Analysis of Three-Dimensional Image Data: Display and Feature Tracking

Zvi Kam, Hans Chen, John W. Sedat, and David A. Agard

1. Introduction ................................................................. 237
2. The Image Display and Processing Environment .................. 239
3. Processing Data for Display ............................................. 243
4. Image Modeling, Analysis, and Quantitation ....................... 245
   4.1. Interactive 3D Model Building Using PRISM ................. 245
   4.2. Computerized Recognition and Tracking of Objects .......... 247
5. Summary ........................................................................ 255
References ........................................................................... 256

1. INTRODUCTION

Microscopes have been used for almost two centuries to produce a wealth of biological information based on direct observation and photography. Recent advancements in digital image processing and display have created a revolution in biological and medical imaging. Digital methods have lead to significant improvements in the ease of image acquisition and data quality for both light and electron microscopy. It is now possible to routinely obtain quantitative image data and not simply pictures. The advent of quantitative microscopy is a major step forward in the application, and utilization of structural information for biological problems at the level of cells and cell components. Perhaps most significant has been the possibility of examining complex, noncrystalline objects such as supra-
molecular assemblies in three dimensions. Combined with powerful new probes to examine specific molecular components, three-dimensional imaging can provide new insights into the spatial localization patterns of specific molecules and their redistribution, for example during the cell cycle or development.

Electron microscope tomography methods aim to fill the very wide gap between atomic structures defined by x-ray crystallography and the more global patterns of organization observed in the light microscope. The goals of tomographic projects range from reconstructing the three-dimensional surfaces of small particles, such as ribosomal subunits, to examining much larger structures, such as cilia, chromosomes, or even entire neurons. Unlike more conventional structural methods that implicitly depend on symmetry or averaging over a very large number of examples, microscopic imaging methods typically examine individual structures. As a consequence, there can be considerable variation in the details for each specimen examined. This variability results from intrinsic variations in the structures themselves, and variability resulting from inadequate preservation, heterogeneity in staining, and electron beam damage. The smaller, particulate structures such as ribosomal subunits are undoubtedly considerably more homogeneous than much larger structures such as chromosomes. Such structural variability extends to the organellar level where the details of Golgi or endoplasmic reticulum structure are dynamic and hence not precisely reproducible.

Much of the challenge in structural cell biology is to be able to extract biologically relevant structural information from the vast amount of data present in a three-dimensional reconstruction. For many cell biological problems (using either EM tomography or three-dimensional light microscopy), what is desired is the elucidation of general patterns of three-dimensional organization. Depending upon the problem, accomplishing this goal may require analysis at many different levels. In some cases, being able to trace the paths of structures in three dimensions is the best way to extract the relevant information. For other problems, different forms of semiquantitative analysis such as surface area or volumetric measurements, pairwise distance measurements, etc., may be the most important. While for still other problems a qualitative examination may be all that is required; thus, simply providing stereoscopic images may be sufficient.

In this chapter, we will emphasize approaches that provide the capability for a quantitative analysis of three-dimensional volumetric data in addition to the more conventional display aspects. A direct method of quantitatively analyzing the complex three-dimensional images typical of biological structures is to first identify the structural component of interest using continuous intensity images and to then construct graphical objects or models depicting features of interest. The kind of modeling required will vary in its geometrical (topological) nature, depending on the biological problem being investigated. The structure of chromosomes and the tracking of cytoskeletal structures can be well represented by 1D (linear or branched) tracks in 3D images. The topology of membranes (nuclear envelope, cells plasma membrane) and the definition of inside/outside relations of a 3D body require the modeling of 2D surfaces, whereas tracking 3D structures in time requires analysis of 1D or 2D elements within the context of a 4D space. Once a simplified representation of the volumetric data is developed, it can be used directly to guide any desired form of quantitative analysis. Discerning geometrical or topological
patterns within complex structures is greatly simplified by having first abstracted the relevant information in the form of a model. In addition, once defined, models can be used to trace intensity profiles along the model or to directly unfold the image intensity in the region surrounding the model and to display it stretched out. Another very important use of models is to facilitate detailed correlations between features from one specimen to another.

Despite the immense progress in interactive computer power and graphics workstations, and the associated development of artificial intelligence and expert systems for machine vision and image understanding, our experience with structural studies of chromatin suggests that the complexity of biological images and the variability of the structural features studied represents a new class of problems that have only begun to be investigated. It has been our experience, that in general simple three-dimensional display approaches such as the three-dimensional iso-surface cage contours routinely used by the x-ray crystallographers (FRODO; Jones, 1982) or single-contour surface rendering approaches are unable to convey the richness and subtleties inherent in biological three-dimensional reconstruction data. As discussed below and elsewhere in this volume, the use of projection methods that maintain the gray-level density representation provide a powerful and versatile alternative to contouring approaches. Unfortunately, it has been our experience that many biological problems, especially those resulting from EM tomographic reconstructions of large, nonparticulate objects, are too complex to be readily visualized in their entirety. Thus, we think that it is necessary to develop an environment for the interactive display and analysis of complex volumetric image data that allows a variety of visualization methods to be used in concert on interactively chosen subregions of the object being displayed in arbitrary orientations. Furthermore, it is important to design this visualization environment from the outset with the capability for interactive or semiautomated three-dimensional model building. In this chapter we discuss our efforts to develop such a display environment and modeling capability. In addition, we will cover the work of others and ourselves on the challenging problem of automated or semiautomated modeling.

2. THE IMAGE DISPLAY AND PROCESSING ENVIRONMENT

In our laboratory, we are interested in understanding higher-order chromosome structure and the spatial and temporal dynamics of three-dimensional chromosome organization throughout the mitotic cell cycle. Toward this end, we have been developing data collection and reconstruction methods for both IVEM tomography as well as optical sectioning light microscopy. Due to the difference in resolution, nature of the optics, and sources of contrast, electron and light microscopes (EM and LM) are different, but complementary in their structural information contents. In spite of the differences, there are many commonalities in the analysis of both data types. As a necessary part of our overall structural effort, we have also been developing a general purpose display environment and structural analytic methods designed for the examination of complex cellular structures derived from both light and electron microscopic reconstructions. The key features are (i) an image display
environment able to interactively manipulate and display multiple sets of three-dimensional data. Using a multiple-window approach, various combinations of section data and projection data (either raw or image processed) can be viewed simultaneously, and the windowing system keeps track of the geometrical relationship between images; (ii) interactive 3D model building using cursor-controlled graphics overlayed on the original or preprocessed data. Modeling can be performed in multiple windows which greatly facilitates tracing paths in complex objects; and (iii) semiautomatic and automatic modeling based on algorithms which utilize the characteristic features of the structures under study. In this and following sections we will describe this image display environment and modeling methods in more detail.

Before proceeding to discuss the display system, we give a brief overview of important features of the data collection and handling. Regardless of modality (electron microscope tomography or optical section microscopy), high resolution, quantitative three-dimensional microscopy requires the best possible data. We have taken great care in the design of our microscopic imaging systems which employ cooled, scientific-grade charge-coupled device (CCD) cameras having a wide dynamic range (12 bits), exceptional linearity, and geometric stability (Hiraoka et al., 1987). The CCD systems on either the light microscope or our Philips EM430 IVEM are built around compact, MicroVaxIII workstations that use a high-performance array processor (Mercury) and Parallax display system. All aspects of microscope operation are computer controlled. The IVEM system employs a 20-μm-thick YAG single-crystal scintillator (Gatan) directly coupled to a 1024 × 1024 pixel CCD by fiber optics. The array processor permits on-the-fly gain and offset corrections for the individual pixels. Rapid image processing or Fourier transform calculations are useful for on-line critical evaluation of microscope focus or aberrations. For the EM system, we have a computer-controlled ultrahigh-tilt stage (±80°), and complete computer control of x, y, z motor stages, image and beam deflection, and lens excitation parameters. For the light microscope we have employed computer-controlled barrier and excitation filters and double- or triple-wavelength dichroic mirrors to provide rapid wavelength switching free from image translation artifacts. Such problems are inevitable if dichroic mirrors have to be switched. Multicolor data acquisition has been employed in fixed samples and recently also in live developing embryos, which were co-injected with two molecules labeled with fluorescein and rhodamine. This enables the spatial relationships and interactions of several molecular components to be examined both as a static and as dynamic processes changing through time.

In addition to attempting to optimize the quality of the three-dimensional data collected, we have spent considerable effort on developing reconstruction algorithms for both the optical sectioning data (Agard et al., 1989) and for EM tomography (Abbey and Agard, forthcoming). In general, these methods all share a common theme—they seek a solution which is everywhere positive, appropriately spatially bounded and that when projected (in the case of the EM) or convolved with the point-spread function (for the light microscope) will match the observed data. The use of such straightforward a priori information as constraints imposed during the iterations can have a profound effect on the stability of otherwise ill-posed reconstruction algorithms, both ameliorating the effects of missing data as well as conditioning against noise.
Convenient manipulation and trouble-free handling of a large number of three-dimensional data sets requires that the full volumetric data be kept within a single file, with relevant information maintained in a file header. For all of our studies we utilize images files written in the MRC format (for description see Chen et al., 1990). The file begins with a header that contains parameters specifying the dimensions, geometrical scales, origins, orientation, mode of storage (bytes, integer, real, complex), data type (sections, projections, tilt series in mono or stereo, multi-wavelength data, and time lapse 3D series), range of intensities, and text strings describing the file history. The header record, which can be extended for individual picture attributes, is followed by a direct access data structure which is easily used to store 2D images, stacks of serial sections of 3D data, time series of 3D data, and multiple color pictures, etc. A complete set of subroutines is available to facilitate file setup and manipulation. Importantly, the header contains sufficient information so that files generated from different geometric manipulations of the same data (e.g., translations, rotations, projections, subsamplings, magnification changes, etc.) can be related to one another so that models can be displayed in the correct perspective view.

Our three-dimensional imaging and modeling package (PRISM, Chen et al., 1990) employs a high-resolution (1280 × 1024, 60 Hz, noninterlaced) video board

<table>
<thead>
<tr>
<th>TABLE 1. Functions Available within the PRISM DISPLAY Program</th>
</tr>
</thead>
<tbody>
<tr>
<td>Display functions:</td>
</tr>
<tr>
<td>Load a series of images into a window.</td>
</tr>
<tr>
<td>Look-up table manipulations: linear, clipped and power LUTs.</td>
</tr>
<tr>
<td>Zoom, pan, move window position and size.</td>
</tr>
<tr>
<td>Manual paging of images associated with one or all windows.</td>
</tr>
<tr>
<td>Movie mode for one or more windows (like stereo pair movie).</td>
</tr>
<tr>
<td>Stereographics display (top-bottom stereo encoding to be used with the Stereographics polarizing screen on glasses)</td>
</tr>
<tr>
<td>Photography (computer-controlled exposure with film recorder) and low-resolution video recording for videotape and videodisk.</td>
</tr>
<tr>
<td>Image processing functions employing the MERCURY array processor:</td>
</tr>
<tr>
<td>Convolutions, Fourier transforms, and filtering.</td>
</tr>
<tr>
<td>Local contrast enhancement.</td>
</tr>
<tr>
<td>Interactive 3D projections and sections through a cube of data.</td>
</tr>
<tr>
<td>Modeling and presentation utilities:</td>
</tr>
<tr>
<td>Point: interactive, mouse-selected set of 3D points overlayed on image stacks.</td>
</tr>
<tr>
<td>Intensity, distance and angle calculations between points picked by the cursor. Vertical and horizontal intensity profiles through the cursor.</td>
</tr>
<tr>
<td>Label: overlay text, scale bars, arrows, and circles; in different colors.</td>
</tr>
<tr>
<td>Montage: interactive pasting of images from all windows into one page.</td>
</tr>
<tr>
<td>Superposition: Allows RGB color components to be translated independently in color images.</td>
</tr>
<tr>
<td>Spawn: generate a subprocess.</td>
</tr>
<tr>
<td>Model and model Show: an extensive menu for interactive and automatic 3D line model building:</td>
</tr>
<tr>
<td>Add, delete, branch, modify, cut and paste. Models are composed of linear &quot;objects&quot; that can each be labeled, colored, turned off or displayed on top of the images, and interactively rotated in 3D.</td>
</tr>
<tr>
<td>Interactive operations within PRISM can be recorded in a journal file for automatic repetitive execution or for fast recall of complex multiwindow environments. Displayed screens can be saved as image files for future reference, reproduction, and photography.</td>
</tr>
</tbody>
</table>
equipped with 12 Mbytes of memory (Parallax Graphics, Inc. Santa Clara, CA), and implements a very flexible multiple windowing architecture. Based on a combination of software and fast video processing hardware, each window can be arbitrarily sized and positioned, and the images can be independently zoomed. Because there is so much memory available on the display board, it is possible to preload a large number of images. This permits dynamic display under interactive user control (or digital movies under computer control) of preprocessed sets of images like sections through a 3D cube at any angle, tilted projections, and stereo pairs. Based on the file header information, windows presenting different processing manipulations of the same data can be geometrically related. Several windows, each associated with series of images can be activated so that sections and projections from different directions could be viewed and shown in movie mode.

In conjunction with an array processor (Mercury Computer Systems Inc., Lowell MA), PRISM implements on-line image processing (Fourier transforms, Fourier filtering, convolutions, local contrast enhancement, scaling) and projection and sectioning of gray-scale 3D data. Images can be presented in a single window or as stereo pairs. Interactive presentation of three-dimensional gray-level data is a most powerful tool for perception and understanding of these images.

In addition to the ability to present and compare data from different sets of images in simultaneously displayed windows, PRISM contains a true color option

![Image](image_url)

**FIGURE 1.** A photograph of the PRISM screen while displaying stereo pairs using the on-line 3D rotation capability. The actual rotated data is shown as a stereo pair in the two lower windows, while the two top windows show the orientation of the displayed data with respect to the original data. To the right of the display windows the menu window with mouse-activated commands is shown.
that allows two or three images to be superimposed, independently shifted to align them, and to independently tune the relative intensity and contrast of each color component. This is extremely useful for display and analysis of samples in which several molecules (like DNA, and antibodies to proteins) are all labeled simultaneously using different dyes. A listing of PRISM function options is given in Table 1. Examples of the interactive 3D rotation and 3D cube slicing functions are shown in Figs. 1 and 2.

3. PROCESSING DATA FOR DISPLAY

The initial aim in image processing is to present the data to the scientist with optimal contrast, and a minimum of interfering background and noise. A wide variety of filters have traditionally been used for processing two-dimensional images.
An important feature of our software is the ability to process a sequence of two-dimensional images as a true three-dimensional cube of data taking into account differences in scale and resolution along the axes. A list of typical processing routines is shown in Table 2.

A fundamental difficulty is developing means to adequately convey the information present in a complex three-dimensional reconstruction. The conversion of volumetric data into images on a two-dimensional raster screen is called volumetric rendering. Typically, such rendering techniques fall into two categories: those that extract a surface and then render it using shaded polygons, and those methods that perform various types of projection operations on the voxel data and maintain the gray-level information. Since this topic is covered extensively in Chapter 10, only a cursory overview will be given here.

For relatively simple objects, we have found the most generally informative approach for presenting the data with minimal processing is to use a series of rotated projections (Chen et al., 1990). In this approach, a series of two-dimensional projections of the three-dimensional data are formed at different angles of rotation about a single axis. These projections are rapidly and sequentially displayed to form a digital movie. The apparent back-and-forth rocking of the data gives a realistic perception of the three-dimensional density. This volumetric perception can be further enhanced by using stereo pairs. For more complex images, small subregions need to be chosen and to preserve interpretability. The projections themselves can

<table>
<thead>
<tr>
<th>TABLE 2. Useful Image Processing Methods for 3D Data</th>
</tr>
</thead>
<tbody>
<tr>
<td>Linear Fourier filters (smoothing, enhancement) by high, low, and bandpass Gaussian profile 3D filtering in Fourier space.</td>
</tr>
<tr>
<td>3D Convolutions filters for smoothing, and for enhancement by subtraction of the smoothed data from the original picture. The advantage of these simple local filters is the existence of fast algorithms (the complexity of the algorithm is proportional to the total number of voxels, and is independent of window dimensions. A fast 3D median filter was written by extending the 2D algorithm employing sweeping update of the histogram and the corresponding median (Huang et al., 1979).</td>
</tr>
<tr>
<td>3D gradients, and edge-enhanced pictures (Chen et al., 1990):</td>
</tr>
<tr>
<td>The modulus of the three-dimensional gradient of an image is computed and multiplied by the original image. The new image is a linear combination of the original and the weighted gradient image. This provides significant edge enhancement without building up noise.</td>
</tr>
<tr>
<td>3D interpolation of sections from 3D data cube and rotation of a cube of data by arbitrary angles.</td>
</tr>
<tr>
<td>3D projections of a slice from 3D data cube at any angle, and fast rotations about the X axis or Y axis for movies of tilt series and stereo pairs.</td>
</tr>
<tr>
<td>Entropy maximization filters.</td>
</tr>
<tr>
<td>3D alignment and matching of two sets of images, including translations, rotations, and magnification as free-fitting parameters.</td>
</tr>
<tr>
<td>Specialized image processing algorithms for optical and electron microscopic 3D reconstructions include: Deconvolutions of the microscope CTF to eliminate out of focus contributions. The nearest-neighbor approximation to the deconvolution problem is implemented in near real time during data collection. More sophisticated full 3D algorithms and deconvolution of projections are batch processed (Agard et al., 1989; Hiraoka et al., 1990).</td>
</tr>
<tr>
<td>Tomographic reconstructions from EM tilt series. (Back projections, iterative, constrained methods) (Abbey and Agard, forthcoming).</td>
</tr>
</tbody>
</table>
be computed along an arbitrary view angle. The use of a pure projection operator (sum along the view line) or even selection of the maximum value tends to give the resultant images a transparent, ethereal quality. The perception of solidity can be added by preferentially weighting the images in front of the stack or, more accurately, by considering that each pixel is opaque in proportion to its intensity. Based on the properties of fluorescent objects, Van der Voort and his colleagues (1988) developed an approach that combines aspects of ray tracing with volumetric rendering to provide dramatic images having strong shadows suspended above a uniform background. Because of the shadowing, this technique is probably less useful for very complex images, yet it should find application for some problems in EM tomography.

Edges within images can be effectively enhanced using an approach developed at PIXAR (PIXAR Corp., San Rafael; Driebin et al., 1988). The starting volumetric data is processed by multiplying the image pixel by pixel with the modulus of the three-dimensional gradient of the image. As this greatly enhances surface features, it is advisable to also incorporate opacity into the projection operator. Another approach is, instead of just taking the modulus of the gradient, to take the dot product of the gradient with a directional light source. This procedure produces a pleasant and fast surface rendering that importantly does not rely on the selection of a single contour threshold.

Improved background selection can be done utilizing a procedure called a 3D cookie cutter. This method uses an edge-detection scheme based on the PIXAR approach. By choosing a threshold, a surface boundary is defined and subsequently smoothed. The boundary is then used to zero-out background from the region of interest and thus present the object with better contrast and definition.

Three-dimensional surface rendering algorithms are useful for presenting the shape of solid bodies. Surface polygonization can be carried out by choosing a contour threshold to define the surface and then asking where that surface crosses through the 12 possible edges of each voxel in the image. These intersection points are determined by interpolation and then converted into vertices of a surface triangle (Chen et al., 1990). The triangular skeleton can be presented as line graphics, or the shaded surfaces can be displayed on a raster display. Shading incorporates background light, as well as diffuse and reflective illumination resulting from arbitrarily positioned light sources. Alternatively, ray tracing algorithms can also be used to enhance surface reflectivity and to more easily show overlapping objects by the use of transparent surfaces (Paddy et al., 1990).

4. IMAGE MODELING, ANALYSIS, AND QUANTITATION

4.1. Interactive 3D Model Building Using PRISM

Interactive modeling is probably the most effective mode of extracting general structural data from 3D image data. As discussed in the introduction, interactive modeling provides a means to trace the three-dimensional path of one or more objects that can be visualized within a reconstruction. The general approach used is to be able to step through a stack of images representing a volumetric data set
and to use a cursor to sequentially mark locations along the path of the object being traced.

Within PRISM, a model is defined as a collection of objects where each object is a line drawing representation of the path of a structure in three-dimensional space. Each object is defined as a set of connected nodes, each of which represent a single branch in the structure. A tree structure is used to store the connectivity of all these nodes, and a position array is used to store the coordinates of every node in real space. No limitations aside from total available memory space are imposed on the depth of the tree or on the number of entries at any level. For versatility and clarity, each object has its own set of display attributes which PRISM uses to determine its visibility and color, and any of six predefined marks can be attached to a node. Thus when it is displayed, the attached mark will be shown at the same location and in the same color as the object to which it belongs. While building the model, PRISM maintains a data pointer within the tree structure that determines the node at which modeling commands take place. There are four basic commands for updating a model: ADD, DELETE, BRANCH, and MODIFY. ADD will add a node right after the active node, and make the new node active. DELETE will remove the active node from the structure, and make the previous node active. BRANCH will create a new branch substructure at the active node and set the first node in the new branch as the active node. MODIFY will not affect the tree structure; it simply alters the coordinates of the current active node, allowing the user to reposition the selected point. In addition to these basic building commands, PRISM also provides a set of commands for moving pointers within the tree structure so that any part of model can be retrieved easily.

Building a model is accomplished by determining the locations of all the nodes. When displaying section images, a data point on the screen will correspond to a unique location in the data sampling space, as each image represents a slice in the sample. For such data, model building is accomplished by picking data points displayed on the screen. Stepping forward or backward in sections (which occurs rapidly because they have been preloaded) provides the third coordinate. Being able to rapidly step from one section to another is crucial for ease of use. Preprocessing the data to improve the clarity and contrast of structures being traced can be quite important for successful modeling. In our modeling of optical section data which suffers from nonisotropically due to a missing cone of data, in much the same way as can EM tomographic data, we have found that use of the PIXAR image gradient edge-enhancement approach greatly minimizes modeling problems.

In optimal cases, it may only be necessary to step through a stack of images in one, fixed orientation. However, for more complex images, especially where non-isotropic resolution may contribute to overlap in some directions, it is desirable to be able to model-build into the same data set that has been rotated in different ways. Thus, what appears ambiguous from one direction may be clear from another viewing direction. Although one could first model-build into data oriented in one direction and then correct the model by using another data orientation, it is greatly preferable to model-build while simultaneously viewing the data from several different viewing angles. If the views are related by $10-15^\circ$ tilt about the vertical axis, they will comprise a set of stereo pairs. Alternatively, orthogonal views can be used to allow precise modeling in the perpendicular direction. Especially in the case of
EM tomographic data, it can be important to also show the model superimposed upon the original projection data (or calculated projection data from a subregion). For this reason, PRISM has been designed to allow simultaneous modeling into several windows that are geometrically related. Thus, in the model building function, PRISM will choose what section to display in every window so that they all contain the same point in three-dimensional space as defined by the cursor and the section number in the active window as well as the projected view on top of 3D projections. This approach to window management and model building allows the user to simultaneously construct models from multiple viewing angles.

Models are easy to analyze for structural statistical properties, such as length, correlation, or anticorrelation of particular loci, approximate symmetry, and comparison of shapes. For quantitative analysis of fluorescence intensity in images, we employ routines that facilitate interactive and automatic definition of volumes and surfaces within which data should be evaluated. For example, for fast computations of volume and area averages, closed contour lines generated as a series of ordered contour points, can be converted to a run-length list (that is, first and last pixels of the area defined along each line), and, vice versa, the sequential list of border points enclosing an area required, for example, for contour plots, can be computed from the membership list of area patches. The area enclosed by a polygon is calculated by the equivalent of Green’s theorem for contour integrals (M. Koshy, private communication):

\[
\text{AREA} = \text{SUM}\{\left[ x(i) + x(i+1) \right]/2 \times \left[ y(i+1) - y(i) \right] \}, \quad i = 1, 2, ..., N - 1
\]

where \( x(i), y(i) \) are the coordinates of vertex \( i \), and there are \( N \) vertices. Patching, sorting, and counting of blobs can be done in one pass through the pixels in an image using the Union Find algorithm (Duda and Hart, 1973). This readily allows calculation of shape properties like the centers of mass, eccentricities, and the direction of the principal axes of the second moment ellipsoid (Castleman, 1979).

4.2. Computerized Recognition and Tracking of Objects

Where repetitive modeling of a particular kind of data is required, it is highly desirable to use an automatic or semiautomatic modeling scheme. This is a challenging task, that probably must be optimized for each type of data being analyzed. In general, the more a priori information that can be used to specify the behavior of the model, the more robust will be the entire procedure. Again, preprocessing can dramatically simplify the design and application of the necessary expert algorithms. Since many of these algorithms are not linear, their results can depend on data preprocessing for fast convergence to a visually acceptable estimate of the model. Proper treatment of the image to flatten or minimize background variation and enhance object discrimination will dramatically simplify the task of feature recognition. A trivial example of this is the ability to use threshold methods to define objects, or boundaries. The task of finding a suitable contour threshold can be greatly simplified by preprocessing to edge enhance and then smooth the data to maximize edge continuity. Having the ability to interactively process the data while testing various threshold choices is obviously quite desirable, since it permits
an optimal staring point to be selected for more sophisticated but time-consuming algorithms.

Automatic structural analysis can require complex logistics. It is necessary to reduce the dimensionality and complexity of the algorithms or to restrict them to the relevant area of interest. This is particularly important in three-dimensional image processing given the vast size of the analyzed data sets. Image segmentation can be based on gray levels, gradients, or generalized "energy" functions, which are associated with each voxel and are evaluated in a volumetric neighborhood of arbitrary size. Neighborhood operators provide a method for segmentation based on local properties or averaging to provide smoothing. Analyses at various resolutions—multigrid methods—(Brandt, 1986) not only save computational time by locating the region of interest at low resolution and then confining high-resolution processing to that area, but they also provide a potentially very powerful method for combining local and global interpretation of images.

4.2.1. Ridge Methods for Line Tracking: Tracking of Chromosomes in 3D Light Microscopic Data

A central unsolved problem in modern biology is understanding the complex structure of chromosomes. Related to this problem are questions pertaining to the general architecture of the cell nucleus. For example: How are chromosomes organized and folded within the tight confinement of the spherical nucleus? How does chromosome packing change as a function of the cell cycle or development? Is there a relationship between transcription and the spatial location of a gene? The problem of chromosome structure needs to be approached at two levels. At high resolution, which is best addressed by EM tomographic methods, the question is to determine how DNA and protein become organized into successively higher order chromosomal structures. Also of central importance is understanding how the higher-order structural organization changes during condensation, decondensation, transcription, and replication. At the more global level, which is best examined by 3D light microscopy, are questions of spatial organization of chromosomes within the nucleus and their dynamic behavior during the cell cycle. Structural features that are of importance are the localization of genes along the chromosomes (mapping complex chromosome paths to straight linear structures for comparison with genetic maps and to determine architectural features such as compaction) and the correlation of spatial position with function. Such an analysis would, for example, include examining the association of specific genetic loci with the nuclear membrane, spatial correlation of the two homologous genes, active genes and puffs in some cases, or inactive genetic regions like heterochromatin.

In our light microscopic work, 3D images of nuclei are collected from intact *Drosophila* embryos that were fixed and fluorescently stained using DNA specific dyes, labeled hybridization probes, or antibodies to DNA associated proteins (Hiraoka et al., 1987; Agard et al., 1988; Rykowski et al., 1988; Paddy et al., 1990). Alternatively, many sets of 3D data can be collected as a function of time from living embryos microinjected with fluorescently labeled molecules such as tubulin, topoisomerase II, or histones, which then become incorporated into spindle or chromosomal structures (Minden et al., 1989; Hiraoka et al., 1989). The micro-
scopic images recorded at a series of focal planes are processed to eliminate out of focus contributions, and projected tilt series and orthogonal section series are computed to facilitate visualization (Agard et al., 1989). The processed 3D series are then displayed in separate windows in PRISM. Models are constructed by interactively positioning the cursor along the chromosome paths as judged by the user based on the multifaceted presentation of the 3D density.

Semiautomatic modeling allows processing of a large number of nuclei for statistical analysis. This is especially important when trying to derive what are the biologically significant features of a modeled organization. The semiautomatic algorithms that we have employed use local contrast filtering and three-dimensional gradients to define a sausage-like corridor. Within this corridor, the central ridge of

FIGURE 3. An example of the result of semiautomatic modeling of chromosomes. A stereo pair of an anaphase chromosome figure is shown at top, and the models are shown below. The user initiates the automatic tracking by seeding the model. Models created by random seeding at image voxels having an intensity above a set threshold demonstrate that most of the tracks accurately follow the chromosomal paths. The tracks that did not follow the chromosomes were a result of confusion caused by a combination of overlapping of two chromosomal arms and the lower resolution along the light microscope optical axis.
maximum density is determined. In order to avoid the accumulation of background and noise from regions outside the chromosome track, the data are preprocessed using a three-dimensional cookie cutter. Using spline routines to account for persistence of direction, a suggestion for the track of the chromosome is then made. Crosses between chromosome paths and chromosomal bands which introduce a somewhat fragmented staining pattern can be bridged by interactive editing, based on the known chromosome structure, or by "bridging" the gaps based on persistence of directions from both sides. An example of tracking seeded at a number of randomly chosen points along the chromosomes and automatically drawn from that point to in two directions is shown in Fig. 3.

An automatic fiber-tracking algorithm was developed by Borland et al. (1988) to identify chromatin fibers in EM tomographic reconstructions from expanded mitotic chromosomes. Given an appropriate starting point, their approach is to use the density gradient computed within a small region to determine the appropriate direction of the fiber center. Once the center is determined, they then attempt to travel along the central ridge of maximum density. With their algorithm, closely juxtaposed fibers whose densities overlap will be treated as a single entity. They use a threshold as a way of separating chromatin fibers from background. Although it seems that much manual intervention may have been required, using this approach they have been able to successfully track long segments of loosely packed chromatin fibers within metaphase chromosomes.

Houtsmuller et al. (1990) have described a semiautomated chromosome-tracing approach based on the idea of a "homing 3D cursor." When positioned near a chromosome, the cursor automatically attempts to home-in on the center of the chromosome. This is accomplished by convolving the voxel data with a small sphere centered at the current cursor position. The cursor is then stepped to the location of maximal overlap, and the process repeated until convergence is reached. Furthermore, the user can specify enough parameters (direction, distance, etc.) to allow the cursor to proceed automatically along a region of the chromosome and then return to manual mode. Perhaps the most serious drawback of this approach is the convolution operator used to center the cursor in the density. Since the sphere needs to be somewhat larger than the diameter of the object being traced for this method to work, other objects must be spaced further away than the diameter of the sphere. This will obviously pose problems when examining complex structures having numerous regions in close proximity. However, it should be possible to utilize preprocessing methods such as the cookie cutter to redefine spatial boundaries within which the homing cursor must operate.

4.2.2. Matched Filter Methods for Identifying Objects in Complex Images

Tracking Microtubules in Nerve Growth Cones Microtubules are the most dynamic component of the cell cytoskeleton. In order to study their role in cell locomotion and nerve growth, tubulin can be fluorescently labeled and micro-injected into cells. Its subsequent incorporation into microtubules allows them to be directly visualized by fluorescent microscopy (Kellogg et al., 1988). Because of their long axonal processes and well-differentiated cell bodies, neurons are excellent cells for examining the behavior of microtubules. Since living nerve cells are very sensitive
to the fluorescence excitation light level, images need to be recorded at very low light levels and thus tend to be quite noisy. Obtaining quantitative information on the rate of growth, directionality, etc., requires that the pictures be "distilled" into linear graphs presenting the microtubules. This can be readily achieved by using a matched filter that recognizes rods. The algorithm developed by Soferman (1989) tests for rods in all orientations by calculating the correlation of a box about each image point with rods in a discrete set of orientations. The result can be enhanced by subtracting the correlation obtained with two parallel rods displaced sideways by the characteristic line thickness to be traced. The original pixel is then replaced by the best correlated rod, scaled by the original image intensity. Although slow and time consuming in real space, this process is efficiently performed in Fourier space. The results of this approach shown in Fig. 4 demonstrate a better pattern of continuity along the track than obtainable using edge-enhancement algorithms which tend to produce highly fragmented paths along the microtubules due to noise.

*Locating Beads in Electron Micrographs*  Another example of using matched filters to extract objects from complex images involves finding small beads in electron micrographs. A necessary step in computing tomographic reconstructions from tilt series of a thick sections is to bring the different images to a common coordinate system, so that relative displacements and rotations introduced by not precisely

![Graphs showing fluorescence excitation and microtubules dynamics.](image-url)

**FIGURE 4.** An example of the use of a matched filter (based on oriented rods, Soferman, 1989) for tracking microtubules dynamics in the tips of growth cones. (Taken from E. Tanaka, unpublished data.)
eucentric tilting can be compensated. The most reliable and internally consistent algorithm for tilt series alignment is to record the positions of small colloidal gold beads which are dispersed on the surface of the sample, and to use the projected coordinates to refine values for \((x, y, z)\) bead position, tilt angle, tilt axis, translation, and magnification (see appendix in Luther et al., 1988). Locating the small beads within complex images can be challenging due to their relative low contrast (especially at high tilt angles). The matched-filter approach followed by pattern recognition and classification methods provides a very reliable method for accurate and automatic identification. The matched filter is performed as a Wiener-type inverse filter:

\[
\text{Filtered image} = \text{FT}^{-1} \left( \frac{\text{FT(image)}}{\text{FT(bead)} + k} \right)
\]

where \(\text{FT} = \) Fourier transform, \(\text{FT}^{-1} = \) inverse Fourier transform, and \(k\) is a user-chosen constant to determine the balance between sharpness and noise. The peak positions, indicating sites of best matching to the bead template, are then selected based on local contrast and finally screened on the basis of size and shape. Finally, a procedure has been developed which automatically matches the beads from all of the sections in the tilt series, producing a list of beads that is consistent throughout the data set and can be used for alignment. Images of the original digital micrograph and the processed images are shown in Fig. 5.

![FIGURE 5. The identification of gold beads in electron micrographs using a matched Wiener filter. (a) shows the raw EM data taken from the CCD. (b) is the mass normalized image to convert the data to electron absorbance, (c) is the Wiener matched filter, and (d) is the filtered result after thresholding, demonstrating the high contrast achieved. This approach can accurately detect beads even when there is very little contrast difference with respect to the rest of the image.](image-url)
4.2.3. Tracking Nuclei in Four Dimensions: A 3D Time Series

Tracking cell lineages in embryonic development is of interest in order to study the mechanisms of segmentation of new tissue, development of organs, and the underlying processes of cell differentiation. One of the unique features of the light microscope is its ability to acquire images from living biological specimens. To determine patterns of cell lineage in *Drosophila melanogaster*, we followed a large field of cells in a developing embryo that had been injected with fluorescently labeled histones (Minden et al., 1989; Hiraoka et al., 1989). The histones become incorporated into chromosomes during subsequent rounds of nuclear division, and act as a vital nuclear stain.

During embryonic development there are dramatic changes in the positions of nuclei. For the *Drosophila* embryo at the relevant stages, the cells move in all three dimensions. Because the embryo is quite thick, and the movements are large, three-dimensional imaging is required to follow the large number of nuclei that migrate in and out of any given focal plane. Thus, collecting data during embryonic development necessitates the acquisition of a time series of three-dimensional image stacks, which can be called a four-dimensional stack.

The problems of tracking objects in time series have been intensively studied in relation to fields like robotics, artificial intelligence, and automatic navigation. The special property of the time axis with respect to the other three spatial coordinates is that if the data are collected sufficiently rapidly, the changes in the images are bounded and can be used to identify the motion of defined objects from one time frame to the next. For example, to trace a single bright point in a dark field requires only searching for the object within a certain radius of its position in the previous frame, the radius being defined by the object's velocity and the time increment between the frames.

This approach, unfortunately, cannot be easily exploited when examining images of more diffuse objects. An example would be the definition of the velocity field of clouds, as recorded by a time series of satellite images, where evolving continuous patterns and textures need to be followed rather than a set of isolated points. Correlation methods can be applied to track defined regions which are moving and changing slowly with time. By comparing time sequences at reduced resolutions, the computational load can be reduced, and, more significantly, high contrast can be achieved for specific moving patterns or features within an evolving scene (Baker and Bolles, 1989). Still, the higher the resolution the details of interest, the closer in time the images have to be taken, so that correlations between consecutive time points do not decay for that resolution.

Unfortunately, for imaging developmental events in *Drosophila* embryos, a variety of instrumental limitations place a limit on how fast the three-dimensional data stacks can be acquired. For this reason, the nuclei traced were found to move distances greater than their own diameters between consecutive frames during daughter chromosome separation in anaphase. The conventional approach using correlations would thus not be able to track the lineage of cells through several mitotic divisions. An alternative approach is to first identify the objects to be traced in each time point, based on their known geometric properties, then apply point tracking.
A three dimensional patching routine was written, based on a 2D algorithm developed by Soferman (1989) that can be described as the emergence of mountain tops in a landscape during the decline of the water after a flood. All the voxels in the image are first sorted by height (intensity). Then, starting at the highest, and going down the list, they are in turn "patched" by assigning to each voxel the number of a touching patch or a new patch number for isolated voxels (new tops of hills). As patches grow, the newly added voxel may connect two existing patches. At that point, a decision is made to merge small patches (resulting from noise) into large ones or to leave them as separate patches if both are large enough. Patches can be restricted to some maximum size, although in order to prevent new patches from appearing it is necessary to assign the neighboring voxels to them as the

FIGURE 6. The pattern of nuclear motion following mitosis of the 14th cycle as tracked in time from images of *Drosophila melanogaster* embryos which were injected with fluorescently labeled histones (Minden *et al.*, 1989; Hiraoka *et al.*, 1989). Surprisingly, the patterns of motion were found to be coordinated over quite large areas. Mirror symmetry about the dorsal-ventral line can also be seen.
“water” level continues to descend. This approach allows all nuclei to be identified as local peaks with a typical size, irrespective of the background level or their peak height. The algorithm can be stopped at any fraction of the total number of voxels, corresponding to the fractional volume of the structure of interest (size of a nucleus). The original 2D lake algorithm (Soferman, 1989) exhausted all pixels and registered the lines of patch meetings which are the ridges in the intensity landscape (valleys for descending water). The network of lines, their intersects, and the regions generated in one sweep have interesting topological properties for the purpose of segmentation of pictures.

In our case, the isolated patches represent nuclei and define their center of mass in each time slice. Once each nucleus has been identified, tracking the paths in time was easy to achieve by point proximity criteria. To visualize mitotic events, the analysis must also allow for binary branched paths. The derived nuclear tracks can be used to assign lineages or patterns of parenthood, which is very important for understanding the biology. Interestingly, the tracked motion reveals patterns of motion (see Fig. 6) and changing orientations of cell division which reflect differentiated domains in the developing embryo (symmetry about dorsal-ventral line).

5. SUMMARY

Three-dimensional image reconstructions have the potential for dramatically extending our understanding of cell structure and behavior. Of central importance is the ability to extract relevant information from the complex voxel data derived from light and electron microscopic reconstructions. In general, for biological data this is an extremely difficult problem that is just beginning to be approached.

We have discussed the importance of utilizing a powerful and versatile interactive image display and processing system for extracting relevant information from the voxel data. In addition, the hardware and software that we have been developing and using for these purposes have been described. In many cases, the relevant biological information can be in the form of a line drawing model that represents the path of the structure in three-dimensional space. The key to dealing with such models and voxel data simultaneously are an image file structure that maintains geometrical data for multidimensional images and a display and modeling system that can utilize this information to allow orientation and processing-independent superimposition of model and data. Multiple windows enable simultaneous modeling in several orientations. A variety of approaches for data preprocessing as well as automatic and semiautomatic methods for model generation were also surveyed.

ACKNOWLEDGMENTS

This work was supported in part by NIH GM25101 and NIH GM32803 to JWS and NIH grant GM31627 to DAA, and by the Howard Hughes Medical Institute. D.A.A. was also supported by an NSF presidential young investigator grant.
REFERENCES


