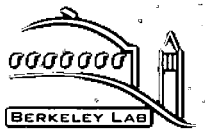
LBNL-42119
Preprint



ERNEST ORLANDO LAWRENCE
BERKELEY NATIONAL LABORATORY

Segmentation of Confocal Microscope Images of Cell Nuclei in Thick Tissue Sections

Carlos Ortiz de Solórzano, Enrique García Rodriguez, Arthur Jones, Dan Pinkel, Joe W. Gray, Damir Sudar, and Stephen J. Lockett

Life Sciences Division

July 1998

Submitted to
Journal of Microscopy

DISCLAIMER

Segmentation of Confocal Microscope Images of Cell Nuclei in Thick Tissue Sections

[a]Carlos Ortiz de Solórzano, [a]Enrique García Rodriguez, [a]Arthur Jones, [b]Dan Pinkel, [b]Joe W. Gray , [a]Damir Sudar and [a]Stephen J. Lockett

[a]Ernest Orlando Lawrence Berkeley National Laboratory
University of California
Berkeley, CA
USA

[b]Cancer Center
University of California at San Francisco
San Francisco, CA
USA

Address of Correspondence:

Carlos Ortiz de Solorzano
Lawrence Berkeley National Laboratory, MS 84/171
1 Cyclotron Road
Berkeley, CA 94720
USA
Phone: (510) 486 5422
Fax: (510) 486 7209
email: carlos@rmc.lbl.gov

Supported by the Director, Office of Energy Research, Office of Health and Environmental Research of the U.S. Department of Energy under contract NO.DE-AC03-76SF00098, NIH grant CA-67412, a contract with Carl Zeiss Inc. and the University of California Systemwide Biotechnology Research and Education Program for Training in Genome Research.

ABSTRACT

Segmentation of intact cell nuclei from three dimensional (3D) images of thick tissue sections is an important basic capability necessary for many biological research studies. However, segmentation is often difficult because of the tight clustering of nuclei in many specimen types. We present a 3D segmentation approach that combines the recognition capabilities of the human visual system with the efficiency of automatic image analysis algorithms. The approach first uses automatic algorithms to separate the 3D image into regions of fluorescence-stained nuclei and unstained background. This includes a novel step, based on the Hough transform and an automatic focusing algorithm to estimate the size of nuclei. Then, using an interactive display, each nuclear region is shown to the analyst, who classifies it as either an individual nucleus, a cluster of multiple nuclei, partial nucleus or debris. Next, automatic image analysis based on morphological reconstruction and the watershed algorithm divides clusters into smaller objects, which are re-classified by the analyst. Once no more clusters remain, the analyst indicates which partial nuclei should be joined to form complete nuclei. The approach was assessed by calculating the fraction of correctly segmented nuclei for a variety of tissue types: Caenorhabditis *elegans* embryos (839 correct out of a total of 848), normal human skin (343/362), benign human breast tissue (492/525), a human breast cancer cell line grown as a xenograft in mice (425/479) and invasive human breast carcinoma (260/395). Furthermore, due to the analyst's involvement in the segmentation process, it is always known which nuclei in a population are correctly segmented and which not, assuming that the analyst's visual judgement is correct.

Key words: Confocal microscopy, image segmentation, image analysis, three dimensional analysis.

1. INTRODUCTION

Cytology shows that cells in a tissue become increasingly heterogeneous in their structural properties during carcinogenesis, while histology shows increasing disorganization of the cells. In order to understand the underlying molecular mechanisms of these structural alterations, it is necessary to analyze the cells individually and within their natural tissue context. Since many of the structural and molecular changes occur within the cell's nucleus, the particular ability to segment individual nuclei in intact tissue is an important technical capability.

To obtain quantitatively accurate measurements of individual nuclei, the nuclei must be intact. However, using standard 4 to 6 μm thick sections only an arbitrary fraction of each nucleus is present. To overcome this problem, thick (> 20 μm) sections must be employed, but this requires fluorescence to keep tissue transparent, three-dimensional (3D) confocal microscopy image acquisition (Wilson, 1990) and 3D image analysis.

In order to facilitate the segmentation of nuclei from images, the specimen is labeled with a fluorescent DNA stain. Because of the abundant DNA in the nuclei, staining produces very high contrast images consisting of high intensity nuclear regions and low intensity non-nuclear (background) regions (Lockett *et al.*, 1991).

Segmentation of nuclei can be achieved by interactive or automatic algorithms. Interactive methods, based on drawing around nuclei in sequential (Czader *et al.*, 1996; Rigaut *et al.*, 1991) or orthogonal (Lockett *et al.*, 1998) 2D slices are superior in performance (defined as the fraction of nuclei correctly segmented based on visual judgement) compared to automatic algorithms. However, they are slow, typically taking minutes per nucleus and are thus limited in their practical application to situations where only tens of nuclei require analysis. Rodenacker (Rodenacker *et al.*, 1997) presented a less interactive method, which was based on thesholding followed by 3D volume visualization and interactive division of clusters of nuclei. Division was performed by manual marking of the centers of nuclei, followed by automatic growing from the markers limited by size and shape constrains, and final manual refinement by fitting the nuclei to ellipsoidal model. However, interaction is still considerable for densely packed nuclei and the shape

constraints imposed could produce undesirable results on samples with highly irregular nuclear shape, such as those which occur in cancer tissue.

Automatic algorithms on the other hand are much faster, enabling convenient analysis of 100s to 1000s of nuclei per study, but they do not correctly segment all of the nuclei. For example, Rigaut (Rigaut *et al.*, 1991) developed a segmentation method composed of gray level opening, top hat transformation, geodesic reconstruction, particle deagglomeration by the watershed algorithm, and final discrimination by adaptive thresholding to segment 3D tissue samples of normal rat liver and in situ carcinoma of the esophagus. The algorithm performed well for the separated nuclei characteristic of liver samples, but performance deteriorated for cancerous tissue where nuclei were clustered, and caused them to use an interactive method. Ancin et al (Ancin *et al.*, 1996) presented a segmentation that first performed thresholding, followed by refinement of the results using a split and merge algorithm and then a watershed to divide clusters of nuclei. The authors reported 92% correct segmentation on rat liver, but did not report results for the hipocampal brain region of the rat, where nuclei are more densely packed. Irinopoulou (Irinopoulou *et al.*, 1997), presented a 3D segmentation method that used global thresholding followed by morphological filtering and segregation by the watershed algorithm to segment each 2D slice of the 3D image. Then, nuclei deagglomeration in the Z (depth) direction was performed based on the convexity criterion of the object contours. It was tested on four different types of prostatic lesions: hyperplasia, prostatic intraepithelial neoplastia, well-differentiated and poorly differentiated carcinoma. Since the study focused on the measurement of DNA volume, quantitative results about the number of correctly segmented nuclei were not given. In summary, the performance of the automatic methods is only high (> 90%) for specimens containing well-separated nuclei. Performance significantly deteriorates for many cancer specimens, because the cells are structurally dominated by their nuclei leaving little separating cytoplasm and thus the images show clustered nuclei and because nuclei are more variable in their sizes and shapes.

Here we report a 3D segmentation approach designed to deal with the problem of clustered nuclei in cancer specimens. It combines the recognition capabilities of the human visual system with the efficiency of computational image analysis algorithms, thus

achieving the correct segmentation of a high proportion of individual nuclei in a variety of intact tissue specimens.

2. MATERIALS AND METHODS

2.1. Segmentation Algorithm

2.1.1 Outline of the algorithm

The main steps of the segmentation are summarized in Figure 1. The input image is a 3D confocal microscope image of the fluorescent DNA staining of the specimen. The image is composed of a stack of 2D images acquired at equidistant planes through the specimens. First, the 3D image is automatically segmented into regions corresponding to nuclear (positive staining) and non-nuclear, background regions (no staining). Next for each nuclear region (object) its size and shape is measured and then it is classified by visual inspection. Those objects classified as being clusters of multiple nuclei are split into smaller objects using an automatic cluster segmentation algorithm and are returned for measurement and reclassification.

2.1.2 Software environment

The algorithms were programmed in ANSI C language and compiled into SCIL_Image v. 1.3 (TNO Institute of Applied Physics & University of Amsterdam, The Netherlands) under UNIX Operating Systems (Solaris v. 2.6 and IRIX v. 5.3). SCIL_image provided basic algorithms for point operations, linear filtering, morphological operations, etc. All the algorithms were anisotropic in that they accounted for the difference in the physical size of a voxel in the axial direction versus the lateral directions. This avoided increased computation that would be associated with working with interpolated images with equaly sized voxels in the axial and lateral directions. The eigenanalysis for some of the object measurements (see below) was done using the Linear Algebra Package (LAPACK) v. 2.0 (Anderson *et al.* 1994). The 3D visualization and user interface software was written in C++, using the Motif widget toolkit, a widget scripting language developed in house and the OpenGL library (Sun Microsystems, Mountain View, CA). For visualization purposes, the images were interpolated with third order polynomials to make them isotropic.

We developed a linked-list style structure (see appendix) for databasing the acquired images along with the intermediate and final results of the segmentation process. The structure enabled increased efficiency in the visualization and cluster division steps in part

by restricting analysis to only the relevant parts of the image, and by enabling the segmentation process to be interrupted at any stage..

2.1.3 Automatic Segmentation

Figure 2 shows the sequence of image analysis steps for automatically separating the acquired image into nuclear objects and non-nuclear background. Each of these steps is described below.

Median Filtering. Median filtering with a 3*3*3 kernel was applied to the original image to remove shot noise, which was introduced by the photomultiplier tubes of confocal microscopes.

Estimation of the average radius of nuclei. An estimate of the average radius of the nuclei was needed for subsequent steps in the analysis. It was calculated directly from the median filtered 3D image using an iterative procedure where each iteration was a combination of two steps: the first step was based on the *Hough Transform* (HT) (Ballard, 1981; Lockett *et al.* 1997) and attempted to shrink nuclei. The second was an automatic focusing algorithm, which measured the average amount of shrinkage of the nuclei present in the image. The iterative procedure searched for the maximal level of shrinkage, from which the average radius of the nuclei could be estimated.

Shrinking was done as follows for the case of the image having isotropically-sized voxels and the spatial resolution being isotropic. First the magnitude and direction of the gradient of the intensity at every voxel was calculated using the 1st order Gaussian Derivative filter (van Vliet, 1991) applied in the x,y and z directions. The magnitude is high at significant intensity transitions, i.e. at nuclear surfaces, and low in the background and interior of nuclei. Then, starting with an initially empty new image and a guess of the radius, (R) of the nuclei, for each voxel, the magnitude of the gradient is added to the intensity of the voxel in the new image at a distance R from the voxel in the original image in the direction of the gradient.

However in our images voxels were anisotopic in size and the spatial resolution was also anisotropic. The latter is a property of the confocal microscope. Thus the direction of the gradient could not be determined directly from the output of the 1st order Gaussian filter, and because of the anisotropic voxel size, the value of R required adjustment,. The components of the direction of the gradient in the lateral (x & y) directions, G_x and G_y were corrected for the anisotropioc voxel size by dividing them by the square of the relative voxel size (r) in the lateral direction (relative to the voxel size in axial (z) direction) to produce two corrected components G'_x and G'_y. R was adjusted for the anisotropic voxel size using the equation

$$R' = R\sqrt{\frac{2r^4 + w}{2r^2 + w}} \quad (1)$$

where $w = G_x^2 + G_y^2$. The magnitude of the gradient was adjusted to take account of the anisotropic Point Spread Function (PSF) of the microscope. These adjustment were determined empirically by applying the shrinkage to computer-generated images of spheres that had been convolved with an estimate of the PSF. The estimated PSF was a Gaussian kernel of standard deviation (σ) equal to 0.5 voxels in the lateral directions and 1.25 voxels in the axial direction. This kernel corresponded roughly to the anisotropic PSF of our image acquisition system. The images of the spheres after convolution were interpolated in the axial direction to create the actual anisotropy in pixel size. Shrinkage was applied to this image with different values of filter width in the lateral and axial directions (σ'_{xy}, σ'_z) of the 1st order Gaussian filter until we obtained a combination that gave an isotropic response in the output image.

This last showed that the correct filter widths for the 1st order Gaussian filter were related to the filter widths of the Gaussian filter simulating the PSF by the following equation

$$\sigma'_{xy} = \frac{\sigma_z}{\sigma_{xy}} \cdot \frac{1}{r} \cdot \sigma'_z \quad (2)$$

where σ'_{xy} and σ'_z are the filter widths of the 1st order Gaussian filter in the lateral, and axial directions, and $\sigma_{x,y}$ & σ_z are the widths for the Gaussian simulating the PSF. The experimental results on images of actual nuclei confirmed this relationship.

The average amount of shrinkage achieved by R was determined by measuring the degree of "focus" of the output image. The measure of focus was taken as the sum of the gradient calculated by the *Sobel* operator (Castleman 1979a) magnitudes squared at every voxel in the output image (Yeo *et al.*, 1993). Then a binary iterative search was undertaken to locate the value of R, R_{OPT} which produced the maximal degree of focus and thus the maximal degree of shrinkage. R_{OPT} equaled the average radius of the nuclei.

Figure 3 illustrates this procedure for estimating the average radius of nuclei. Figure 3a is a 2D slice from a median-filtered confocal image of the breast cancer cell line grown in mice (see below). Figure 3b shows the magnitude of the response from the 1st order Gaussian filter. Figures 3c, 3d & 3e show respectively the output images for $R<R_{OPT}$, $R=R_{OPT}$ and $R>R_{OPT}$. Figure 3f is a graph of the "focus" values as a function of R of the image in figure 3a. If R_{OPT} was known from the previous analysis of similar specimens, it could be entered as an input parameter to the overall segmentation procedure and this step would be omitted.

Threshold-based Segmentation. Intensity thresholding was considered appropriate for segmenting the nuclear regions because of the high voxel intensities in nuclei labeled with fluorescent DNA stain versus the low intensity of the background. However, the intensity of nuclear regions was not constant across the image, especially in the axial direction because of increasing photobleaching and increasing spherical aberration caused by the refractive index of the specimen being less than oil. This prevented a single, global, threshold being applied to the whole image. Therefore, an adaptive thresholding algorithm was used (Lockett *et al.*, 1991). Adaptation was achieved by dividing the image into smaller, cuboidal volumes, and calculating a single threshold for each volume. The size of each volume (in voxels) was ($3R_{OPT}$ * $3R_{OPT}$ * 1) in the (x, y, z) directions. This size was large enough in the lateral directions to ensure that each cuboidal volume usually contained at least one nucleus and some background, which is necessary for the correct threshold to be calculated, yet small enough in the z direction to account for changing nuclear intensity.

The gradient weighted threshold T (MacAulay & Palcic, 1988) was calculated for each volume. $T_{a,b,z}$, the threshold for the a^{th} volume in the x direction, the b^{th} volume in the y direction and at the slice z, *was*

$$\frac{\sum_{x=3aR_{OPT}}^{3(a+1)R_{OPT}} \sum_{x=3bR_{OPT}}^{3(b+1)R_{OPT}} I_{x,y,z} \cdot G_{x,y,z}}{\sum_{x=3aR_{OPT}}^{3(a+1)R_{OPT}} \sum_{x=3bR_{OPT}}^{3(b+1)R_{OPT}} G_{x,y,z}} \qquad (3)$$

where (x, y, z) is the voxel coordinate in the image, I is the voxel intensity from the median-filtered image and G is the gradient magnitude calculated from the median-filtered image using the *Sobel.*

Occasionally, volumes contained only background and no nuclei. For these volumes, the calculated threshold was approximately equal to the mean intensity of the background in the volume, and thus some of the voxels were incorrectly classified as object. In order to avoid this error, for each volume, the mean intensity, mean weighted intensity (numerator in equation 3), and mean gradient (denominator in expression 3) were calculated and compared to the average of these means over all the volumes of the image. If all three values were lower than the averages, the threshold value for the volume was recalculated as being the average threshold of the neighboring volumes where this condition was not met.

In order to obtain a gradually changing threshold across each lateral slice image, the values $T_{a,b,z}$ were assumed to be true only at the center voxel of their respective volume and bi-linear interpolation was used to calculate a threshold at all the other voxels. Actual thresholding was accomplished by setting all voxels with intensities above their threshold intensity to *1* and all others to *0*. Holes inside the nuclear regions, presumably corresponding to nucleoli, were filled. This was done by setting the voxels of background, non-connected regions whose areas were much less than that of a nucleus,to object voxels.

Morphological Filtering. The binary image obtained above was morphologically filtered to eliminate small objects which were debris and to divide slightly touching clusters of nuclei. First binary erosion was applied using an ellipsoidal kernel a third of the average size of the nuclei in each dimension. This size guaranteed substantial erosion without loss

of the nuclei, but with loss of small debris. Next, the skeleton (Verwer, 1988) of the eroded binary image was calculated to find bisecting surfaces between slightly touching nuclei that had been separated by the erosion. These surfaces were superimposed on the original binary image to split slightly touching nuclei. The result was a set of binary, unclassified objects that could represent single nuclei, clusters of nuclei or large pieces of debris.

2.1.4 Object Measurement

The following parameters were measured for each of the binary objects:

Volume, equal to the total number of voxels inside the object.

Surface Area, which was measured accurately a method based on Chamfer metrics (Mullinkin & Verbeek 1993).

Shape Factor. This parameter indicates how spherical the object is.

$$Shape\,Factor = \frac{(Surface\ Area)^{\frac{3}{2}}}{Volume} \qquad (4)$$

Center of Mass. This provided a central reference point for each object. The coordinates of the center of mass are,

$$\left(\frac{\sum_{i \subset Object} x_i}{Volume}, \frac{\sum_{i \subset Object} y_i}{Volume}, \frac{\sum_{i \subset Object} z_i}{Volume} \right) \qquad (5)$$

where x_i, y_i, z_i are the image coordinates of voxel i.

The *Eccentricity* measures the combined elongation and flattening of the object relative to a sphere. We used the following functional definition of eccentricity:

$$Eccentricity = \frac{2\sqrt{L_P}}{\sqrt{L_S} + \sqrt{L_T}} \qquad (6)$$

where L_P, L_S, *and* L_T ($L_P \geq L_S \geq L_T$) are the eigenvalues calculated from eigenanalysis on the matrix of second-order central moments of the binary object (Lo *et al.*, 1989). The eccentricity of a sphere is 1. Objects with $L_P = L_S$ have eccentricity values ranging from 1 to

2 ($L_T = 0$ represents a perfectly flat disk). Objects with $L_S = L_T$ have eccentricities between 1 and ∞ ($L_S = L_T = 0$ represents a straight line)

2.1.5 Visual Classification.

After segmentation and measurement, the objects were rendered using our software program called DAta Visualization aNd Computer Interaction (*daVinci*). The user interface to daVinci is shown in Figure 4.

DaVinci interpolates objects to make them spatially isotropic and displays them to the user. Each "unclassified" object in turn is displayed and the user classifies it as a "nucleus", "nucleus on the edge" of the image, "cluster of nuclei", "cluster on the edge" or "debris". To help the user, several rendering options are available: objects can be freely rotated by mouse interaction in the OpenGL window (top left part of figure 4); automatic ("Fit to window" in the "Rendering options" window in the top right of figure 4) and manual can be adjusted, and the option to simultaneously overlay any intersecting *xy*, *xz* or *yz* slice from the original acquired image is available. This is important when the nuclei are tightly clustered, leaving little evidence in terms of the shape of the surface rendering, and when the surface rendering is suspected of not accurately representing the nuclear surfaces. In addition, the objects surrounding the current object can be displayed ("Show neighbors"). This is useful for reclassifying objects after cluster division (section 2.1.6) when individual nuclei might have been erroneously divided into several objects.

The user makes the classification decision based on visual examination of the shape of the surface rendered object in combination with the intersecting slices and using the measurements of the object ("Info" window in the bottom left corner). Then the user informs the program of the decision by clicking a button in the "Classify" window in the bottom right of figure 4. The program provides and "Undo" option under the "Edit" menu so that the user can correct their errors and an option under the "Cell" menu to rejoin objects erroneously divided by the cluster division algorithm. Further information about the interface options is given in the caption of Figure 4.

2.1.6 Cluster Division

After visual classification, clusters of nuclei were divided into individual nuclei using a two-step algorithm. Figure 5 is a flow diagram of the cluster division. In the first step, the algorithm attempted to find an internal marker for each nucleus in the cluster.

Then, in the second step the algorithm attempted to find a surface between marked nuclei in the cluster. After division, the resulting objects were classified by the user (Section 2.1.5) as an individual nucleus, a cluster of nuclei (in which case cluster division would be repeated) or a partial nucleus (in which case it would be later rejoined to its counterparts). This cycle was repeated until no clusters remained.

The method to calculate nuclear markers was a 3D extension of the method of (Malpica *et al.*, 1997). First, the *Distance Transform* (DT) of the object (see Figure 5) was determined using a modified vector distance transform (Mullikin, 1992) to avoid both the expensive computational load of the Euclidan distance transform (Castleman, 1979b) and the errors associated with the use of the city-block distance transform (Borgefors, 1983).

We use the peaks of the DT image to indicate the centers of nuclei. However, this image contained many small, noise peaks that arose from the rough surface of the cluster. These noise peaks would ultimately lead to the cluster dividing into a large number of small objects rather than a few nuclei. Thus the DT image was smoothed by *Gaussian* to leave only significant peaks. The standard deviation of the filter was one third of R_{OPT}. At this stage, a new estimate of the radius of the nuclei inside the cluster could be calculated as the value of the peaks in the filtered DT. When only one peak was found, the above process was repeated, with a standard deviation of the Gaussian filter of half the previous value.

Surfaces between peaks were found using the watershed algorithm (Beucher & Meyer, 1992) where the peaks were used as the initiation points for "flooding" either the inverse of the acquired image after median filtering or the DT image before filtering. The rationale for using the inverse of the original image after median filtering was that voxels just outside the surfaces of nuclei should have low intensity (high intensity in the inverse image) because the DNA stain should be absent at these points. However that is not necessarily true if the nuclei are closely clustered. In those cases, the morphological information encoded in the DT image was used, since it shows bright peaks at the centers of the nuclei and 'dark' necks between objects. We normally used the inverse of the original image, and only used the DT if the cluster did not divide.

2.2 Evaluation of the Segmentation Procedure

2.2.1 Evaluation using Computer-Generated Objects

Evaluation of the algorithm using computer generated objects was carried out to establish the accuracy of the segmentation and its robustness against noise.

Six objects were generated. See (Lockett *et al.*, 1998) for details of the first five. These were two spheres of radius 10 and 25 voxels, a curved disk, a normal ellipsoid and a curved ellipsoid. The sixth object consisted of another curved disk. It was created by intersecting two of the bigger spheres, whose centers were 10 pixels apart and selecting the smallest of the not-intersecting volumes. The objects were of the equivalent size of the nuclei, where the voxel size was 0.2μm. The objects were blurred using a Gaussian filter of standard deviation ($\sigma_{x,y}$ = 0.1μm, σ_z = 0.25μm), which roughly approximated the blurring effect of the PSF of the confocal microscope. Next we subsampled the objects by a factor of two in the axial direction, to simulate a voxel size of 0.2μm in the lateral direction and 0.4μm in the z direction. Then we added Gaussian noise such that the Signal to Noise Ratio (SNR) equaled 3.5, which was the same as the SNR of the tissue images. The objects were then segmented and the accuracy of the segmentation was determined by calculating the average distance between the true surface of the objects and the surface given by the segmentation.

Next, we determined the limits of our cluster division algorithms, by calculating the minimum distance between clustered objects required for the algorithm to divide it. We prepared two computer generated clusters consisting of two "nuclei" each. One was composed of two truncated spheres facing each other at their flat side and the other was the larger sphere and the larger of the two curved disks. These two clusters were representative of cases found in real specimens. The first, composed of two truncated spheres (Fig. 6a & 6b), could be envisioned as two nuclei flattened in their contact surface, with cytoplasm in between. The second (Fig. 6c & 6d) is an extreme situation representing a cluster of two nuclei where one is wrapped around the other. The increasing size of the gap between the surfaces in each cluster represented different degrees of clustering of the nuclei. The interface surface between the objects in the clusters was first parallel to the *xy* plane and then parallel to the *yz* plane and the spacing between the two parts of each cluster ranged

from 0 to 10 voxels. Since the voxel size was 0.2μm, the distance between objects ranged from 0.2 to 2μm. Figure 6 shows the clusters.

The clusters were subsampled and anisotropic Gaussian filtered as described above for the single objects, but in this case we added three different amounts of Gaussian Noise, which corresponded to no noise (SNR=∞), SNR=3.5 and 2. The goal of this experiment was to find the minimum distance necessary for correct segmentation of the cluster, as a function of the SNR.

2.2.2 Evaluation using Tissue Specimens

Five types of thick tissue specimens were used that were considered to range from easy: *Caenorhabditis elegans (C. elegans)* embryos which had nuclei that were spherical in shape and well separated, to difficult: invasive carcinoma of the human breast, where nuclei were highly irregular in shape and size and highly clustered. Intermediate specimens were normal human skin, a human benign breast tumor, and a human breast cell line grown as a xenograft in a mouse.

Several hundred nuclei from each specimen type were imaged and performance was measured as the fraction of nuclei correctly segmented, based on visual judgment.

Sample Preparation and Image Acquisition

C. elegans embryos (CE) were obtained from 10-20 adult worms gravid with embryos. An incision was made at the vulva to release the embryos, which were then prepared and fixed as explained in (Chuang *et al.*, 1994). After fixation, the embryos were treated with RNase and stained by incubation with the DNA dye propidium iodide (PI) (2 μg/ml) for 1 hour to label the nuclei. Then the embryos were washed and mounted for confocal microscopy. The embryos were 30 μm thick and contained between 50 and 300 cells.

Normal human skin specimens (NS) were obtained from the archives of the Dermatopathology Section of the Departments of Pathology and Dermathology, University of California, San Francisco.

Human breast specimens, which contained benign parts (BP) and invasive carcinoma parts (IC) were obtained from the Department of Pathology, California Pacific Medical Center (CPMC), San Francisco.

The skin and breast specimens had been fixed in 10% neutral buffered formalin and paraffin-embedded before receipt. They were cut into 20 μm sections and stained with PI at 0.1 μg/ml. In addition, the centromeric regions of chromosome 1 were also labeled using fluorescence in situ hybridization (FISH) (Thompson *et al.*, 1994), although the analysis of those signals is outside the scope of this paper. However, the FISH protocol did cause some degradation of the nuclei.

Formalin-fixed, paraffin-embedded MCF7 cells (a human breast cancer cell line, BC) that had been grown in nude mice as a xenograft were provided by Dr. Gail Colbern (Geraldine Brush Cancer Research Institute, CPMC). They were cut to 30μm thickness, FISH was performed using probes for chromosome 1 centromeric region and the 20q13.2 locus and they were counterstained with 100μl of 1μM YO-PRO-1 (Molecular Probes, Eugene, OR, USA) containing 1mg/ml RNase A in phosphate buffer and incubated in the dark at 37°C for 2 hours. Afterwards the sections were washed twice in phosphate-buffered saline at room temperature for 5 minutes and rinsed in distilled water. After air drying, the sections were covered with a coverglass (No.1), and sealed with nail polish.

Specimens CE, BP, IC & BC were imaged using a laser scanning confocal microscope 410 (Carl Zeiss Inc., Thornwood,NY) equipped with an Axiovert 100 microscope (Zeiss), a 63X, 1.4 NA plan-APOCHROMAT objective lens (Zeiss) and an Argon/Krypton (Ar/Kr) laser. NS was imaged with an MRC-1000 confocal imaging system (Bio-Rad Microscience Ltd., Hemel Hempstead, England) equipped with a Diaphot 200 microscope (Nikon Inc., Instrument Group, Garden City, NY), a 60X, 1.4 NA planapo objective lens (Nikon) and an Ar/Kr laser.

The PI in specimens CE, NS, BP and IC was imaged using the 568 nm laser line and collecting emissions longer than 590 nm. The YO-PRO-1 in BC was exited using the 488 nm laser and emissions were detected using a band-pass filter in the range 515-565 nm.

The distance between adjacent 2D slices was 0.3 μm for CE and BC and 0.5 μm for the others. The voxel size was 0.1 in the x and y dimensions for CE, 0.2 μm for BP, IC and BC, and 0.18 μm for NS.

The acquired images were stored in the ICS image format (Dean *et al.*, 1990) and transferred to a UNIX workstation for archiving and analysis.

Figure 7 shows lateral slices from each the specimen types to illustrate the differences in tissue organization, cellular homogeneity and nuclear size and shape.

3. RESULTS

3.1 Results of the experiments with computer-generated objects

The results of the study using computer-generated objects to assess the performance of the segmentation procedure are in tables 1 to 3.

Table 1 shows that the average distance between the actual and segmented surfaces of the isolated objects was always less than 1 voxel (0.2 μm), suggesting that nuclei were accurately segmented by the procedure. In all cases the average distance was at or below the limiting spatial resolution.

Table 2 shows the minimum distance between the two objects in each cluster that was necessary to achieve segmentation at various noise levels. The cluster composed of two truncated spheres was segmented, no matter the distance between the objects, including the 'no distance' situation. This is due to the fact that the binary mask of the clusters provided by the thresholding algorithm always contained a "neck" at the interface of the objects which the cluster division step could detect. The other cluster required a minimum distance of up to 1.6 μm. At closer distances, the binary mask of the cluster generated was purely convex and only one nuclear marker could be extracted. The table also shows the influence of the increasing noise and the orientation of the object interface in the cluster relative to the image.

Table 3 shows the accuracy of the surfaces at the interface of touching objects in the cluster for the intermediate noise level (SNR=3.5). Even though the accuracy is not as high as in the case of isolated objects, it still remains close to the resolution limit.

3.2 Results of the experiments with tissue specimens

The performance of the segmentation procedure based on visual judgment for tissue specimens is shown in Table 4. It shows, for each type of specimen, the number of correctly segmented nuclei, the number of nuclei in cluster that could not be divided, the number of nuclei that were lost due to poor DNA staining, and the number of objects that the user could not identify as either individual, clustered nuclei or debris.

Segmentation was considered highly accurate for the C. elegans specimens (Correct nuclei=99%). This was believed to be caused due to their high nuclear homogeneity in terms of shape and relatively large separations between nuclei.

The segmentation accuracy for normal skin and benign breast tissue was 94.7 % and 93.7 % correctly segmented nuclei respectively. Most clusters that could not be divided touched at least one of the edges of the image, such that an incomplete nucleus was present. These partial nuclei were not sufficient to generate a nuclear marker in the cluster division step. We propose a solution to this problem in the discussion. Other clusters were so compact that no suitable peaks could be extracted. Images in Figure 8 are a stereo pair of segmented normal skin nuclei.

In the specimens of cancer cells grown in mice, 88.7 % of the nuclei were correctly segmented. In these specimens, the number of nuclei remaining in clusters was higher, due to higher compactness of the nuclei and to the existence of unstained zones inside the nuclei, presumably nucleoli, which generated incorrectly placed nuclear markers in the cluster segmentation algorithm. A hole-filling step was performed after thresholding, but in some cases this step filled some background areas between nuclei, making the segmentation of those clusters more difficult.

The proportion of correct nuclei dropped to 65.8 % for invasive carcinoma. This was caused by the high degree of clustering and increased heterogeneity of the nuclear shape and size. The images of these specimens showed many nuclear regions where the individual nuclei could not be recognized by visual examination. These were classified as undefined objects.

4. DISCUSSION

A new program for segmenting DNA stained nuclei from 3D confocal images has been developed and tested. In this paper we have described the steps that comprise the program, and we have presented the results of the evaluation of the algorithm using computer generated objects and using five different types of tissue specimens that dramatically varied in their complexity.

The program performs automatic segmentation, but includes an interactive classification step, where the users confirms, corrects and classifies the results of the segmentation. By taking this approach, user interaction was kept to a minimum, by only requiring one mouse click per object to accomplish the classification task. Most importantly however, we always obtain a subset of nuclei that we know are correctly segmented (based on visual judgment).

Our algorithm is highly accurate, as shown by the experiments with computer generated objects, since all the surfaces were segmented with an average error below the spatial resolution level. The results with tissue specimens were obtained from analysis of 2548 nuclei, and show that our algorithm correctly segmented a greater fraction of nuclei than previously published methods.

The significant differences between this algorithm and the previously published ones are: 1) the thresholding used to define regions of the image containing nuclei is adaptative, and therefore could vary across the image; 2) the filter size for the morphological filtering is related to the size of the nuclei, and is automatically estimated from the image using a combination of a Hough-like transform of the original image and an automatic focusing method; 3) Objects were classified visually by the operator, which provides higher accuracy than automatic classification.

In the future work, we will refine the segmentation for highly clustered nuclei. Our initial approach consists of the use of a cluster segmentation method based on the Hough transform. This approach combines edge information from the original image with *a priori* shape information about the nuclei, in a procedure to shrink and thereby separate objects (Lockett *et al.*, 1997). Initial tests of this method show that it identifies a greater proportion

of individual nuclei than our current method, but its utility is limited by imprecise surface definition. To overcome this limitation, we plan to apply denoising and shape recovering methods based on anisotropic diffusion and level set curvature flow (Malladi & Sethian 1995), which have been shown to provide significant improvement in a preliminary study (Sarti *et al.*, 1998). A second strategy, for highly complex clusters will include the use of specific nuclear surface markers (lamin antibodies) to explicitly delineate nuclear surfaces.

5. ACKNOWLEDGEMENTS

We thank Mr. Carlos Fernandez and Mr. Anton Rutten for their work on the early stages of this research; Dr. Gail Colbern for providing the MCF7 cell line, Dr. Koei Chin for preparing and staining the xenografts, Mr. Ramin Khawan for preparing the normal skin and human breast tissue and Dr. Jason Lieb for providing the C. *elegans* embryos, testing the segmentation software and for his suggestions that helped us improving the visualization program (daVinci).

REFERENCES

Ancin H., Roysam B., Dufresne T.E., Chesnut M.M., Ridder G.M., Szarowski D.H., Turner J.N. (1996) Advances in Automated 3-D Image Analysis of Cell Populations Imaged by Confocal Microscopy. *Cytometry* **25**, 221-234.

Anderson E., Bai Z., Bischof C., Demmel J., Dongarra J., Du Croz J., Greenbaum A., Hammarling S., McKenney A., Ostrouchov S., Sorensen D. (1994) LAPACK Users' Guide - Release 2.0, ed. Society for Industrial and Applied Mathematics (SIAM), September 1994.

Ballard D.H. (1981) Generalizing the Hough Transform to detect arbitrary shapes. *Pattern Recognition*, **13**(2), 111-22.

Beucher S., Meyer F. (1992) The morphological approach to segmentation: The watershed transformation. In: Mathematical Morphology in Image Processing. E. Dougherty (ed.). Marcel Dekker Inc., pp. 433-481.

Borgefors G. (1983) Chamfering: A fast methods for obtaining approximations of the Euclidian distance in N dimensions. In Proceedings, 3rd Scandinavian Conference on Image Analysis, Copenhagen, Denmark, 250-255.

Castleman K.R. (1979a) Digital Image Processing. Prentice-Hall, Englewood Clliffs, New Jersey, 1979, p.465.

Castleman K.R. (1979b) Digital Image Processing. Prentice-Hall, Englewood Clliffs, New Jersey, 1979, p.475.

Chuang P.T., Albertson D.G., Meyer B.J. (1994) DPY-27: a chromosome condensation protein homolog that regulates C. *elegans* dosage compensation through association with the X chromosome. *Cell* **79**(3), 459-74.

Czader M., Liljeborg A., Auer G., Porwit A. (1996) Confocal 3-Dimensional DNA Image Cytometry in Thick Tissue Sections. *Cytometry* **25**, 246-253.

Dean P., Mascio L., Ow D., Sudar D., Mullikin J. (1990) Proposed standard for image cytometry data files. *Cytometry* **11**, 561-569.

Irinopoulou T., Vassy J., Beil M., Nicolopoulou P., Encaoua D., Rigaut J.P. (1997) Three-Dimensional DNA Image Cytometry by Conical Scanning Laser Microscopy in Thick Tissue Blocks of Prostatic Lesions. *Cytometry* **27**, 99-105.

Lo C.H., Don H.S.(1989) 3-D Moment Forms: Their Construction and Application to Object Identification and Positioning. *IEEE Transactions of Pattern Analysis and Machine Intelligence* **11**(10), 1053-1064.

Lockett S.J., O'Rand M., Rinehart C., Kaufman D.G., Herman B., Jacobson K. (1991) Automated fluorescence image cytometry: DNA quantification and detection of chlamydial infections. *Analytical and Quantitative Cytology and Histology* **13**, 27-44.

Lockett S.J., Rodriguez E.G., Ortiz de Solórzano C., Sudar D., Pinkel D., Gray J.W. (1997) Automatic Hough Transform-based 3D segmentation of cell nuclei in thick tissue sections. Proceedings of Microscopy and Microanalysis'97, Cleveland, Ohio, August 10-14, 1997. *Microscopy and Microanalysis* **3** (supp 2), 1121-1122.

Lockett S.J., Sudar D., Thompson C.T., Pinkel D., Gray J.W. (1998) Efficient, interactive, three-dimensional segmentation of cell nuclei in thick tissue sections. *Cytometry* **31**, 275-286.

MacAulay C., Palcic B., (1988) A comparison of some quick and simple threshold selection methods for stained cells. *Analytical and Quantitative Cytology and Histology* **10**, 134-138.

Malladi R., Sethian J.A. (1995) Image processing via level set curvature flow. *Proceedings of the National Academy of Sciences USA*, **92**, 7046-7059.

Malpica N., Ortiz de Solorzano C., Vaquero J.J., Santos A., Vallcorba I., Gracia-Sagredo J.M., del Pozo F. (1997) Applying watershed algorithms to the segmentation of clustered nuclei. *Cytometry* **28**, 289-297.

Mullikin J.C. (1992) The vector distance transform in two and three dimensions. *CVGIP: Graphical Models and Image Processing* **54**(6), 526-535.

Mullikin J.C., Verbeek P.W. (1993) Surface area estimation of digitized planes. *Bioimaging* **1**(1), 6-16.

Rigaut J.P., Vassy J., Herlin P., Duigou F., Masson E., Briane D., Foucrier J., Carvajal-Gonzalez S., Downs A.M., Mandard A-M. (1991) Three-Dimensional DNA Image Cytometry by Confocal Scanning Laser Microscopy in Thick Tissue Blocks. *Cytometry* **12**, 511-524.

Rodenacker K., Aubele M., Hutzler P., Umesh Adiga P.S. (1997) Groping for quantitative digital 3-D image analysis: an approach to quantitative fluorescence in situ hybridization in thick tissue sections of prostate carcinoma. *Analytical Cellular Pathology* **15**,19-29.

Sarti A., Ortiz de Solorzano C., Lockett S.J., Malladi R. (1998) A Unified Geometric Model for 3D Confocal Image Analysis in Cytology. Accepted for SIBGRAPI'98

(International Conference on Computer Graphics, Image Processing and Vision) Rio de Janeiro, Brazil, October 20-23th , 1998.

Thompson C.T., LeBoit P.E., Nederlof P.M. and Gray J.W. (1994) Thick-Section Fluorescence in Situ Hybridization on Formalin-Fixed, Paraffin-Embedded Archival Tissue Provides a Histogenetic Profile. *Am. J. Patholol.* **144**, 237-243

Verwer B.J.H. (1988) Improved metrics in image processing applied to the Hilditch skeleton. 9th International Conference on Pattern Recognition (IEEE Cat. No88CH2614-6), Rome, Italy, 14-17 Nov, 1988. IEEE Computer Society Press, vol 1, pp 137-142.

Vliet, L.J. van (1991) Gray-Scale Measurements in Multidimensional Digitized Images. Ph.D. thesis. Delft University Press. Delft, The Netherlands.

Wilson T. (1990) Confocal Microscopy, T. Wilson ed., Academic Press, London 1990.

Yeo T.T.E., Jayasooriah, Ong S.H., Sinniah R. (1993) Autofocusing for tissue microscopy. *Image and Vision Computing* **11**, 629-639.

APPENDIX. Description of the Linked List Structure for databasing original images with segmentation results.

The linked list provides an easy way to store basic dependence relationships in a tree-linked structure that can be accessed and modified in a much more efficient way than other list types. It takes advantage of the dynamic addressing capabilities of C language to create series of 'objects' mutually linked by pointers. The relationships that can be created using these pointers are *child_of/parent_of/brother_of.* All of these relationships are bi-directional. In our applications we use it to store the relations between objects obtained at different stages of the analysis and the original 3D image.

The basic element of the list is the type *T_Object*, which is a C structure whose fields contain the information in the object (different types of objects can store different types of information) and pointers to other objects. Each object has pointers to a *Parent*, a list of *Children*, a *Previous* and *Next* object. As an example, nuclei obtained from a segmented DNA stained image are *Children* of an object of type *Image*, and themselves may have *Children*, which could for example be the results of the segmentation of FISH-labeled DNA sequences inside the nuclei. Besides the pointers, objects can be addressed using a unique index, calculated in a recursive way through the linked list, starting from the first object in the list (index=1), thus making the implementation of the list hardware independent. At this moment we have defined 9 types of objects: *Empty* (transitory object); *Image:* Original 3D color image composed of the different fluorecent channels acquired with the microscope in a single (lateral) field of view; *Unclassified:* Result of the segmentation process, before classification; *Nucleus*; *Nucleus on the edge:* Nucleus touching one of the borders of the image; *Cluster:* Unresolved group of touching nuclei; *Cluster on the edge*; *Debris*; *FISH Signal (Channel 1)*; *and FISH signal (Channel 2).*

Besides the indexes and links to other objects, every objects contains information about the coordinates (relative to the *Parent* object) of a prismatic box which includes the object's domain, and pointers to the original gray values and a binary mask of the object. It also includes the original voxel size and the interpolation factor used, if any, to correct the anisotropic image acquisition on that particular object. Finally the objects contain information, particular to each type of object.

Functions were written for the management of the linked list, such as creating new empty objects, deleting existing ones, creating copies of objects, moving through the list, linking a new object to an existing list in a particular place, replacing one object in the tree by a new object and saving/loading the tree to/from a file. We also wrote some interface functions to Scil-Image for converting objects to the SciIImage basic Image type and for displaying objects. All the functions were organized into libraries and compiled and linked under UNIX (Solaris v 2.6, IRIX v. 5.3).

Using this linked list structure, information from different sources, such as labeled genes, antibodies bound to expressed proteins, etc.- can be included into the tree without disturbing the basic data structure.

FIGURES

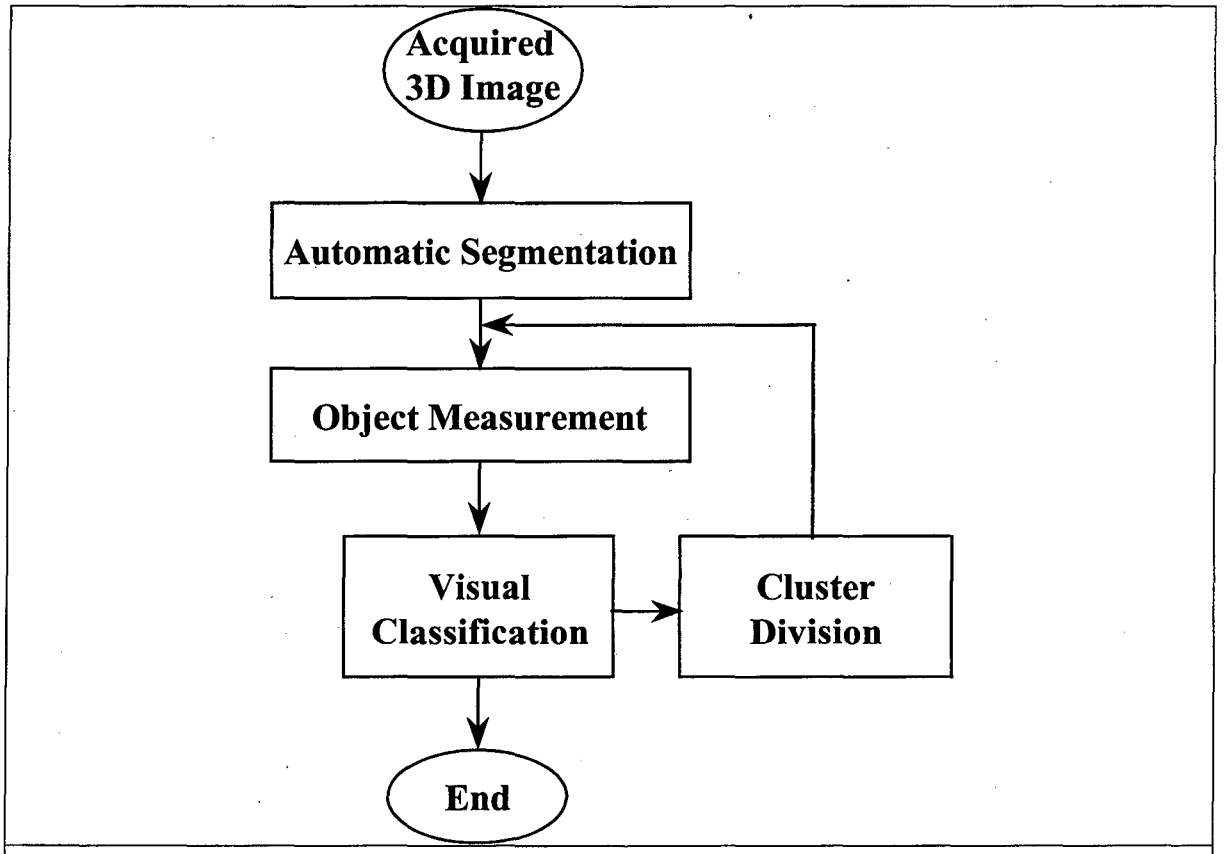


Figure 1

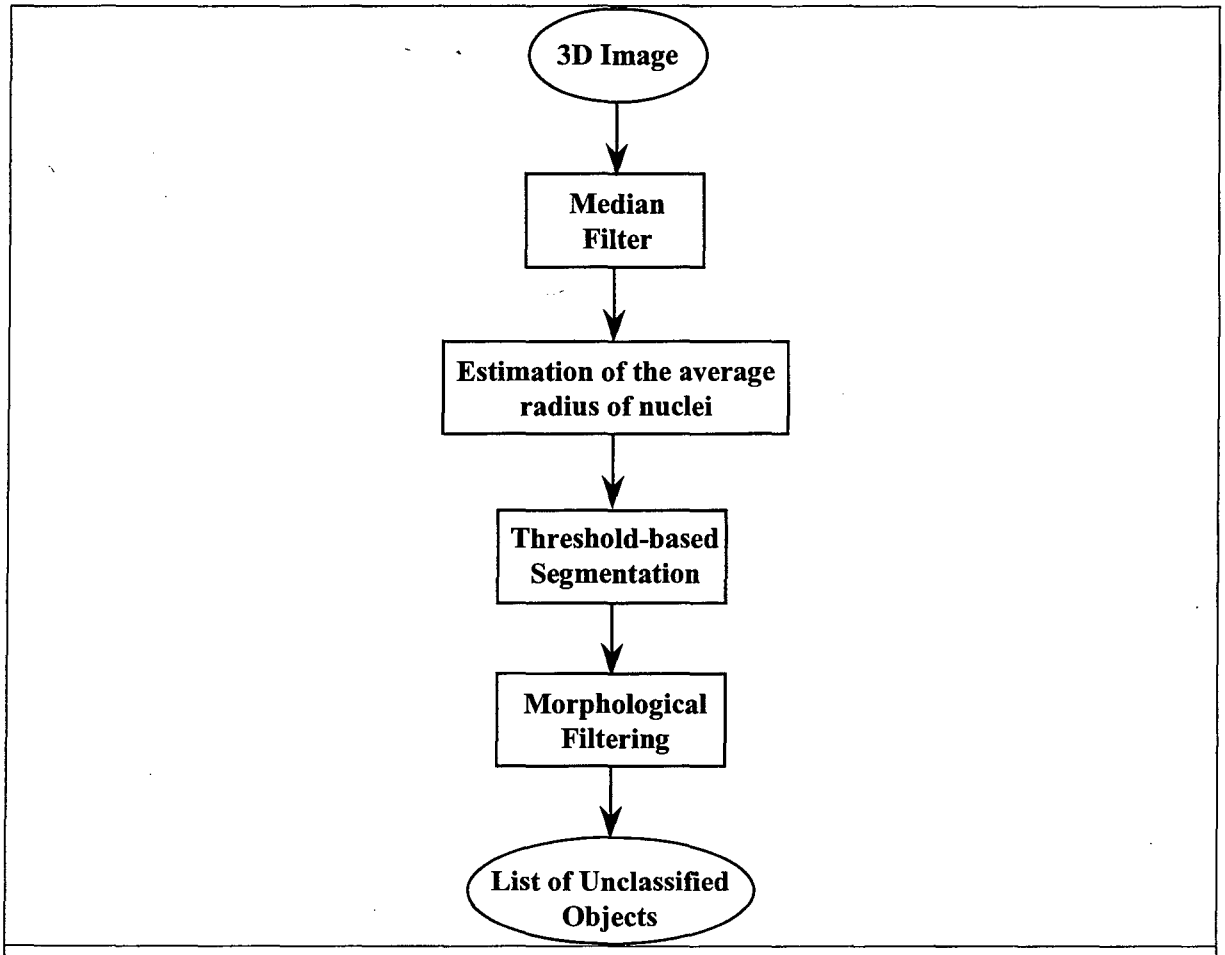

3D Image
Median Filter
Estimation of the average radius of nuclei
Threshold-based Segmentation
Morphological Filtering
List of Unclassified Objects

Figure 2

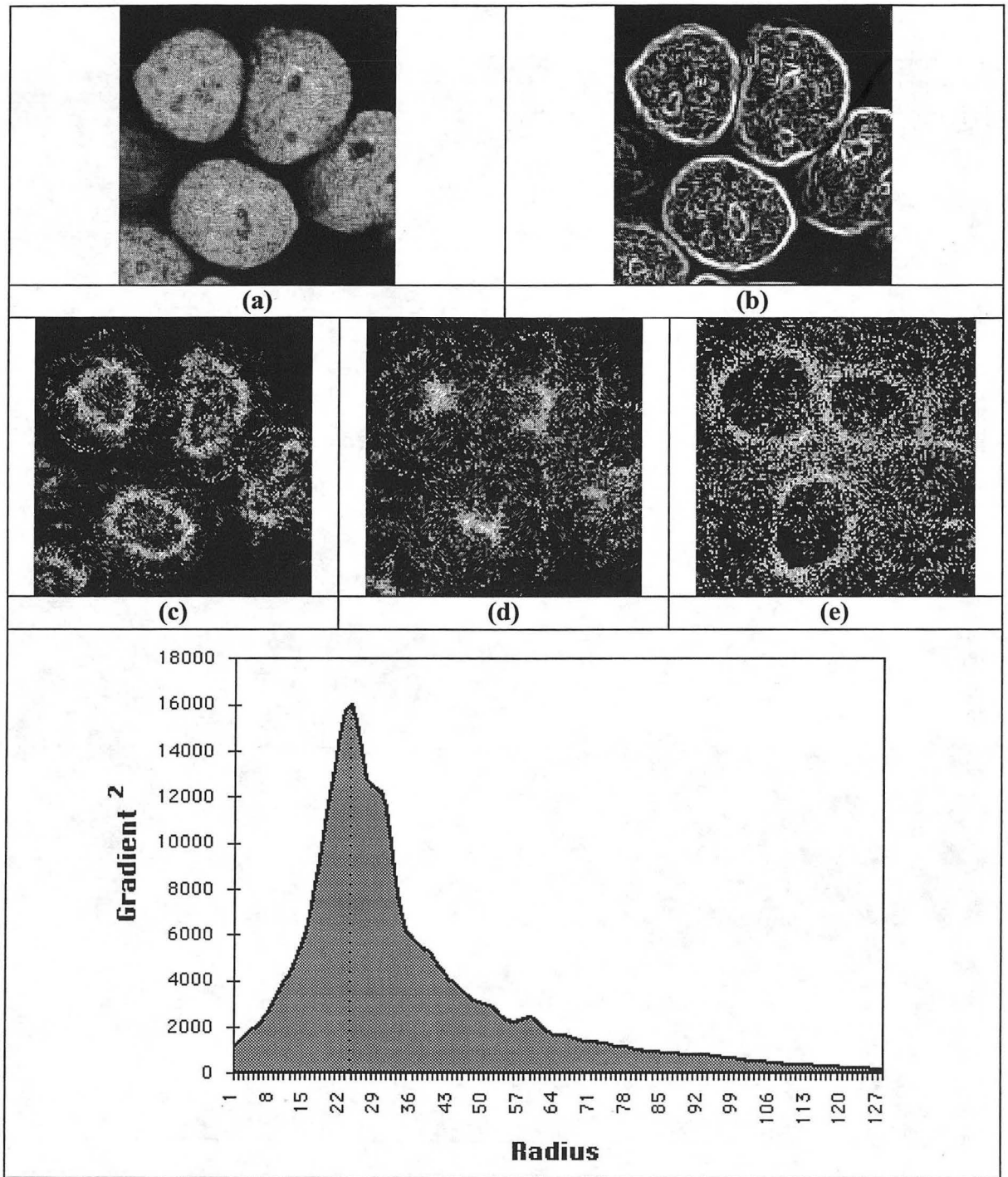

(a)
(b)
(c)
(d)
(e)
18000
16000
14000
12000
10000
8000
6000
4000
2000
0
Gradient 2
1
8
15
22
29
36
43
50
57
64
71
78
85
92
99
106
113
120
127
Radius
(f)

Figure 3

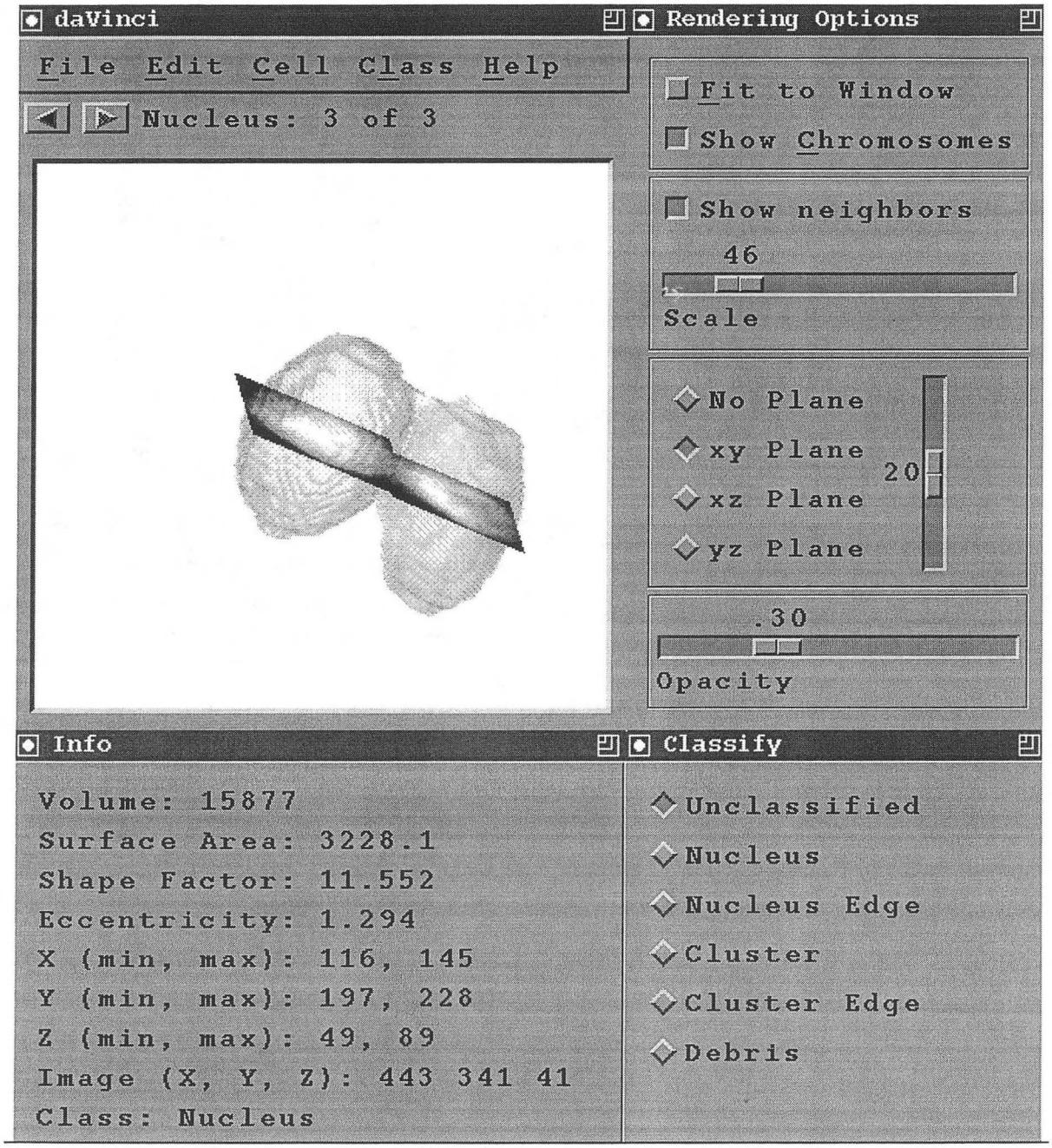

daVinci
File Edit Cell Class Help
Nucleus: 3 of 3
Rendering Options
Fit to Window
Show Chromosomes
Show neighbors
46
Scale
No Plane
xy Plane
xz Plane
yz Plane
20
.30
Opacity
Info
Volume: 15877
Surface Area: 3228.1
Shape Factor: 11.552
Eccentricity: 1.294
X (min, max): 116, 145
Y (min, max): 197, 228
Z (min, max): 49, 89
Image (X, Y, Z): 443 341 41
Class: Nucleus
Classify
Unclassified
Nucleus
Nucleus Edge
Cluster
Cluster Edge
Debris

Figure 4

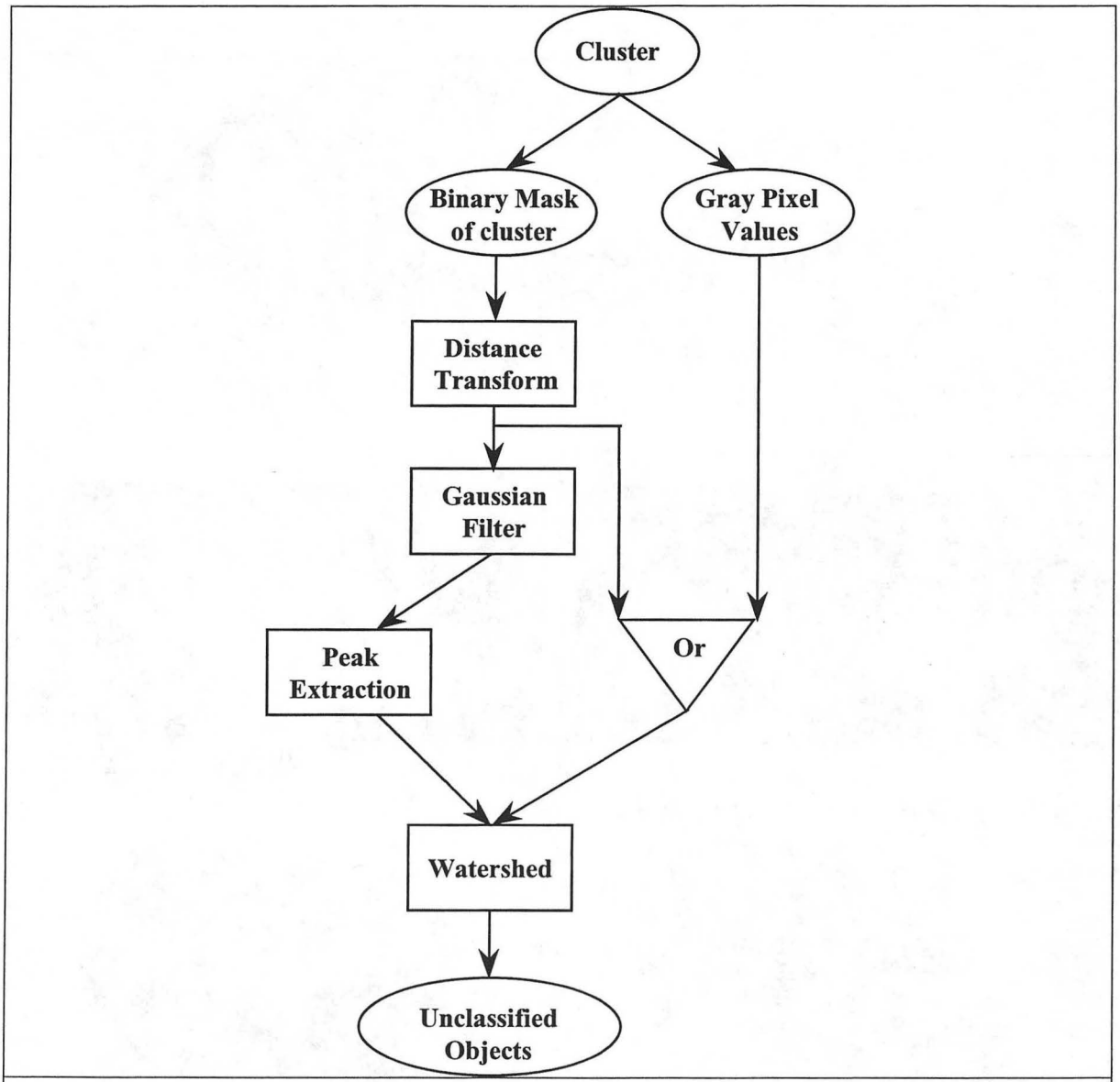

Cluster
Binary Mask of cluster
Gray Pixel Values
Distance Transform
Gaussian Filter
Peak Extraction
Or
Watershed
Unclassified Objects

Figure 5

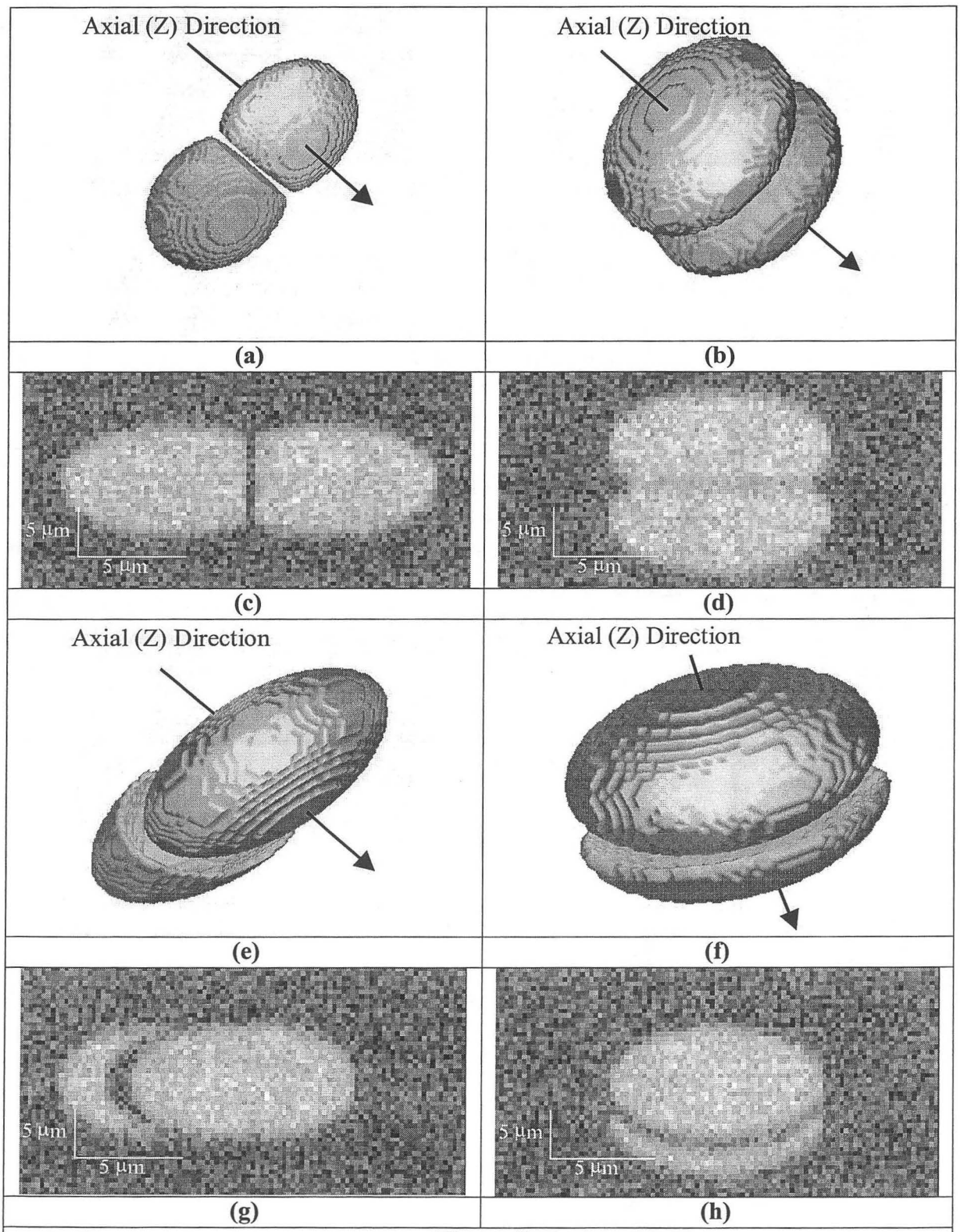
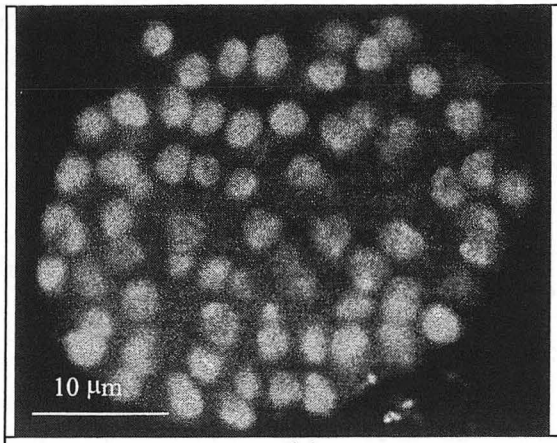
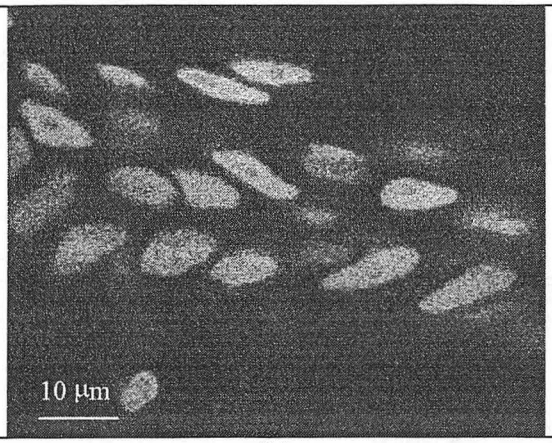

Axial (Z) Direction
Axial (Z) Direction
(a)
(b)
5 µm
5 µm
5 µm
5 µm
(c)
(d)
Axial (Z) Direction
Axial (Z) Direction
(e)
(f)
5 µm
5 µm
5 µm
5 µm
(g)
(h)

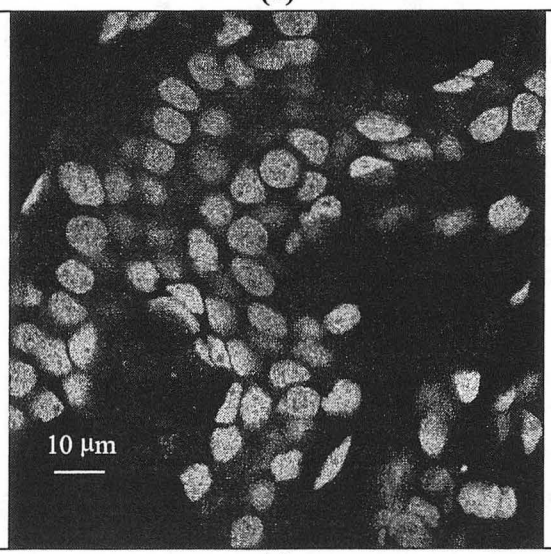
Figure 6



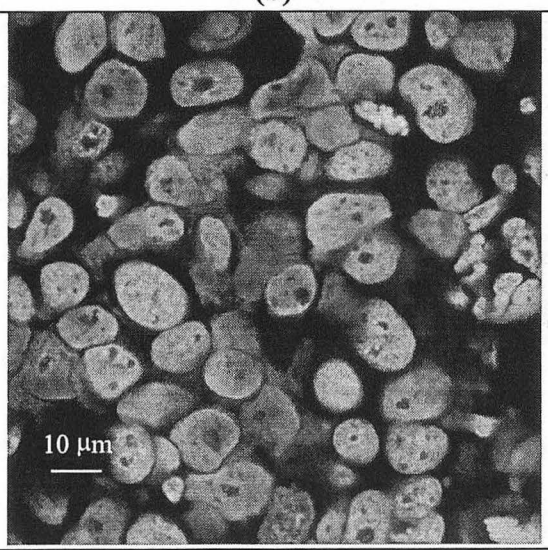
(a)



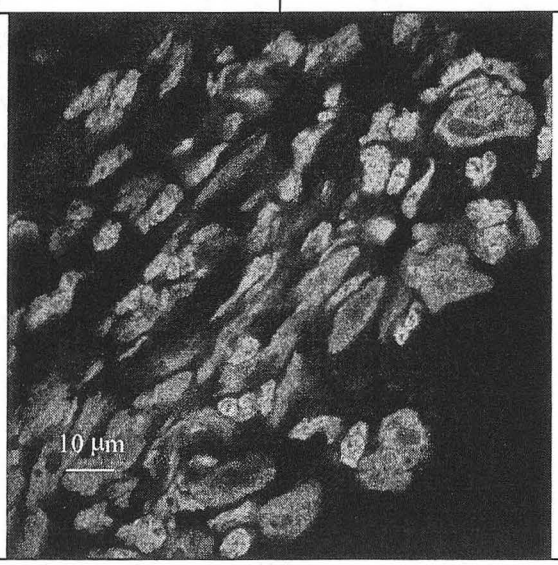
(b)



(c)

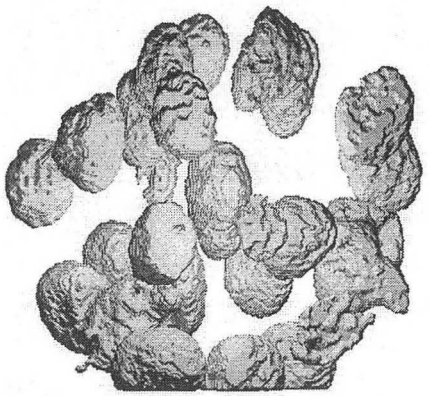


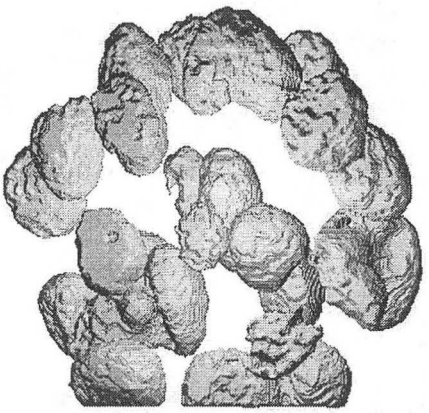
(d)



(e)

Figure 7





10 μm

Figure 8

TABLES

Table 1
Evaluation of the segmentation procedure on isolated computer generated objects. Surface accuracy

Object	Average distance in μm between true and segmented surfaces
Sphere (Radius=10)	0.18
Sphere (Radius=25)	0.15
Curved Disk (Radius=10)	0.10
Curved Disk (Radius=25)	0.17
Curved Ellipsoid	0.18
Test Ellipsoid	0.13

Table 2
Evaluation of the segmentation procedure using two clusters of two objects each. Minimum distance between objects (in μm) necessary to achieve segmentation for different values of Signal to Noise Ratio (SNR)

	SNR		
Cluster	**∞**	**3.5**	**2**
Two Truncated Spheres (oriented as in Fig. 6a)	0	0	0
Two Truncated Spheres (oriented as Fig. 6b)	0	0	0
Sphere+Disk (oriented as Fig. 6e)	0.4	0.6	0.8
Sphere+Disk (oriented as Fig. 6f)	0.8	1.4	1.6

Table 3

Evaluation of the segmentation accuracy for the touching parts of the surfacesof the objects in the clusters and for SNR=3.0

Cluster	Average distance in μm between true and segmented surfaces
Two Truncated Spheres (oriented as in Fig. 6a) *Distance between objects: 0 μm*	First Half: 0.15 Second Half: 0.14
Two Truncated Spheres (oriented as in Fig. 6b) *Distance between objects: 0 μm*	First Half: 0.11 Second Half: 0.17
Sphere+Disk (oriented as in Fig. 6e) *Distance between objects: 0.6μm*	Sphere: 0.24 Disk: 0.16
Sphere+Disk (oriented as in Fig. 6f) *Distance between objects: 1.4 μm*	Sphere: 0.23 Disk: 0.32

Table 4

Performance of the Segmentation Procedure on Tissue Specimens

Specimen	Correct Nuclei	Nuclei remaining in Clusters	Undefined	Lost
C.elegans	839 (98.9%)	2 (0.02 %)	0 (0 %)	7 (0.08 %)
Human Skin	343 (94.7 %)	16 (4.4 %)	1 (0.3 %)	2 (0.6 %)
Benign Breast Tumor	492 (93.7 %)	29 (5.5 %)	3 (0.6 %)	1 (0.2 %)
Xenografts	425 (88.7 %)	49 (10.2 %)	1 (0.2 %)	4 (0.9 %)
Invasive Carcinoma	260 (65.8 %)	75 (19.0 %)	56 (14.2 %)	4 (1.0 %)

FIGURE CAPTIONS

Figure 1. Outline of the Segmentation Procedure. First the 3D confocal image is automatically segmented into objects corresponding to contiguous regions of fluorescent DNA staining; The size and shape of these objects are measured and then the objects are classified by the user as being individual nuclei, debris or clusters of nuclei. Objects classified as clusters are divided into smaller objects that are returned for re-classification.

Figure 2. Flow chart of the main steps in the Automatic Segmentation to separate images into nuclear and non-nuclear regions: 1) Median filtering to remove shot noise; 2) Estimation of the average radius of nuclei (used in subsequent steps); 3) Thresholding to separate the image into nuclear regions (objects) and background; 4) Morphological filtering to remove small objects (debris) and separate slightly touching nuclei.

Figure 3. Example of the procedure to determine the average radius of the nuclei: (a) part of an *xy* slice from a 3D image of breast cancer cell line grown in mice. (b) Image of the gradient (magnitude) calculated from image (a). (c),(d) & (e) Examples of transformed images for R= 12, 24, 48. R=24 corresponds to the maximal shrinkage for the image. (f) Graph of the degree of focus as a function of R showing a peak at R=24.

Figure 4. The user interface of DAta VIsualization aNd Computer Interaction (DaVinci). **Top left window**: Menus and 3D visualization panel: The panel shows a surface rendered object which is two touching nuclei. It can be arbitrarily rotated by mouse interaction. The object is intersected with the 20th *xy* slice from the acquired image. The pull-down menus contain the following options:

a) *"File"* (General Options): 1) "Open" a list of stored objects (see the appendix for information about the file format); 2) "Save" the list of objects in a file; 3)

"Export_Cell_Data" creates a text file which contains the information about all the objects in the tree structure; 4) "Exit" terminates the execution of the program.

b) *"Edit"*: 1) "Undo" last action; 2) "Redo" last action; 3) "Delete" selected object(s); 4) "Join" selected objects; 5) "Select" centers the graphic panel on the selected object.

c) *"Cell"* opens the following windows: "Rendering Options", "Classify" and "Info" (see bellow).

d) *"Classify"* allows selective visualization of objects that belong to a given class: Unclassified, Nucleus, Nucleus on the Edge, Cluster, Cluster on the Edge or Debris.

e) *"Help"* provides instructions about the operation of the program.

The **top right window** ("Rendering Options") provides controls for rendering the objects: "Fit to Window" zooms the object to the size of the display panel; "Show Chromosomes" is not relevant to the study reported here; "Show neighbors" of the current objects; "Scale" manually changes the size of the object (zoom). The option bellow enables the rendered object to be intersected with any *xy*, *xz* or *yz* slice from the acquired gray image. "Opacity" changes the rendered surface.

The **bottom left window** ("Info") shows measurements of the current object: Volume, Surface Area, Shape Factor, Eccentricity, range of image Coordinates containing the object, Size of the image and object Class, if known.

The **bottom right window** allows the user to classify the current object as "Unclassified", "Nucleus", "Nucleus Edge", "Cluster", "Cluster on the Edge" or "Debris".

Figure 5. Cluster Division: 1) Calculation of the distance transform (DT) of the binary mask of the cluster; 2) Gaussian filtering to remove noise peaks from the DT image; 3) Extraction of marker peaks that are the center of nuclei; 4) Watershed algorithm to find surfaces between nuclei using the inverse of either the acquired image after median filtering or of the DT image.

Figure 6. Examples of the computer generated clusters of two objects that were used to assess the segmentation procedure. The rendered objects are not-interpolated and thus show the anisotropic voxel size. The lateral (*xy*) slices beneath the surface renderings show the central slice from the 3D images, and include the Gaussian blurring & noise (SNR=3.5). a) The two truncated spheres, two voxels (0.4 μm) apart, with contact surface parallel to the *yz* plane; b) two truncated spheres, two voxels (0.4 μm) apart, with contact surface parallel to the *xy* plane; c) central *xz* slice through (a); d) central *xz* slice through (b); e) curved disk and sphere, six voxels (1.2 μm) apart, with contact surface 'parallel' to the *yz*; f) curved disk and sphere, six voxels (1.2 μm) apart, with contact surface 'parallel' to the *xy*; g) central *xz* slice through (e); h) central *xz* slice through (f).

Figure 7. Lateral (*xy*) slices from 3D images of DNA stained specimens. a) *C. elegans* embryo, b) normal skin, c) benign breast tumor, d) breast cancer cells grown in mice, e) invasive carcinoma.

Figure 8. Stereo pair image of segmented normal skin nuclei.

TABLE CAPTIONS

Table 1. Evaluation of the segmentation procedure on isolated computer generated objects. The table shows for each object the average distance in μm between the true surface and the surface obtained by segmentation.

Table 2. Evaluation of the segmentation procedure using two clusters of two objects. Shown is the minimum distance (in μm) necessary for the algorithm to correctly divide the cluster into its objects, for three different noise levels.

Table 3. Evaluation of the segmentation accuracy at the surfaces between the computer generated objects of clusters. Shown is the average distance (in μm) between the segmented and true surfaces of the objects.

Table 4. Performance of the segmentation procedure for the tissue specimens. It shows, for each specimen type, the number of nuclei that were segmented correctly based on visual judgement (Correct nuclei); number of nuclei in clusters that could not be divided (Nuclei in Clusters); number of objects that the user could not identify (Undefined) and number of lost nuclei (Lost), due to poor DNA staining.

Ernest Orlando Lawrence Berkeley National Laboratory
One Cyclotron Road | Berkeley, California 94720

Prepared for the U.S. Department of Energy under Contract No. DE-AC03-76SF00098