

# 3D Segmentation of Mammospheres for Localization Studies <sup>\*</sup>

Ju Han<sup>1</sup>, Hang Chang<sup>1,2</sup>, Qing Yang<sup>2</sup>, Mary Helen Barcellos-Hoff<sup>1</sup> and Bahram Parvin<sup>1</sup>

<sup>1</sup> Lawrence Berkeley National Laboratory, Berkeley, CA 94720

<sup>2</sup> Institute of Automation, Chinese Academy of Sciences, Beijing, China

**Abstract.** Three dimensional cell culture assays have emerged as the basis of an improved model system for evaluating therapeutic agents, molecular probes, and exogenous stimuli. However, there is a gap in robust computational techniques for segmentation of image data that are collected through confocal or deconvolution microscopy. The main issue is the volume of data, overlapping subcellular compartments, and variation in scale and size of subcompartments of interest. A *geometric* technique has been developed to bound the solution of the problem by first localizing centers of mass for each cell and then partitioning clump of cells along minimal intersecting surfaces. An approximate solution to the center of mass is realized through iterative spatial voting, which is tolerant to variation in shape morphologies and overlapping compartments and is shown to have an excellent noise immunity. These centers of mass are then used to partition a clump of cells along minimal intersecting surfaces that are estimated by Radon transform. Examples on real data and performance of the system over a large population of data are evaluated. Although proposed strategies have been developed and tested on data collected through fluorescence microscopy, they are applicable to other problems in low level vision and medical imaging.

## 1 Introduction

Current models of high throughput and high content screening are based on two dimensional cell culture assays that are grown either on plastic or glass. Although such a model system may be appropriate as an initial step toward discovery or certain aspect of biological studies, the knowledge may not readily extensible to *in vivo* models. On the other hand, animal studies are expensive and time consuming and as a result cannot scale for high throughput studies that is necessary to build a space-time continuum of responses in the presence of biological heterogeneity. An intermediate step is three-dimensional cell culture

---

<sup>\*</sup> The Research was supported by National Aeronautics and Space Administration Grant no. T6275W, NASA Specialized Center for Research in Radiation Health Effects, the low dose radiation research program and medical imaging program at the Office of Biological Effects Research U.S. Department of Energy, Grant No. DE-FG03-01ER63240. PubID is LBNL-61402.

model systems that has been demonstrated to have some of functionalities of the *in vivo* models [11]. However, such a model system introduces significant computational challenges: (i) imaging is in 3D and not in projection space, (ii) subcellular compartments often overlap and delineation is made difficult, and (iii) variations in subcellular scale imposes a more complex segmentation problem at the object level. In this paper, we present a series of unique geometric steps for segmentation of 3D cell culture models, also known as acini, that enables subsequent localization studies and protein-protein interactions.

Research in the analysis of subcellular structures spans texture-based features for classifying patterns of protein expression [5], and geometric methods [8,9,12] and surface evolution methods [4,6] for delineation of nuclear compartments. Segmentation provides context for quantifying protein localization in fixed samples with an antibody or a nucleic acid based probe in living cell studies [8]. To our knowledge, previous techniques are not applicable for automated segmentation mammospheres. Proposed solution is to bound an inherently ill-posed problem through geometric constraints. A key observation is that nuclear regions are often convex and form a positive curvature maxima when they overlap each other. This feature was used earlier in 2D segmentation of nuclear regions [9]. However, evaluating 3D convexity and estimating 3D surface curvature is hindered with a significant computational complexities. Convexity can be viewed as a salient feature, which is an important perceptual cue for localization and segmentation of subcellular regions. In biological image understanding, saliency can be driven by continuity, symmetry, convexity, or closure. Among these, it is well known that symmetry is a pre-attentive process [1] that improves recognition, provides an efficient mechanism for scene representation, and aids in reconstruction and description. Radial (spherical in 3D) symmetry is a special class of symmetry, which persists in nature at multiple scales. Robust and efficient detection of inexact radial symmetries facilitates the semantic representation, interpretation, and prioritization for higher level processes. Yet, the notion of radial symmetry is used in a weak sense since the basic geometry can deviate in aspect ratio and convexity. Localization of approximate centroid of each nucleus in a three dimensional cell culture assay enables partitioning a mammosphere along the planes that generate minimum surface cross sections and are possibly aligned with points of maximum curvature along the surface.

The novelty of the proposed method is in specific geometric steps designed to bound the solution through seeding and subsequent partitioning. The basis for seeding (e.g., estimating centers of mass) is through geometric voting and perceptual grouping, and is implemented through the refinement of specifically tuned voting kernels [13]. Localization of centers of mass for each nucleus provides a bound on overall stable segmentation. Partitioning of adjacent nuclei is performed by finding planes that best separate adjacent cells, and the actual methodology is based on the Radon transform.

The organization of this paper is as follows. Section 2 provides a brief review of the previous research. Section 3 describes the basic idea and detailed

implementation of approach. Section 4 demonstrates the experimental results. Section 5 concludes the paper.

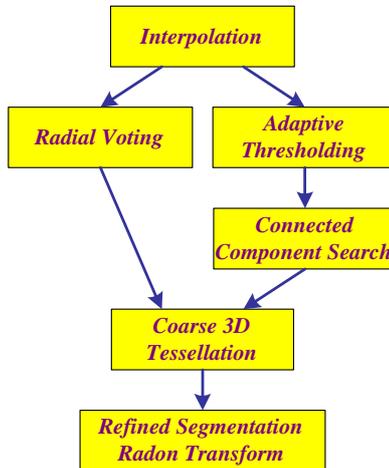
## 2 Previous research

Current state of art for delineation of nuclear regions from 3D multicellular systems and mammospheres leverage intensity information with limited amount of inherent geometry. Some of these methods are interactive and serve as a computed aided tool to increase operational throughput. In [7], background and nuclear regions are automatically delineated using a thresholding mechanism, Hough transform and automatic focusing are applied to estimate the size of the nuclei, the user labels each object as a single nucleus or a cluster of nuclei, and the process ends with watershed method to partition potential clumps of cells. In [10], limitations in [7] were identified in several computational steps: (1) initial thresholding, (2) noise, and (3) low gradient in some of the nuclear regions. These limitations were then addressed using level set methods for improved performance. In [12], nuclear regions were modeled as elliptic features and fragmented features were grouped together to form a convex hull. The method produces a segmentation that is not very accurate along the surfaces with potential fragmentation of nuclear regions. Recently, in [2] mammosphere slices were segmented in 2D and then segmented 2D slices were merged together. However, combination of assay and high resolution imaging produced a morphological nuclear signature that tend to be more separable in 2D (e.g., little overlap) while maintaining similar scale in nuclear size. One novelty of this system is that the analysis does not require isotropic representation of the data volume.

In contrast to previous approach, proposed method uses high level geometric features to delineate a multicellular system. Geometrically, nuclear regions are almost convex; however, scale (e.g., size) is heterogeneous. Furthermore, when two nuclear regions overlap, they form folds corresponding to positive curvature maxima. Partitioning adjacent nuclear regions along points of curvature maxima or a variation of that is the final step of the process.

## 3 Approach

Proposed steps in delineation of nuclei in a mammosphere system are shown in Figure 1. Starting from the interpolated 3D image, the solution first bounds the problem by computing seeds that estimate the centroid of each nucleus through iterative radial voting in 3D. Simultaneously, the colony is thresholded in 3D, which produces an erroneous segmentation of the clump of nearby cells by merging them. Each clump is subsequently labeled for further analysis, and any connected volume with more than two seeds is subject for further analysis. Partitioning is performed by finding planes that best separate adjacent nuclei, and the actual methodology is based on the Radon transform. However, Radon transfer precedes by a coarse segmentation from adjacent seed locations. Ideally, such a coarse segmentation should be realized through voronoi tessellation, which is

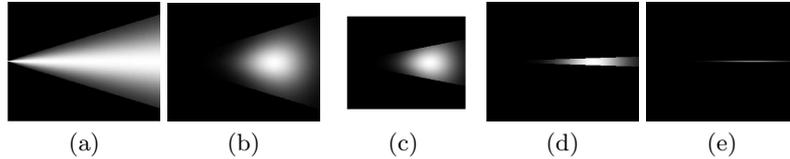


**Fig. 1.** Detailed representation of methodology for segmentation of a 3D mammosphere.

compute intensive in 3D. A simpler approximation to voronoi tessellation is implemented to provide a rough segmentation of nuclear regions. This segmentation is further refined by Radon transform. Details of seed selection through iterative voting and partitioning adjacent connected nuclei through radon transform are included below.

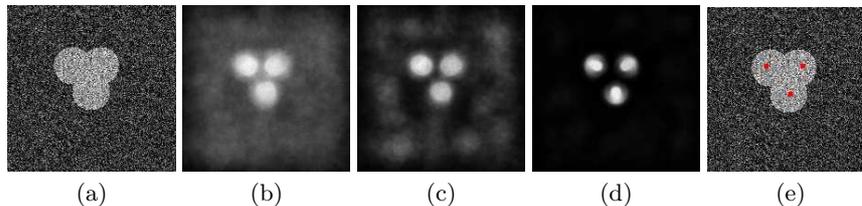
### 3.1 Seed estimation with iterative voting

The basis for seeding (e.g., estimating centers of mass) is through geometric voting and perceptual grouping, and is implemented through the refinement of specifically tuned voting kernels [13]. In general, voting operates on the notion of continuity and proximity, which can occur at multiple scales, e.g., points, lines, lines of symmetry, or generalized cylinders. The novelty of our approach is in defining a series of kernels that vote iteratively along the radial or tangential directions. Voting along the radial direction leads to localization of the center of mass, while voting along the tangential direction enforces continuity. At each iteration, the kernel orientation is refined until it converges to a single focal response. Voting kernels have a cone-shaped with an initial scale and spread (e.g., height and base) that is refined iteratively. These kernels are initially applied along the gradient direction, then at each consecutive iteration and at each grid location, orientation is aligned along the maximum local response. The method has excellent noise immunity, is tolerant to variations in target shape scale, and is applicable to a large class of application domains. Figure 2 shows a subset of voting kernels that vary in topography, scale, and orientation.



**Fig. 2.** Kernel topography: (a-e) Evolving kernel for the detection of radial symmetries (shown at a fixed orientation) has a trapezoidal active area with Gaussian distribution along both axes.

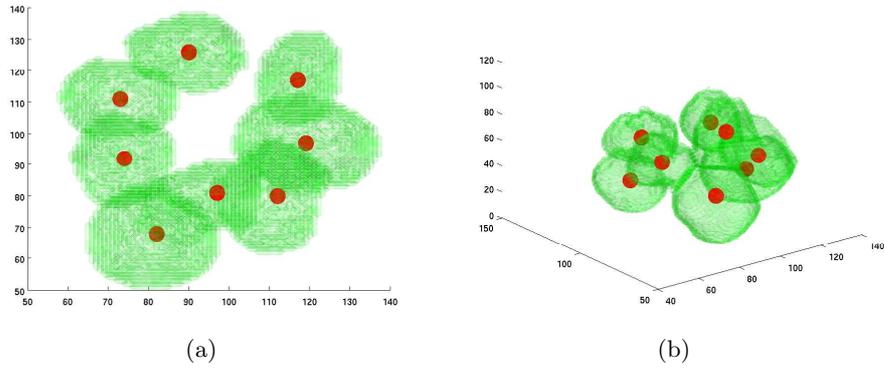
The iterative voting algorithm is presented for the 2D case in [9]. Our current implementation extends the 2D algorithm to volumetric data. An example of the application of radial kernels to overlapping 2D objects is shown in Figure 3 together with the intermediate results. The voting landscape corresponds to the spatial clustering that is initially diffuse and subsequently refined and focused into distinct islands. Figure 4 shows two views of the voting results for a 3D clump of mammosphere.



**Fig. 3.** Detection of radial symmetries for three overlapping blobs with with the signal-to-noise ratio of  $6dB$ : (a) original image; (b)-(d) the voting landscape at intermediate iterations; and (e) final localization following thresholding. The voting landscape is initially diffused in the background region, but it becomes more localized at the foreground in subsequent iterations.

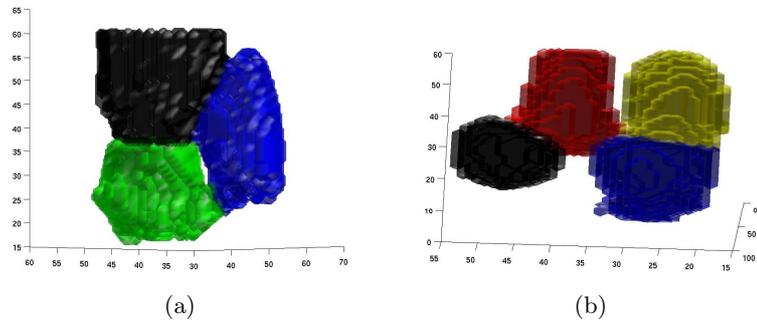
### 3.2 Partitioning of a mammosphere from seeded nuclei

The process is initiated by a coarse segmentation of nuclei with a simplified 3D voronoi tessellation. Tessellation facilitates (1) identification of a local neighborhood where each nuclear region is contained within its own space, and (2) improved computational performance for each mammosphere prior to Radon transform. The first aspect has to do with constrained locality, which eliminates error and reduces ambiguities. Without tessellation, Radon transfer will fail because two neighboring nuclei may have a third nucleus that sits at the fold of the two touching nuclei. This condition is shown and visualized in Figure 5 from



**Fig. 4.** Two views of voting results of a 3D clump of mammosphere: (a) top view; (b) side view.

real data. The second point has to do with the fact that not all adjacent nuclei are connected and that there is a clear empty space between them. Under this condition, there is no need to refine the segmentation further.



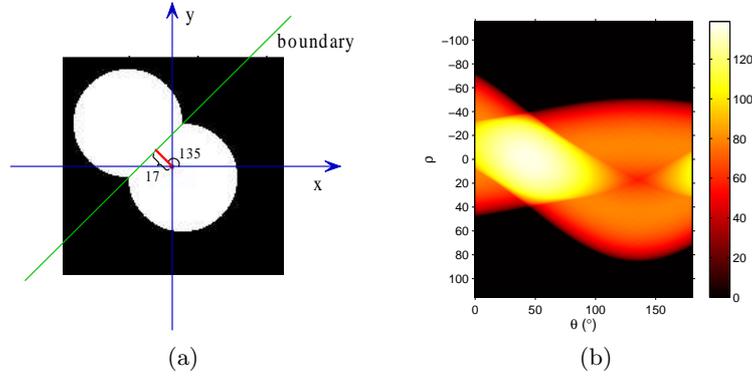
**Fig. 5.** A segmented example of the nuclear configuration where tessellation enforces locality: (a) the blue nuclei resides at the fold between the green and black nuclei and without an initial tessellation subsequent Radon transform refinement will fail; and (b) empty spaces between black and blue nuclei eliminates the need for Radon transform refinement.

The details of Radon transform is as follows; however, for simplicity the 2D version is first described. The Radon transform represents an image as a collection of projections in a function domain  $f(x, y)$  along various lines defined by the shortest distance  $\rho$  from the origin and the angle of inclination  $\theta$  with the

$y$  axis:

$$R(\rho, \theta) = \int \int f(x, y) \delta(\rho - x \cos \theta - y \sin \theta) dx dy.$$

Properties of the Radon transform enable delineation of nearby touching objects. For example, two adjacent objects, represented by circles in Figure 6(a), and the corresponding Radon transform shown in Figure 6(b), has a local minimum at  $\rho = 17$  and  $\theta = 135^\circ$ . This local minimum corresponds to the integration over the line that separates the two objects with the smallest cross section.

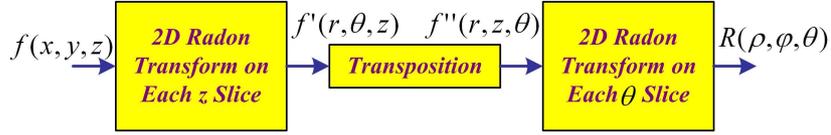


**Fig. 6.** An example of 2D object segmentation using the Radon transform: (a) synthetic object composed of two circles; and (b) corresponding Radon transform with local minimum at  $\rho = 17$  and  $\theta = 135^\circ$ .

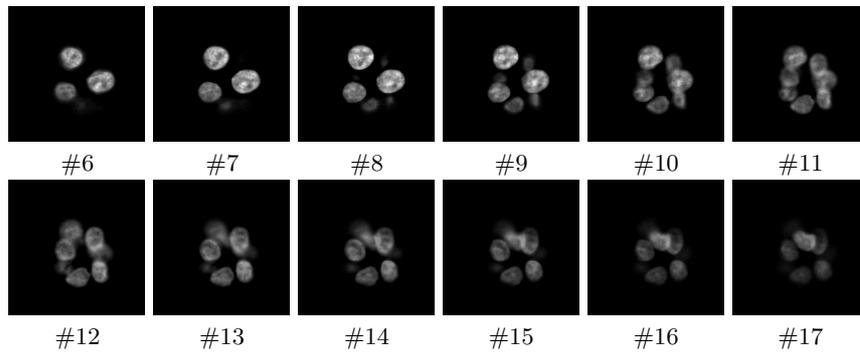
Similarly, 3D Radon transform represents a 3D volume as a collection of projections in a function domain  $f(x, y, z)$  along various planes defined by the shortest distance  $\rho$  from the origin, the angle of azimuth  $\phi$  around the  $z$  axis and the angle of elevation  $\theta$  around the  $y$  axis:

$$R(\rho, \phi, \theta) = \int \int \int f(x, y, z) \delta(\rho - x \cos \phi \cos \theta - y \sin \phi \cos \theta - z \sin \theta) dx dy dz.$$

The Radon transform is a separable transform and its implementation is shown in Figure 7. A fast method for computing 3D radon transform via a direct Fourier method can be found in [3]. Given a local cube containing two nearby adjacent cells, each of which is bounded by a seed, the optimal plane separating these two cells should be located between the two seeds and have the smallest cross section. The local minimum in the 3D Radon transform corresponds to the integration over the optimal plane in the local cube.



**Fig. 7.** The implementation of 3D Radon transform through 2D Radon transform.

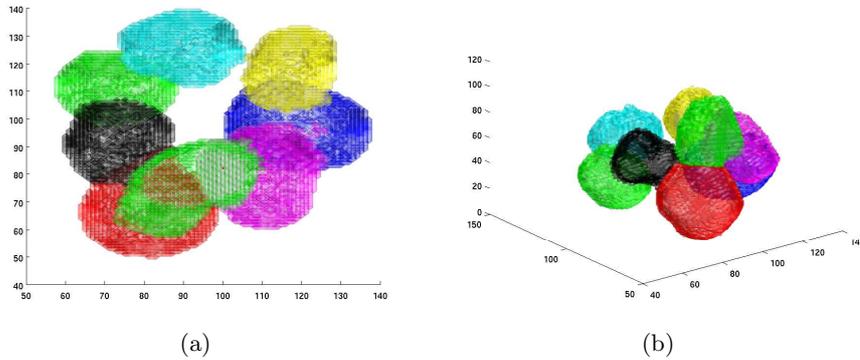


**Fig. 8.** Slices from a 3D cell colony in the order of  $z$  direction.

## 4 Experimental results

The proposed approach was implemented and applied in a data set of mammospheres, imaged with deconvolution microscopy. The image resolution along the  $x$  and  $y$  directions is  $0.15 \text{ micron}$ , and the resolution along the  $z$  direction is  $1.32 \text{ micron}$ . Figure 8 shows several slices from a mammosphere, and the corresponding segmentation result is shown in Figure 9. 151 colonies in the data set with an average of 11 seeds per colony were processed using the proposed approach and the same input parameters. 1771 seeds were estimated through iterative radial voting; however, 77 nuclei did not register any corresponding seeds, which indicates a detection error rate of 4%. This is presumably due to abnormal scale and shape of the nuclear volume and the exact conditions are as follows:

- *Low contrast between overlapping nuclei:* Absence of gradient information between overlapping nuclei coupled with their accidental morphological properties provide ambiguous voting evidence that produces one fixed point instead of two.
- *Morphological abnormality:* Often a single nucleus has an abnormal elongated shape and radial voting merges multiple seed points into a single fixed point. This condition is highly correlated with previous case.



**Fig. 9.** Two views of final segmentation of a mammosphere for the stack shown in Figure 8: (a) top view; (b) side view.

- *Incomplete information:* This is an imaging problem where imaging is incomplete and part of the nuclei is missing from the volumetric image, as shown in Figure 5.
- *Low sampling resolution in Z axis:* The current interpolation algorithm is linear for making an volumetric stack homogeneous in its X, Y, and Z dimension. Linear interpolation smooth the gradient in the Z direction and reduces contribution of the corresponding gradient information. An improved model will use some form of spline interpolation.

Finally, partitioning accuracy was compromised for 35 pair of overlapping nuclei from a total of 1850 pairs, which indicates an error rate of approximately 2%. These errors occur when the optimum planes for separating two nuclei is not the desired plane for partitioning two neighboring nuclei. The notion of desired planes has to do with those planes that bisect neighboring nuclei along points of maximum curvature. In this case, the error rate can be reduced through improved seed localization.

## 5 Conclusion and future work

This paper presented a series of geometric steps for segmentation of mammospheres from volumetric data. The first step localizes an approximation to center of mass for each nucleus and then partitioning clump of nuclei along minimal intersecting surfaces. Approximate solution to the center of mass is realized through iterative spatial voting, which is tolerant to variation in shape morphologies, perceptual surfaces, noise, and overlapping compartments. These centers of mass are then used to partition a clump of cells along minimal intersecting surfaces that are estimated by Radon transform. The technique has been tested on 151 colonies and their corresponding 3D volumes, and error rate is fully char-

acterized. The system is being planned to be used for subsequent localization studies.

## References

1. F. Attneave. Symmetry information and memory for patterns. *American Journal of Psychology*, 68:209–222, 1955.
2. D. Knowles, D. Sudar, C. Bator-Kelly, M. Bissell, and S. Lelievre. Automated local bright feature image analysis of nuclear protein distribution identifies changes in tissue phenotype. *Proceedings of National Academy of Science*, 103:4445–4450, 2006.
3. S. Lanzavecchia and P. Luigi Bellon. Fast computation of 3d radon transform via a direct fourier method. *Bioinformatics*, 12(2):212–216, 1998.
4. R. Malladi, J. Sethian, and B. Vemuri. Shape modeling with front propagation: a level set approach. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 17:158–175, 1995.
5. R. Murphy. Automated interpretation of subcellular locatoin patterns. In *IEEE Int. Symp. on Biomedical Imaging*, pages 53–56, April 2004.
6. C. Ortiz De Solorzano and et. al. Segmentation of nuclei and cells using membrane protein. *Journal of Microscopy*, 201:404–415, March 2001.
7. C. Ortiz De Solorzano, R. Garcia, A. Jones, D. Pinkel, W. Gray, D. Sudar, and S. Lockett. Segmentation of confocal microscope images of cell nuclei in thick tissue section. *Journal of Microscopy*, 193(3):193–212, 1999.
8. B. Parvin, Q. Yang, G. Fontenay, and M. Barcellos-Hoff. Biosig: An imaging bioinformatics system for phenotypic studies. *IEEE Transactions on Systems, Man and Cybernetics, Part B*, 33:814–824, 2003.
9. S. Raman, B. Parvin, C. Maxwell, and M. Barcellos-Hoff. Geometric approach segmentation and protein localization in cell cultured assays. In *Int. Symposium on Visual Computing*, pages 427–436, 2005.
10. A. Sarti, C. Ortiz De Solorzano, S. Lockett, and R. Malladi. A geometric model for 3-d confocal image analysis. *IEEE Transactions on Biomedical Engineering*, 47(12):1600–1610, December 2000.
11. V.M. Weaver, A.H. Fischer, O.W. Petersen, and M.J. Bissel. The importance of the microenvironment in breast cancer progression: recapitulation of mammary tumorigenesis using a unique human mammary epithelial cell model and a three-dimensional culture assay. *Biochemical Cell Biology*, 74(12):833–51, 1996.
12. Q. Yang and B. Parvin. Harmonic cut and regularized centroid transform for localization of subcelular structures. *IEEE Transaction on Biomedical Engineering*, 50(4):469–476, 2003.
13. Q. Yang and B. Parvin. Perceptual organization of radial symmetries. In *Proceedings of the Conference on Computer Vision and Pattern Recognition*, pages 320–325, 2004.