Interfacing Image Processing and Computer Graphics Systems Using an Artificial Visual System

James M. Coggins, Kevin E. Fogarty, and Fredric S. Fay

Biomedical Imaging Group
Department of Physiology
University of Massachusetts Medical School
Worcester, MA
and
Computer Science Department
Worcester Polytechnic Institute
Worcester, MA

ABSTRACT

An Artificial Visual System (AVS) has been developed to simplify three-dimensional microscope images for presentation and manipulation in an interactive computer graphics system. The AVS consists of several sets of spatial filters that decompose an image along three different measurement continua. A recombination algorithm processes the filter outputs to detect objects, to eliminate noise, and to map the detected objects into points in a multidimensional feature space. Recent discoveries regarding the geometry of the points in the feature space are described. One recent result simplifies the AVS by decreasing the number of filters required to obtain the same measurements. Not only are accurate measurements possible, but certain image distortions can be modelled and counteracted in the feature space.

Key words: computer vision, pattern recognition, interactive computer graphics

Introduction

Difficulties in the analysis of natural images arise from random noise, aliasing from the digitization grid, systematic distortions such as optical blur, and from an excess of information -- accurate but irrelevant data, relevant but ambiguous data, or simply too much relevant data -- that we call "information overload". Many techniques exist for correcting noise and distortion [1], but the information overload problem requires an understanding of the aspects of the image that are important for the particular application and techniques for specifying and extracting the relevant aspects of the image. Once the information is extracted, then interactive computer graphics can be applied to present the extracted information and to provide powerful interactive facilities to support image interpretation activity [2].

We have developed an interface between image processing and computer graphics systems that both provides a mechanism for specifying the salient aspects of the image and extracts that information in a form that can be used to build an interactive graphics model of an image. The interface takes the form of an artificial visual system (AVS).

The development of the AVS as a means for addressing the information overload problem has been motivated by a biomedical research problem involving interpretation of three-dimensional fluorescence images. In this paper we will present an overview of the biomedical research problem and then show how the AVS for this problem was designed. Additional applications of AVS techniques will be discussed.

Background

Our goal is to elucidate the contractile mechanism of smooth muscle cells [3,4]. One of the proteins believed to play a role in that mechanism is a-actinin, which occurs in two types of discrete bodies of concentration: irregular plaques on the cell membrane, and oblong bodies distributed throughout the cytoplasm and oriented within 30 degrees of the long axis of the cell. Organizational patterns such as strands of these bodies branching and twisting through three dimensions may be discerned if the locations of the bodies and the orientations of the oblong bodies are known. Different kinds of organizational patterns could support different hypotheses of cell structure and function.

A three-dimensional image of the protein distribution in a single, isolated cell is obtained by acquiring a series of 2-D optical sections of the cell using Fluorescence Digital Imaging Microscopy [3,5]. Several types of noise are minimized using averaging and normalizing operations during image aquisition [3].

There remains a serious optical distortion in the direction of focus arising from fluorescence sources from out-of-focus planes above and below the focal plane. This distortion has been empirically modelled by imaging a

fluorescent bead smaller than a voxel. The image obtained serves as an empirical estimate of the point spread function of the overall optical system and is used in a constrained iterative restoration procedure to reduce the distortion in the cell images [3]. The restoration reverses about 80% of the optical distortion, but significant distortion remains in the direction of focus. This residual distortion elongates the image of a spherical object in the direction of focus yielding an apparent axial ratio of about 3:1.

The restored image is still difficult to analyze due to information overload. A 64x64x64 cell image can contain over 200 discrete concentration bodies. The oblong bodies are about 1 voxel wide and about 5 voxels long, but their oblique orientation and the residual distortion spread the image of each body over a larger volume. Locating the bodies manually is a tedious task requiring constant flipping between adjacent image planes, correlating traces of the bodies through the planes. Estimating the orientations of the bodies from these traces is even more difficult. In addition to the residual optical distortion, the small apparent size of the bodies makes aliasing from the digitizing grid a serious concern; the difference in the digitized image between two small bodies at nearby orientations involves a subtle shift of energy among a few voxels.

The three-dimensional nature of the data adds more information overload. Since the structures we seek twist through three dimensions, no single two-dimensional view can capture all of the relevant information about the structures.

We need, then, a system to simplify the restored three-dimensional image by locating the protein bodies and determining the orientation of the oblong bodies. The system we have developed is called a three-dimensional artificial visual system.

Designing an Artificial Visual System

An Artificial Visual System (AVS) is a set of filters along some equivalence dimension (e.g. spatial frequency, size, orientation) and a recombination algorithm for mapping the filter responses into a perceptual feature space [4,6]. Defining an AVS involves selecting approprate equivalence dimensions, designing filters sensitive to different values along those dimensions, and defining the recombination algorithm to perform the visual task at hand.

The equivalence dimension is a continuum along which measurements can be made. Objects in the image will be treated as stimuli to be represented or measured along these continua. Complex structures may require measurements on several equivalence dimensions to adequately characterize the structure of the stimulus.

A sequence of filters is defined for each equivalence dimension. Each filter is sensitive to a different range of values along the continuum. Normally, the sensitivity profiles of the sequence of filters on their equivalence dimension are designed to be tapered and overlapping. Such an ensemble of filters provides a unique sequence of responses for every single-valued stimulus on the equivalence dimension [4,7]. The filter responses serve as coordinates of the stimuli in a multidimensional feature space [8]. Measurements of the stimuli are based on the geometry of the mapping of stimuli into this feature space.

The purposes of the filters are as follows: (1) to decompose the image into separate bands of information so that important and useful information can be identified more easily; (2) to represent the a priori knowledge concerning the objects being sought and the precision of measurement required and (3) to define a feature space into which the image will be mapped [6]. Use of a priori knowledge in the filter design enhances both the sensitivity and the efficiency of the analysis. Sensitivity is enhanced because the filters can be tuned to detect the structures of interest. The filters can also embody the degree of uncertainty in the a priori knowledge, as demonstrated in the present study where the orientations and sizes of the oblong dense bodies are known a priori only to an approximation. Efficiency is enhanced because the a priori knowledge decreases the number of filters required for the task; a more generalized analysis or higher measurement precision requires a finer or more complete decomposition of the image, requiring more filters. The AVS allows a priori information to be incorporated in the design of the filters rather than in the design of new problem-specific heuristic algorithms.

The recombination algorithm merges the information from each set of filters to perform the visual task required. The algorithm may involve thresholding to eliminate noise, averaging to compute a measurement, location of relative extrema to create a representation or detect a particular kind of stimulus. Other recombination algorithms provide edge detection or texture representations [6,9]. If the objective of the recombination algorithm is to create a representation of the stimulus to be operated upon later by other processes, then the AVS serves as the "low-level vision" component of the vision system. If the objective of the recombination algorithm is to produce the required measurement, then the task is called a "pre-attentive" operation.

The AVS for the Cell Study

For the smooth muscle study, three equivalence dimensions are defined:

length, declination, and azimuth. The filters for each dimension are constructed in the spatial domain and convolved with the 3-D image using Fourier Transform methods. The filtering operation produces three series of filtered images. Fortunately, we can arrange the computations so that the filtered images can be created, processed, and discarded, so it is not necessary to store the entire ensemble of 3-D filtered images at once.

The filters along the length dimension are used to locate the bodies, differentiate them from noise phenomena in the image, and to measure the lengths of the bodies. The filters are shaped (approximately) as truncated cones of length R=3, 5, 7, and 9 voxels [4]. See Appendix 2 for details of filter creation. The width of the cone is determined by a priori information concerning the expected variability in the orientations of the oblong bodies.

When convolved with the fluorescence image of the cell, a local relative maximum response occurs when the filter contains a locally maximum fluorescence intensity. These local maxima include the centers of all of the concentration bodies due to the symmetry both of the filters and of the bodies, along with maxima caused by noise or fluorescence hot spots unrelated to the dense bodies we seek.

The recombination algorithm for processing the output of the R series filters involves two steps. First, the local maximum responses are located and thresholded to eliminate miniscule fluorescence hot spots and random noise. Second, if a real concentration body has been found, then the energy captured in the sequence of R filters increases until the filter size exceeds the size of the body, at which point the filter response remains constant. Thus, the length measurement works as follows: If, at a particular location in the image, the smallest filter (R=3) has an superthreshold maximum and the R=5 filter response at the same location is also a superthreshold maximum that is significantly (1.2 times) greater than the R=3 response, then a valid body (length at least 4) has been found. The estimated length will be increased to the size of the next larger R filter as long as the next R filter has a superthreshold maximum at the same location with an intensity greater (1.2 times) than the response of the current R filter.

The orientations of the oblong bodies are measured in terms of the declination (0<=theta<=90) of the long axis of the body from the y axis of the 3-D image and the azimuth (0<=phi<360) about the y axis using the x axis as phi=0. (The 3-D images are acquired with the long axis of the cell oriented vertically in the microscope image, corresponding to the y axis of the 3-D image.)

The filters for the theta and phi equivalence dimensions are constructed by partitioning the R filters into "hollow cones" for theta measurements or "wedges" for phi measurements [4]. Theta filters are centered at theta=0, 5, 10, 15, 20, 25, and 30 degrees. Phi filters are centered at phi=0, 60, 120, 180, 240, and 300 degrees. See Appendix 2 for details of the creation of the phi filters. The R filters at R=5, 7, and 9 are thus partitioned into theta and phi filters giving a total of 40 filters covering all combinations of values in all three equivalence dimensions.

The 36 theta and phi filters are convolved with the input image. For each body identified by the R filters, we record the responses at the body center of the twelve theta and phi filters corresponding to the size of the body. We use only the theta and phi filters best matching the body size to avoid incorporating responses to nearby structures or noise into the orientation measurements. (There is some evidence that responses to fluorescence signals outside the body can be eliminated mathematically without using separate filter sequences for each possible length. We plan to investigate this possibility later.) These filter responses form two six-dimensional vectors that characterize the theta and phi orientations of the bodies. The feature vectors are normalized by the responses to the R filter matching the body's size so that the sum of the values in each feature vector is 1.

The oblong bodies have a single preferred orientation (by virtue of their oblong shape), so the sequence of responses to overlapping, tapered filters defined along the theta and phi equivalence dimensions is unique for every possible orientation. In fact, the normalized response to each filter may be used as a weighting factor indicating the degree to which the body's spatial energy distribution matches the filter's preferred orientation. The recombination algorithm for constructing an estimate of the orientation of the body involves (circularly) averaging the filter center orientations weighted by the responses of the filters. This is equivalent to a sum of vectors whose polar representation (r, alpha) has the r component equal to the normalized filter response for the filter centered at phi=alpha.

Application and Evaluation of the AVS

The performance of the AVS has been extensively studied using artificial images containing model cylindrical bodies at regularly spaced orientations. Model bodies are created at theta angles 0, 5, 10, 15, 20, 25, and 30. For each theta angle (except theta=0) bodies are created with phi angles 0, 30, 60, 90, 120, 150, 180, 210, 240, 270, 300, and 330 degrees. These images have been analyzed themselves, and they have been distorted using the

empirically determined point spread function of the microscope system and partially restored using the iterative restoration algorithm before further analysis.

The objective of this study is to determine from the regular sampling of body angles whether the mapping of the bodies into the feature spaces displays similar regularity. If so, then measurements of body orientations may be reliably performed from the feature space.

With the noiseless images, the R filters correctly locate the bodies, and application of the circular weighted averaging procedure on the sequences of theta and phi filter responses obtains orientation measurements correct within 2 degrees in theta and 5 degrees in phi. The errors remaining are due to aliasing effects and roundoff errors.

With the blurred/restored images, the interpretation of the filtered images is far less straightforward. The residual z-axis distortion, which elongates the images in the phi=90 and phi=270 directions, ruins the theta filter measurements. The problem appears to be that the z-axis distortion causes theta measurements of bodies at a fixed theta orientation to vary through a wide range of values as phi varies. Thus, bodies at different theta orientations are indistinguishable unless the phi angle is already known. We could compute the phi angle first and establish for each phi angle appropriate thresholds for interpreting the theta data, but a more elegant approach has been found.

We have discovered that both the theta and phi measurements can be reliably obtained from the phi filter data alone. This simplification is possible because as the theta angle of a body increases, an increasing proportion of the volume of the body moves away from the theta=0 axis, resulting in increased energy in a preferred phi direction. Therefore, we can measure phi by computing the circular weighted average of the phi filter responses, and we can compute theta by measuring the intensity of the phi angle preference.

When this approach is applied to noiseless model images, the pattern of points in the feature space is a series of concentric circles. Each circle corresponds to a particular theta value, and the polar angle alpha of each point along the circle is precisely the phi angle of the corresponding body.

When the same approach is applied to blurred/restored bodies, the pattern of responses in the feature space is a series of concentric ellipses. The minor axes are in the phi=90 and phi=270 directions, corresponding to the direction of the z axis distortion in the images. The effect of the distortion on the feature space, then, is to decrease the sensitivity of the filters in the direction of the distortion. We can correct this problem in the feature space by converting the polar (r,alpha) coordinates to Cartesian form and boosing the y component of the Cartesian vectors. The scaling factor was determined by measuring the eccentricities of the ellipses in the feature space and for our current data is about 3.2, which corresponds to the axial ratio induced by the residual distortion in the image of a sphere. This scaling factor must be recomputed only if the imaging system changes.

Evaluation of the approach is carried out by measuring the angle between the actual and estimated body orientation vectors. That is, the (theta,phi) values of a set of model bodies and the corresponding estimates are converted to Cartesian form and the angle between the two vectors was computed. The average error angle on our model images is less than 2 degrees. Additional tests are in progress on noisy model images. Preliminary results indicate that in the presence of noise, the errors in the orientation measurements increase gradually as the signal/noise ratio decreases. Further work on interpreting these results is in progress.

Interfacing with a Graphics System

The information extracted by the AVS consists of a list of (x, y, z) locations where a dense body was found along with (r, theta, phi) measurements on each of the oblong dense bodies. This position and orientation data has been used to create a graphic model of the 3-D distribution of dense bodies. The bodies are represented as lozenge-shaped solid objects having the measured length and orientations. The cell image is created by projecting prototype bodies into space and then subjecting the projected model to the required viewing transformations [2].

The user of the graphics system can specify any view position, including positions inside the graphic model that correspond to positions inside a cell. The three-dimensional distribution of the dense bodies can be explored by marking (in color) particular dense bodies in order to trace a network or follow a strand of bodies through the cell [4].

Interaction is provided by a 3-D wire-frame arrow cursor whose movement is controlled by a three-dimensional joystick. Joystick movements can be interpreted as translation or rotation commands depending on a switch setting.

At present, a single view from a single viewpoint is presented by the graphics system [4]. We are considering implementation of a dual viewport system

that could enable presentation of the view as a stereo pair or as a proximal/distal view pair.

Discussion

The artificial visual system has proven to be a powerful tool for image analysis due to the following properties [4,6]:

1. A priori knowledge can be effectively incorporated into the design of the filters and the recombination algorithm. The artificial visual system can be tuned and experiments can be performed by changing only the filter data and not the analysis algorithms. This enables rapid prototyping and optimization of the artificial visual system with a minimum of reprogramming and algorithm development. In addition, since a few simple algorithms suffice for much of the processing requirements, special devices such as array processors can be brought to bear to enhance the speed of execution.

2. Spatial filtering is intuitively understandable. Filters can be defined either in the spatial domain or in the frequency domain. Either way, the filters and their effects on images can be determined and understood easily since the filter is applied uniformly over the image. Understandability is especially important in applications since decisions will be based on the results of the computer procedures and those decisions must be defended based on an understanding of the computer's results.

3. Fast algorithms exist for performing spatial filtering. Spatial domain convolution, spatial frequency domain multiplication, recursive filtering, and in-place filtering are all well-known algorithms for performing spatial filtering. Depending on the hardware support and the nature of the filtering to be performed, any of these equivalent algorithms can be chosen. These algorithms are amenable to parallel processing to enhance execution speeds.

4. The most important property of spatial filtering is that a suitably constructed ensemble of filters can be used to decompose an image along any of several continua (e.g. size, orientation, spatial frequency, shape, etc.). Thus, the ensemble of filters in a visual system can be constructed so as to define a meaningful feature space.

Moreover, a stimulus can be located along a continuum (orientation, size, spatial frequency) by an ensemble of tapered, overlapping filters. With suitably defined filters, every possible stimulus along the continuum yields a unique pattern of responses from the ensemble of filters, and thus a unique location in the feature space. Increased accuracy in the measurements of stimuli requires a finer decomposition of the relevant continuum involving a larger number of more narrowly-defined filters. Thus, the tradeoff between cost and measurement accuracy is explicit and measurable.

Conclusion

An Artificial Visual System has been developed to simplify three-dimensional fluorescence microscopy images. The AVS locates bodies of interest in the 3-D image, discriminates the bodies from noise, and measures the 3-D orientation of each body. The measurements are made by using the outputs of a series of spatial filters to map each body into a point in an abstract feature space. The geometry of the mapping allows the orientation angles to be computed directly from the mapping.

Moreover, distortions in the 3-D image due to the image acquisition system that were not corrected by noise reduction or image restoration algorithms appear as systematic distortions of the geometry of the feature space. This residual distortion can be measured and corrected in the feature space, enabling accurate measurements in spite of the imperfections in the image data.

The measurements are then used to create a simplified graphical image that can be viewed and manipulated using interactive graphics tools. Viewpoints corresponding to locations inside a cell may be constructed. The graphics system user can interact with the simplified image to record organizational patterns that may explain the operation of the contractile machinery of the cell.

Appendix 1: Coordinate Conversions

This appendix gives the algorithms for converting Cartesian vectors to (theta,phi) orientation vectors and vice versa where the (0,0) direction is the y axis and the (90,0) direction is the x axis.

Convert a (theta,phi) orientation to a 3-D Cartesian unit vector [x,y,z] as follows:

```
x = sin(theta)*cos(phi)
y = cos(theta)
z = sin(theta)*sin(phi)
```

The following algorithm converts a 3-D Cartesian vector [x,y,z] to a (theta,phi) orientation vector:

```
Let r1 = sqrt(y^2+z^2)
theta  = arctan(r1/y)  if y is not zero
       = 0             if y=0 and r1=0
       = 90            if y=0 and r1>= 0
phi    = arctan(z/x)   if x is not 0
       = 0             if x=0 and z=0
       = 90            if x=0 and z>=0
```

Appendix 2: Filter Definitions

Filters defined as geometric cones and segments of cones [4] were found to be subject to errors such as false positives or mislocated maxima. Superior results have been obtained with filters created by integrating images of a cylinder rotated about its center. The resulting filters have higher sensitivity where many of the rotated cylinder images overlap (i.e. at the filter center).

Each filter is formed by a weighted sum of images of a rotated cylinder 9 voxels long and 1 voxel in diameter corresponding to the longest apparent size of the dense bodies in our images. The weighting factor applied to each cylinder is based on the difference between the cylinder orientation (t,p) and the filter's preferred orientation. We define two utility functions as follows:

```
WT(t,p;tmax) = 1      if t<=tmax
             = max(0,1-[(t-tmax)/10])
                      if t>tmax

WP(t,p;pcen) = max(0,1-[|pcen-p|/60])
```

The WT function assigns a weight of 1 to cylinders whose theta orientation is less than or equal to tmax and attenuates cylinders with larger theta values with the weight decreasing linearly with (t-tmax) and reaching 0 at a theta orientation of tmax+10 degrees. The WP function attenuates the cylinder images linearly as the phi orientation differs from pcen with the weight reaching zero 60 degrees away from pcen. Note that the difference (pcen-p) must be computed mod 360. Now we use the utility functions to define the cylinder weighting functions for an R filter (a cone) and filters P1-P6 (wedge-shaped phi filters):

```
R(t,p)  = WT(t,p;30)
P1(t,p) = WT(t,p;30)*WP(t,p;0)
P2(t,p) = WT(t,p;30)*WP(t,p;60)
P3(t,p) = WT(t,p;30)*WP(t,p;120)
P4(t,p) = WT(t,p;30)*WP(t,p;180)
P5(t,p) = WT(t,p;30)*WP(t,p;240)
P6(t,p) = WT(t,p;30)*WP(t,p;300)
```

The R filter is equal to the sum of the Pi filters, making the response of the R filter a reasonable normalization factor for the sequence of Pi responses. This normalization eliminates the effects of different overall intensity in different bodies.

Shorter filters are created by multiplying the above filters by a sphere of an appropriate radius. R and Pi filters have been created at lengths 3, 5, 7, as well as 9. When the size of a body is determined, the P series filters corresponding to that size are used to estimate the theta and phi angles.

References

1. A. Rosenfeld and A. C. Kak, Digital Picture Processing, second edition, New York : Academic Press, 1981.

2. J. D. Foley and A. vanDam, Fundamentals of Interactive Computer Graphics, Reading, MA:Addison Wesley, 1984.

3. F. S. Fay, K. E. Fogarty, and J. M. Coggins, "Analysis of Molecular Distributions in Single Cells Using a Digital Imaging Microscope," in Optical Methods in Cell Physiology, P. de Weer and B. Salzburg, editors, John Wiley and Sons, New York, 1985.

4. J. M. Coggins, K. E. Fogarty, and F. S. Fay, "Development and Application of a Three-Dimensional Artificial Visial System," Proc. Symposium on Computer Applications in Medical Care, Nov. 10-13, 1985, Baltimore, MD. pp. 686-690. Also in press in Journal of Computer Methods and Programs in Biomedicine.

5. D. J. Arndt-Jovin, M. Robert-Nicoud, S. J. Kaufman, and T. M. Jovin, Fluorescence Digital Imaging Microscopy in Cell Biology," Science, vol. 230, no. 4723, 18 Oct 1985, pp. 247-256.

6. J. M. Coggins, A Framework for Texture Analysis Based on Spatial Filtering, Ann Arbor:University Microfilms, dissertation for Ph.D. Michigan State University, 1982.

7. W. Richards, "Quantifying Sensory Channels: Generalizing Colorimetry to Orientation and Texture, Touch, and Tones," Sensory Processes v. 3, 1980.

8. R. O. Duda and P. E. Hart, Pattern Classification and Scene Analysis, New York:Wiley.

9. J. M. Coggins and A. K. Jain, "A Spatial Filtering Approach to Texture Analysis," Pattern Recognition Letters, v. 3, 1985, pp. 195-203.