gaussians to the image of Figure 19. This shows eight images, each differing in σ by 0.5 from the one before. Figure 20 left shows the zero-crossings of the smallest Gaussian, and its right shows the zero-crossings of the largest Gaussian – these are the extremes of which Figure 21 represents the continuum. The most stable representation of a feature in this space may be at that part of its evolution exhibiting minimum spatial velocity with respect to $\Delta\sigma$. The connected scale-space surface makes this stability explicit.

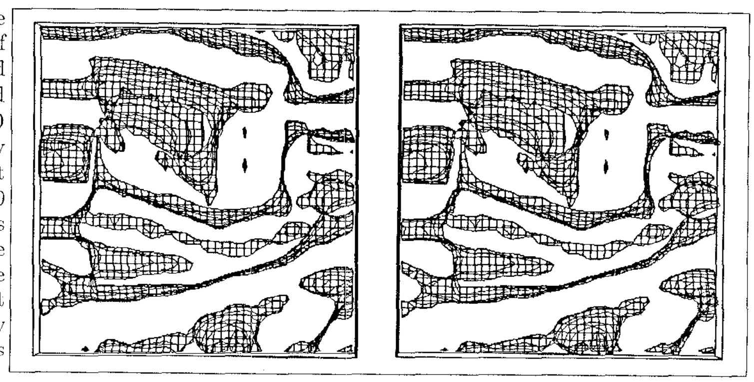

Figure 21. Scale-Space Surface

We will soon be modifying the Weaving Wall to produce four-dimensional surfaces, where the first three are the spatial and temporal as before, and the fourth is Gaussian scale. Our intention is to use the most stable representation of a feature as its instantiation to be tracked. The linear estimators will then use these more appropriate σ values in determining observation weights and in estimating the resulting spatial precisions. Tracking will be occurring at all scales at once. In other applications, such as the processing of magnetic resonance imagery, we will be attempting to use the stability of a contour in this scale dimension as a measure to combine with gradient strength in estimating both contour scale and significance. Higher dimension surface descriptions will also have importance in medical imaging in general: We plan to use the fourth dimension in modelling temporal change, as in hearts beating and vertebrae rotating.

4.4 Analytic Functions

The Weaving Wall has proved useful in object representation studies. Our research group explores issues of representation for object modelling, and has developed a representational facility based on superquadrics [13]. Hanson [14] has developed a hyperquadric generalization permitting the description of shapes with arbitrary polyhedral bounds; superquadrics, in contrast, permit only shapes having the three orthogonal Cartesian bounds. Use of these hyperquadrics enriches the modelling by yielding a superset of the superquadric primitives. Experimenting with these higher dimension objects is complicated by the difficulty in visualizing them. To facilitate this, we display sequences of three-dimensional projections of these n-dimensional objects ($n > 3$), using the Weaving Wall in this rendering. Figure 22 shows a selection of such projections through a four-dimensional surface. Viewing such frames as a sequence gives us insight into the structure of these objects. Banchoff [15] discusses similar issues in assessing the structure of higher dimension functions.

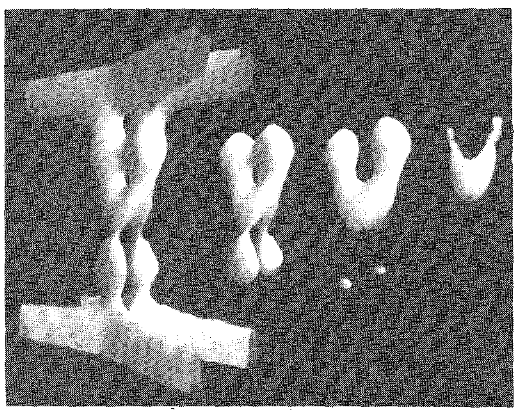

Figure 22. Four 3-D Projections through 4-D Surface

Other applications of the surface-building process described here include representation of surfaces from other two-dimensional sensing domains (*eg.*, geology), display of surface fracture reconstructions (2-D over time), and the colorization of black and white film. In general, this can be used in any application where a depiction or computational representation is desired of the evolution of a two-dimensional pattern varying in a continuous manner over a third dimension – be it time, space, viewing position, resolution, or any other dimension.

5. CONCLUSIONS

The development of the Weaving Wall surface-building process was a necessity for our further progress in sequence analysis. The structures it produces provide for our current needs in tracking features over the full range of possible camera attitudes, and they also seem to be most appropriate for several problems that lie ahead: camera solving and extension of the technique to nonlinear camera paths. We believe that this explicit use of space-time structure will prove to be a significant development for robotic vision. With applications in medical and biomedical imaging and other areas, we feel the surface building process will be seen to be a general and powerful tool for many applications in image processing.

The current implementation, running on a Symbolics 3600, processes the spatiotemporal surfaces at a 1KHz voxel rate, with the associated intersecting, tracking, and estimation procedures bringing this down to about 150Hz, 75% of which is consumed in the surface intersection (the surface intersection would not be required if we had a sensor of the appropriate geometry). Both the feature tracking and the surface construction computations are well suited to MIMD (or perhaps SIMD) parallel implementation. With these considerations, and the process's inherent precision and robustness, spatiotemporal-surface-based epipolar-plane image analysis shows great promise for tasks in real-time autonomous navigation and mapping. Benefit will equally be found in real-time surface building for the other applications we have discussed.

Acknowledgements

Some of this work has been done in collaboration with Bob Bolles. David Marimont, currently with Xerox PARC, has been crucial in the development, providing insights for both the geometry and mathematics of the tracking process, the design of the surface builder, and the approach taken to higher dimension convolution. Lynn Quam has provided excellent image manipulation and graphics tools, and, with Alex Pentland, was very helpful in the production of the rendered surface displays. Andy Hanson provided the data and motivation for the hyperquadric analytic surface depictions. The medical CT data is courtesy of Cort Cuttings M.D., New York University, and CEMAX Corporation, Santa Clara, California; the electron-microscopy data was provided by David Agard and John Sedat of the Department of Biochemistry and Biophysics at the University of California at San Francisco; and the ultrasound data was courtesy of Philip Green of the Bioengineering Research Laboratory here at SRI.

References

[1] R. C. Bolles, H. H. Baker, and D. H. Marimont, "Epipolar-Plane Image Analysis: An Approach to Determining Structure from Motion," *International Journal of Computer Vision,* Kluwer Academic Publishers, Vol.1, No.1, June 1987, 7–55.

[2] H. H. Baker and R. C. Bolles, "Generalizing Epipolar-Plane Image Analysis on the Spatiotemporal Surface," in press, *International Journal of Computer Vision,* Kluwer Academic Publishers, Vol.2, No.4, 1988.

[3] **The Perception of The Visual World**, J. J. Gibson, Houghton Mifflin, Boston, 1950.

[4] D. H. Marimont, "Projective Duality and the Analysis of Image Sequences," *Proceedings of the Workshop on Motion: Representation and Analysis*, IEEE Computer Society, May 1986, 7–14.

[5] E. Artzy, G. Frieder, G. T. Herman, "The Theory, Design, Implementation, and Evaluation of a Three-Dimensional Surface Detection Algorithm," *Computer Graphics and Image Processing*, 1981, 15, 1–24.

[6] G. Wyvill, C. McPheeters, and B. Wyvill, "Data Structure for Soft Objects," *The Visual Computer*, (1986) 2: 227–234.

[7] D. H. Marimont, "Segmentation in Acronym," *Proceedings of the DARPA Image Understanding Workshop*, September 1982, 223–229; amplified by personal communication and code.

[8] W. E. Lorensen and H. E. Cline, "Marching Cubes: A High Resolution 3D Surface Construction Algorithm," *Computer Graphics*, Vol.21, No.4, July 1987, 163–168.

[9] H. H. Baker, "Building Surfaces of Evolution: The Weaving Wall," in press, *International Journal of Computer Vision,* Kluwer Academic Publishers, Vol.2, No.4, 1988.

[10] **Applied Optimal Estimation**, A. Gelb, editor, The Analytic Sciences Corporation, MIT Press, Cambridge, Massachusetts, 1974.

[11] D. A. Agard, "Optical Sectioning Microscopy: Cellular Architecture in Three Dimensions," *Annual Reviews of Biophyshics and Bioengineering,* 1984, 13, 191–219.

[12] A. P. Witkin, "Scale Space Filtering," *Proceedings of the Eighth International Joint Conference on Artificial Intelligence (IJCAI-83),* August 1983, 1019–1021.

[13] A. P. Pentland, "Perceptual Organization and the Representation of Natural Form," *Artificial Intelligence*, **28**, 1986, 293–331.

[14] A. J. Hanson, "Hyperquadrics: Smoothly Deformable Shapes with Convex Polyhedral Bounds," in press, *Computer Vision, Graphics, and Image Processing*, 1988.

[15] T. F. Banchoff, "Visualizing Two-Dimensional Phenomena in Four-Dimensional Space: A Computer Graphics Approach," in **Statistical Image Processing and Computer Graphics,** E. Wegman and D. Priest, editors, Marcel Dekker, Inc., New York, (1986) 187–202.