Detecting and Segmenting Nanodiscs in Immuno-Electron Micrographs

Tianjun Huang[1]
t.huang@dundee.ac.uk
Christian Hacker[2]
ch84@st-andrews.ac.uk
John Lucocq[2]
jml7@st-andrews.ac.uk
Stephen McKenna[1]
stephen@computing.dundee.ac.uk

[1] School of Computing
University of Dundee
Dundee, UK

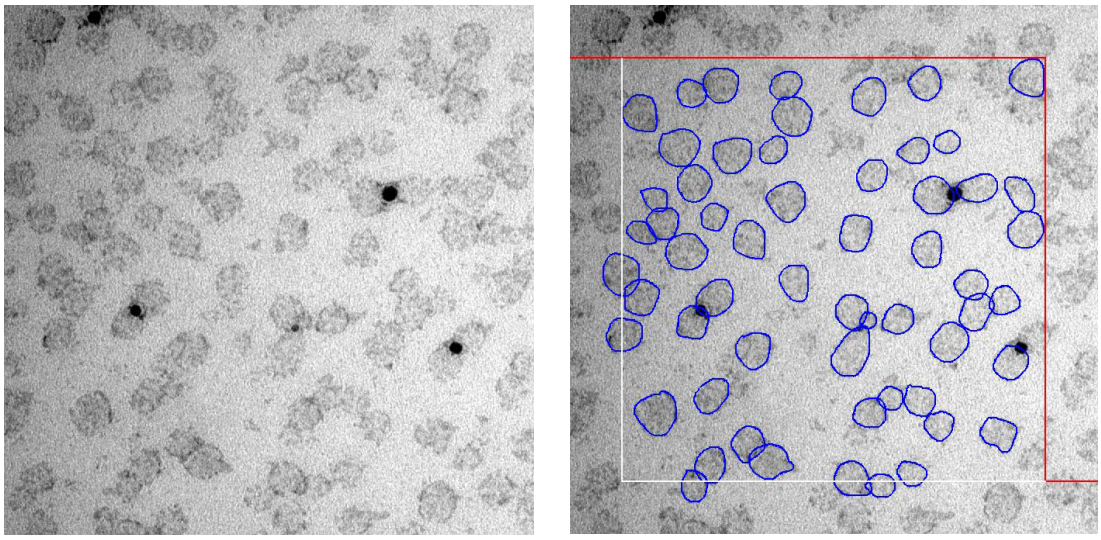
[2] School of Medicine
University of St Andrews
St Andrews, UK

Abstract

Nanodiscs are soluble nanoscale phospholipid bilayers with applications in drug delivery and the study of membrane proteins, for example. They can be imaged using electron microscopy, along with immunogold markers indicating locations of proteins of interest. We describe and evaluate methods for automatically detecting and segmenting nanodiscs in electron micrographs. The detection method modifies aspects of the Fast Radial Symmetry Transform to detect nanodiscs that exhibit approximate radial symmetry against noisy but predominantly lighter background. Detected nanodiscs are then segmented using radial active contours. Experiments on micrographs both with and without immunogold markers indicate promising detection and segmentation performance, and that information on nanodisc quantities, locations, size distributions, and co-location with proteins of interest could be extracted automatically.

1 Introduction

Nanodiscs are soluble nanoscale phospholipid bilayers composed of a genetically engineered membrane scaffold protein and phospholipid [2]. They are useful for studying the function and structure of membrane proteins and have applications such as vehicles for transport of hydrophobic drugs [7]. Electron microscopy (EM) can be used to image nanodiscs, as well as to locate proteins via immunogold markers that appear as dark spots in the electron micrographs. This paper investigates automating the detection and segmentation of nanodiscs, in the presence of immunogold markers, to enable information such as nanodisc quantities, locations, size distributions, and co-location with immunogold-marked proteins to be measured efficiently during studies of protein structure and function or of applications such as drug delivery. We are not aware of any previously published literature describing how to automatically detect and segment nanodiscs using EM.



(a) (b)

Figure 1: (a) Nanodisc image with gold markers. (b) Manual annotation using forbidden lines.

2 Detecting Nanodiscs

Figure 1 shows an example EM image containing nanodiscs as well as five immunogold markers. When imaged in this way, nanodiscs tend to be characterised by approximate radial symmetry, noisy interiors which are on average darker than the background, and darkening around their boundaries. We detect them using a method inspired by the Fast Radial Symmetry Transform (FRST) that was proposed by Loy and Zelinsky to detect interest points exhibiting local radial symmetry [5]. FRST has been used previously as a component in cell nuclei detection [8]. Ni *et al.* [6] extended FRST to cope with affine transformations and applied it to nuclei detection in histopathology. Here we modify FRST to detect nanodiscs.

The detector searches for nanodiscs at radii $n \in N$ where N is a set of radii sampling a known range of nanodisc radii (in pixels). Image gradients are computed (using Farid and Simoncellis' algorithm [3]) and pixels at which the gradient magnitude exceeds a threshold τ_{mag} vote to produce an *orientation projection image* O_n and a *magnitude projection image* M_n. Specifically, a pixel at location $\mathbf{p}$ votes for the pixel location nearest to $\mathbf{p} - n\frac{g(\mathbf{p})}{||g(\mathbf{p})||}$, where $g(\mathbf{p})$ is the gradient at $\mathbf{p}$. After voting is complete, the value of $O_n(\mathbf{p})$ is the number of pixels that voted for $\mathbf{p}$ and $M_n(\mathbf{p})$ is the sum of the gradient magnitudes of all the pixels that voted for $\mathbf{p}$. Similarly to [5] these are combined to give $G_n(\mathbf{p}) = M_n(\mathbf{p})(\min(k, O_n(\mathbf{p})))^2$ where k was set to approximately 20. However, rather than average the response at different radii as in [5], we find the maximal response $S(\mathbf{p}) = \max_n G_n(\mathbf{p})$ and treat this as evidence for a nanodisc of radius $n^* = \arg\max_n G_n(\mathbf{p})$ centred at $\mathbf{p}$. A nanodisc will tend to have high response, S, near its centre but this response will be spread and can even be multimodal. We apply a local averaging filter to S and then set response values not exceeding a detection threshold τ_{detect} to zero. Non-maximum suppression is then used to locate candidate nanodisc centres (with the radius of the suppression window set to the known minimum nanodisc radius).

Before running the nanodisc detector, any immunogold markers in the image are detected using the method described by Wang *et al.* [9]. Immunogold marker regions are then dilated

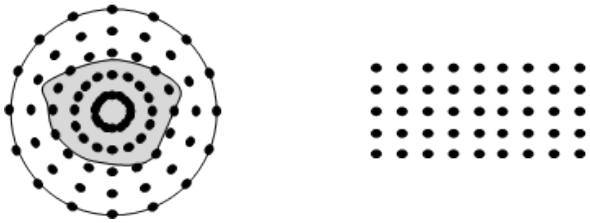


Figure 2: Left: Sampled annular search space. Right: Corresponding trellis.

and any pixels within these dilated regions are excluded from voting for nanodisc centres. This is important to avoid the presence of immunogold markers adversely affecting nanodisc detection. Impurities in the electron micrograph can result in very dark regions; these are ignored by simply removing any candidate nanodisc centre with an image intenisty below a conservative threshold ε. If the distance between the centres of two candidates is less than any one of the candidates' radii, the candidate which has higher symmetry response is kept. Finally, any remaining candidates are labelled as nanodiscs. In the experiments, we set τ_{mag} to 2% of the maximum gradient magnitude, $\tau_{detect} \approx 0.1$ and $\varepsilon = 20$.

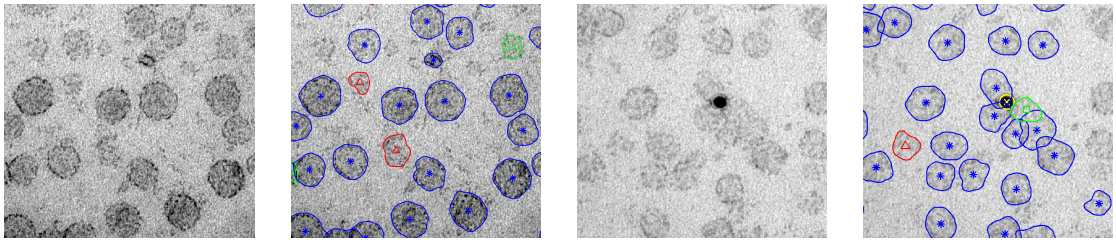
3 Segmenting Detected Nanodiscs

Each nanodisc detection is used to initialise a radial active contour segmentation similar to that used by Bamford and Lovell to segment cell nuclei [1]. The radius and the centre location estimated by the nanodisc detector are used to determine an annular search region within which the nanodisc's boundary contour is assumed to lie (see Figure 2). The annulus is sampled along M radial lines at regular angular spacing, at N equally spaced points on each radial line. The resulting search space can be thought of as an $N \times M$ trellis (right of Figure 2) and the search for the nanodisc contour posed as the search for an optimal left-right path across the trellis (where the left and right sides of the trellis are treated as adjacent). The Viterbi algorithm finds an optimal left-right path. Path optimality is with respect to a cost function that combines internal contour energy, E_{int}, and external energy, E_{ext}, as in Equation (1) where $\mathbf{v}_i$ represents a point in the i^{th} column of the trellis. The internal energy is given in Equation (2); it encourages smooth contours and is minimised by circles.

$$C_i(\mathbf{v}_{i+1}, \mathbf{v}_i) = \min[C_{i-1}(\mathbf{v}_i, \mathbf{v}_{i-1}) + \lambda E_{int}(\mathbf{v}_{i-1}, \mathbf{v}_i, \mathbf{v}_{i+1}) + (1-\lambda)E_{ext}(\mathbf{v}_i)] \tag{1}$$

$$E_{int} = \frac{\|\mathbf{v}_{i-1} - 2\mathbf{v}_i + \mathbf{v}_{i+1}\|^2}{\|\mathbf{v}_{i-1} - \mathbf{v}_{i+1}\|^2} \tag{2}$$

The parameter $\lambda \in [0,1]$ balances internal and external energies. We set $\lambda = 0.7$. At each stage, the minimum cost at the current point and also its corresponding path from the point of the previous stage will be recorded. The final path is obtained by back-tracking from the point in the last column of the trellis. In order to obtain a closed contour we adopt a technique proposed by Gunn [4]; two points at the middle of the trellis on the open contour obtained by an initial Viterbi search are fixed as endpoints for a second Viterbi search. Furthermore, after each contour search, a new nanodisc centre can be estimated from the contour and used to determine a new annular search region. We iterated the search twice.



(a) (b) (c) (d)

Figure 3: (a) Part of a nanodisc image without gold markers and (b) its result. (c) Part of a nanodisc image with gold markers and (d) its result. Blue: correctly detected nanodiscs. Red: falsely detected nanodisc. Green: missed nanodiscs. Yellow: gold marker detections.

Using the inverse of the gradient magnitude at $\mathbf{v}_i$ as the external energy $E_{ext}(\mathbf{v}_i)$ will encourage the contour to lie on image edges but will cause problems when a part of another nanodisc is included in the search space (as in Figure 5). Similarly to [1] $E_{ext}(\mathbf{v}_i)$ can be modified to instead use the component of the image gradient in the radial direction (*i.e.*, projected onto a vector from the hypothesised nanodisc centre to the image point being considered). This encourages the contour to lie on image edges that are approximately in the radial direction and which are consistent with an object that is darker than the background. This method is referred to as the *directional gradient* method.

Nanodiscs are sometimes quite far from circular and their contour directions can then sometimes by poorly approximated by the radial direction. This can result in the *directional gradient* method giving poor results. Therefore, we tried a third external energy which combines the gradient magnitude and the directional gradient:

$$E_{ext} = w|\nabla_{directional}| + (1-w)|\nabla| \tag{3}$$

where $|\nabla|$ is the gradient magnitude, $\nabla_{directional}$ is the directional gradient, and $w \in [0,1]$ is a weight parameter which we set as $w = 0.7$.

4 Experiments

Nanodiscs and immunogold markers were manually annotated by C. Hacker in a data set consisting of five nanodisc images with immunogold markers (dataset A) and five without immunogold markers (dataset B). Images were 1024×1024 pixels and forbidden lines were placed at a distance of 10% of image width from the image border (see Figure 1). Dataset A had 304 nanodiscs and 15 immunogold markers annotated. Dataset B had 199 nanodiscs. Manual annotation took approximately 12 minutes per image in dataset A and 7 minutes per image in dataset B. A Matlab implementation of the proposed method took approximately $90s$ per image.

Figure 3 shows an example cropped result for each dataset. The curves in Figure 4, obtained by varying the symmetry response threshold, τ_{detect}, show false positive and false negative nanodisc detection rates per 100 annotated nanodiscs. The segmentation of each correctly detected nanodisc was compared to its manual segmentation using the Jaccard Index, $J = \frac{A_a \cap A_s}{A_a \cup A_s}$, where A_a and A_s are the sets of pixels in the manual and automatic segmentations respectively. Table 1 compares the results of the proposed segmentation method with

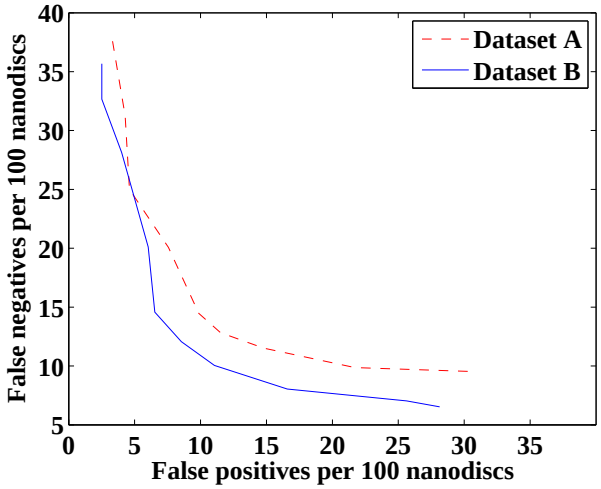


Figure 4: False positives and false negatives per 100 nanodiscs.

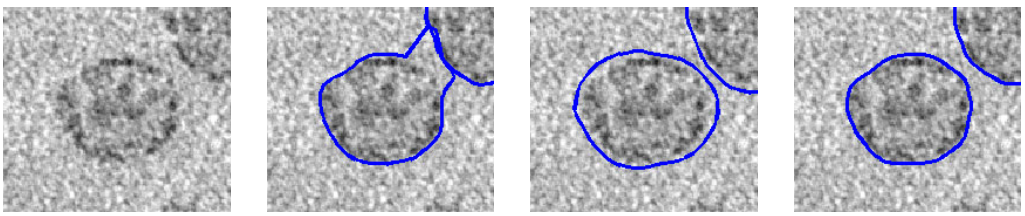
	Dataset A	Dataset B
Gradient magnitude	0.74 (0.12)	0.75 (0.10)
Directional gradient	0.69 (0.12)	0.69 (0.10)
Gradient magnitude + Directional gradient	0.76 (0.12)	0.77 (0.09)

Table 1: Mean Jaccard indices when comparing automatic and manual nanodisc segmentations (standard deviations in parentheses).

the method using only magnitude or directional gradient in the external energy. The proposed method compares favourably and Figure 5 shows an example where this is apparent. Adjacent nanodiscs can cause difficulties in distinguishing the contour in the overlapping regions; however, reasonable results can still be obtained as shown in Figure 6. All immunogold markers were detected successfully.

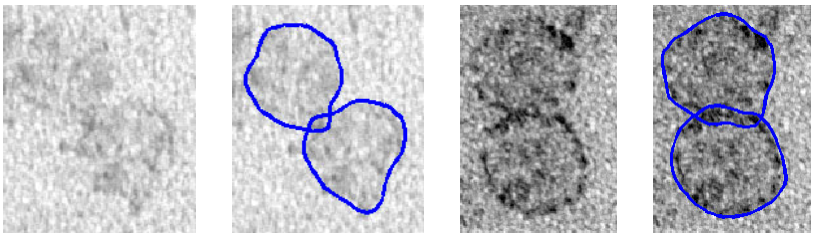
5 Conclusion

The experiments show that nanodiscs can by detected and segmented with reasonably high accuracy, including in the presence of immunogold markers. Detection failures occurred in



(a) (b) (c) (d)

Figure 5: (a) Example of cropped input, (b-d) results using gradient magnitude method, directional gradient method, and proposed combined method respectively.



(a) (b) (c) (d)

Figure 6: Segmentation of adjacent nanodiscs: (b) and (d) are the segmentations of (a) and (c) respectively.

cases where it was difficult to distinguish nanodiscs from background noise even for experts. We are currently investigating inter-observer variability. Two observers were asked to mark independently all the nanodiscs in a set of 10 images (five with and five without immunogold markers). Preliminary data shows 528 nanodiscs were marked by at least one of the two observers. Of these, 454 were marked by both observers. This gives a Jaccard Index of $454/528 \approx 0.86$. Taken in this context, the detection results reported in Section 4 are promising.

References

[1] P. Bamford and B. Lovell. Unsupervised cell nucleus segmentation with active contours. *Signal Processing*, 71(2):203 – 213, 1998.

[2] T. H. Bayburt and S. G. Sligar. Membrane protein assembly into nanodiscs. *FEBS Letters*, 584:1721–1727, 2010.

[3] H. Farid and E. P. Simoncelli. Differentiation of discrete multidimensional signals. *Image Processing, IEEE Transactions on*, 13:496–508, 2004.

[4] S. R. Gunn. *Dual active contour models for image feature extraction*. PhD thesis, University of Southampton, 1996.

[5] G. Loy and A. Zelinsky. Fast radial symmetry for detecting points of interest. *PAMI, IEEE Transactions on*, 25:959–973, 2003.

[6] J. Ni, M. K. Singh, and C. Bahlmann. Fast radial symmetry detection under affine transformations. In *CVPR*, pages 932–939, 2012.

[7] R. O. Ryan. Nanodisks: hydrophobic drug delivery vehicles. *Expert Opinion on Drug Delivery*, 5(3):343–351, 2008.

[8] N. Timilsina, C. Moffatt, and K. Okada. Development of a stained cell nuclei counting system. In *Medical Imaging: Image Processing*, volume Proc. SPIE Vol. 7962, pages 79620K–1, 2011.

[9] R. Wang, H. Pokhariya, S. J. McKenna, and J. Lucocq. Recognition of immunogold markers in electron micrographs. *Journal of Structural Biology*, 176:151–158, 2011.