Detecting and Tracking the Tips of Fluorescently Labeled Mitochondria in U2OS Cells

Eero Lihavainen[1(✉)], Jarno Mäkelä[1], Johannes N. Spelbrink[2], and Andre S. Ribeiro[1]

[1] Tampere University of Technology, Tampere, Finland
{eero.lihavainen,jarno.makela,andre.ribeiro}@tut.fi
[2] Nijmegen Centre for Mitochondrial Disorders, Radboud University Medical Center, Nijmegen, The Netherlands
hans.spelbrink@radboudumc.nl

Abstract. We present a method for automatically detecting the tips of fluorescently labeled mitochondria. The method is based on a Random Forest classifier, which is trained on small patches extracted from confocal microscope images of U2OS human osteosarcoma cells. We then adopt a particle tracking framework for tracking the detected tips, and quantify the tracking accuracy on simulated data. Finally, from images of U2OS cells, we quantify changes in mitochondrial mobility in response to the disassembly of microtubules via treatment with Nocodazole. The results show that our approach provides efficient tracking of the tips of mitochondria, and that it enables the detection of disease-associated changes in mitochondrial motility.

Keywords: Mitochondria · Detection · Tracking · Image analysis

1 Introduction

Mitochondria are involved in many cellular processes, and their dysfunctions have been linked to several diseases. In particular, abnormal mitochondrial dynamics such as an increased rate of fission, have been reported in the case of neurodegenerative diseases (see [2]).

In order to better understand the underlying mechanisms behind abnormal mitochondrial dynamics, it is necessary to analyze time-lapse image data from a large number of cells. So far, studies have relied on qualitative descriptions of mitochondrial movement [3] and manual image analysis [11], which limit the amount of data that can be analyzed. For more detailed studies, e.g. focusing on how interactions may affect the mobility, automatic image analysis methods are needed.

Previously described methods for automatic quantification of mitochondrial motion have mostly been restricted to measuring instantaneous velocity distributions using e.g. Optical Flow estimation [9] among other techniques [1]. Such methods yield no information about long-term dynamics of individual mitochondria. For tracking individual mitochondria, Silberberg et al. [12] applied a

V. Murino and E. Puppo (Eds.): ICIAP 2015, Part II, LNCS 9280, pp. 363–372, 2015.
DOI: 10.1007/978-3-319-23234-8_34

particle tracking method, consisting of the detection of mitochondria and a subsequent tracking step. A limitation of their detection method is that it assumes that the mitochondria appear globular, which is not true in general, as mitochondria often exhibit elongated and networked morphologies.

As mitochondria are similar in appearance to other elongated cellular structures, such as cytoskeletal filaments, when imaged with a fluorescence microscope, methods for tracking such filaments should be applicable to tracking mitochondria as well. For tracking cytoskeletal filaments, active contour tracking methods have been used [13]. These methods have the disadvantage of requiring the adjustment of several, non-intuitive parameters. In addition, methods for tracking the tips of microtubules have recently been proposed [5,6], but they rely on an initial manual detection of the tips.

In this work, we present a novel, automatic approach for detecting the tips of mitochondria, and apply the tracking framework of [8] to track the detected tips. Our detection method is based on supervised learning, namely a Random Forest classifier. Previous methods for automatically detecting the tips of microtubules [10] or mitochondria [9] have relied on segmentation, by applying filters that enhance curvilinear structures, binarizing the filtered image via thresholding, and extracting the tips from the morphological skeleton of the binarized image. Often such approaches will either over- or undersegment parts of the mitochondrial structure, which leads to false positives and misses. In contrast, directly detecting the tips should lead to a more robust method.

Here, we present the method and its validation using synthetic data. We also present a comparison of the method to a segmentation-based approach, and find it more reliable. Finally, we demonstrate the applicability of the method to experimental data, by measuring changes in mitochondrial motility caused by treatment with Nocodazole [14,15].

2 Materials and Methods

2.1 Image Acquisition

We transfected U2OS cells with a vector expressing mitoDsRED2, a red fluorescent protein targeting the mitochondrial matrix. The nuclei were labeled with the Hoechst 33342 fluorescent dye. The images were acquired with a Nikon Eclipse Ti-E with 100x Apo, a Wallac-Perkin Elmer Ultraview spinning-disk confocal system, Andor EMCCD camera, and a Nikon PFS autofocus system.

Prior to imaging, the cells were treated with $5\mu g/ml$ Nocodazole. We then selected four cells to be imaged. At 0, 30, 60, 90, 120 and 180 minutes after the application of Nocodazole, we imaged one optical slice of each cell every 3 seconds, for 10 minutes. This resulted in 6 movies of 10 minutes for each cell.

2.2 Training and Test Data

We selected one representative image of a cell not affected by Nocodazole, in which the sizes and appearances of mitochondria varied widely. In this image,

we manually marked 50 points at the tips of mitochondria, and extracted square patches of size 9×9 around them; we will refer to these as *positive* patches. Next, we extracted 1000 patches at random points, to serve as *negative* examples. This random selection was justified, because less than 1% of the points in the image contain tips of mitochondria. Finally, we manually marked 50 points at non-tip locations that shared visual features with tips: points along the mitochondria filaments, borders between two mitochondria, and curved edges of mitochondria. Figure 1 shows examples of each of these three subsets of the training data.

Although we estimated that 50 patches suffice to cover most of the variation in the appearance of the tips of mitochondria, as well as of the non-tip regions, any particular appearance may be represented in few orientations. In order to make our detector invariant to orientation, we augmented the data set with transformed versions of each manually selected patch. In particular, we applied each of the symmetries of a square to the patch: First, we mirrored the patch horizontally. Second, we rotated both the original patch and its mirrored version by 90, 180 and 270 degrees. This amounted to 7 new patches for each old patch. The reason we selected these transformations instead of, e.g., rotations of arbitrary angles, is that they require no interpolation, and thus do not introduce artifacts to the patch.

After extracting the training data, we further manually extracted 25 positive and 25 negative patches to serve as test data.

2.3 Detecting the Tips of Mitochondria Using a Random Forest Classifier

Our detection method works by classifying each sub-patch of the image in a sliding window, using a binary Random Forest (RF) classifier. As features for the classifier, we use the pixel values of the patches, read in column-major order, and normalized to zero mean and unit variance, in order to achieve invariance to intensity scaling. That is, for an image patch $P_k = \{p_{ij}\}$, where $i \in 1..9$ and $j \in 1..9$ are the row and column indices, respectively, the corresponding unnormalized feature vector is

$$x_k = [p_{11}, p_{21}, \ldots, p_{91}, p_{12}, p_{22}, \ldots, p_{19}, p_{29}, \ldots p_{99}]\,, \tag{1}$$

and the final, normalized feature vector is obtained as

$$y_k = \frac{x_k - \langle x_k \rangle}{\mathrm{std}(x - \langle x_k \rangle)}, \tag{2}$$

where $\langle \cdot \rangle$ denotes the sample mean, *std* denotes the sample standard deviation.

This detection procedure results in a binary image, where each connected component (CC) corresponds to one tip of a mitochondrion. However, the image will also contain some CCs that are false positive detections. We verified by visual inspection that these are typically small, approximately $1 - 2$ pixels in size. Noting that there is some uncertainty to the exact location of a manually marked tip, true positive CCs should contain more pixels in the tip region.

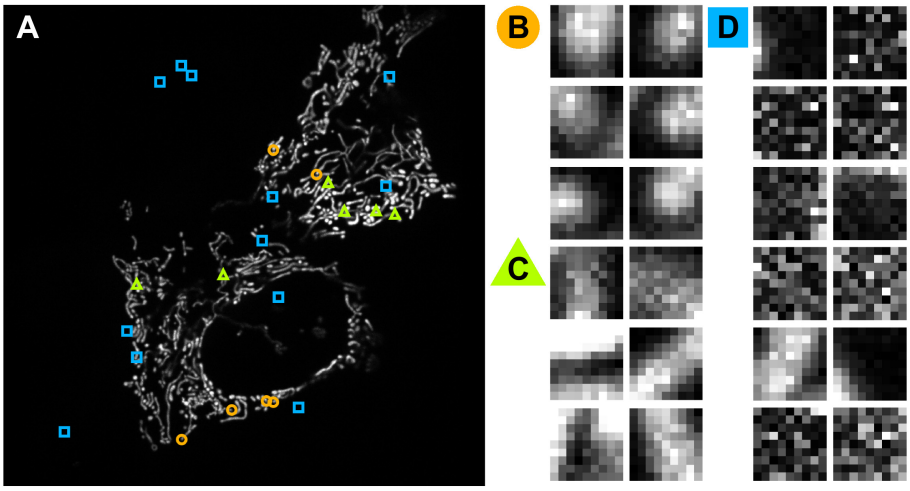

Fig. 1. Panel A shows the image from which the training data was extracted. The markers denote different types of training samples: manually picked positive samples (orange circles), manually picked negative samples (green triangles) and randomly selected negative samples (blue squares). The corresponding patches are visualized in panels B, C and D, where the colors and shapes correspond to the training sample types in panel A.

This is evident from the example classification results shown in Figure 2. Thus, as a post-processing step, we remove connected components that contain less than 3 pixels.

From the remaining CCs, we compute the centroids, which we use as the estimates of the tip locations. Examples of the final detection results are shown in Figure 2C.

2.4 Tracking

For tracking, we adopted the framework of [8], and used the authors' publicly available MATLAB implementation. In short, the method constructs tracks for detected objects in two steps: First, in each pair of subsequent frames, the objects are linked by solving a Linear Assignment Problem (LAP). If an object disappears temporarily, this procedure results in track segments instead of a complete track. Thus, as a second step, another LAP is solved to link track segments from the first step. A more detailed description of the method can be found in [8].

2.5 Generation of Synthetic Image Data

In order to test the performance of the method quantitatively, we needed movies for which the ground-truth location of each tip is known in each frame. To this end, we generated movies with simulated mitochondria. The advantages of using

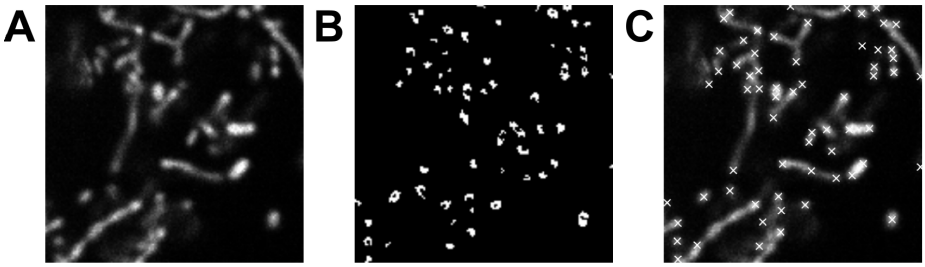

Fig. 2. A. Cropped region of a test image. B. Classification result. C. Tip locations (white crosses) detected as centroids of the connected components in B.

simulated images over manually-analyzed real images are that the ground-truth locations are free from human error, and that all parameters of the simulation can be varied to produce diverse test data.

To generate these synthetic movies, we modeled mitochondria as cubic splines, and subjected them to both brownian-like "wiggling" motion, and translation by applying a single random displacement vector to all control points of a mitochondrion. To simulate the wiggling, we first generated a random displacement vector for each control point. We then averaged the displacement vectors of nearby control points, in order to restrict the movements of the mitochondrial filaments to a realistic level of rigidity. To simulate the flat morphology of the U2OS cells, the distances moved by the mitochondria in the z-dimension were on average 1% of the distances they moved in the x- and y-dimensions.

The imaging process was modeled as follows: first, we rasterized the splines. To the resulting three-dimensional binary image, we added Poisson noise in order to generate variability in intensity inside the mitochondria filaments. Next, we convolved the noisy image with a Gaussian approximation of a fluorescence microscope point-spread-function. Finally, we added Gaussian noise to the image to simulate noise from the imaging system. Figure 3 shows example frames from one synthetic movie, as well as the locations of the true and detected tips.

3 Results

3.1 Performance of Tip Detection

For the RF, we trained 50 trees, and the number of features per split was selected to be 8. The latter parameter was selected using cross validation, by maximizing the area-under-ROC-curve. On the test data set, the classifier correctly classified 20/25(80%) of the positive examples, and 24/25(96%) of the negative examples.

We also compared the detection performance to that of a previous approach [9] based on segmentation. Briefly, the alternative method, here referred to as *SEG*, finds tips in 3 steps: first, it applies a median filter and a morphological

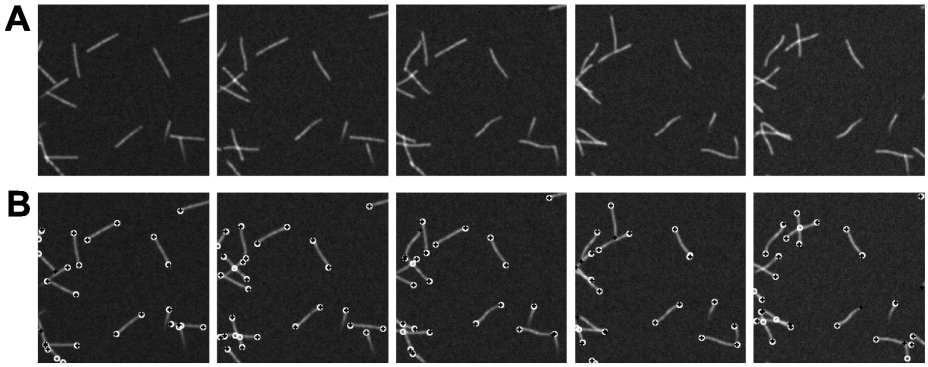

Fig. 3. A. Frames 0, 10, 20, 30 and 40 from a synthetic movie. B. The same frames with ground truth points (black crosses) and detected tips (white circles) overlaid.

top-hat filter to the image, to enhance the mitochondria structures. Next, it segments the image via Otsu's thresholding. Finally, it applies morphological thinning to find the morphological skeleton, from which the tips can be uniquely identified.

The methods were compared by generating 11 synthetic images, each containing approximately 20 tips, and comparing the locations of the detected tips to the ground truth tip locations. Specifically, the detected tips were paired with the ground truth tips via a Linear Assignment Problem (LAP); pairings were only made when the distance between the points was less than X pixels – points lacking such a real pair were paired with dummy points in the LAP. From the pairings, we obtained false positives as unpaired detections, true positives as paired detections, and false negatives as unpaired ground truth points. From these, we computed the true positive rate (TPR) and positive predictive value (PPV), and the F1-score defined as the harmonic mean of PPV and TPR. The results for both methods are shown in Table 1. The proposed method had only a slightly smaller TPR compared to SEG, but a significantly larger PPV. Consequently, the proposed method also yielded a larger F1-score.

Table 1. Detection performance of proposed method and alternative method (SEG) on synthetic images

	Proposed	SEG
TPR	0.91	0.92
PPV	0.89	0.79
F1	0.90	0.85

3.2 Accuracy of Tracking

For validating the method, we generated 11 synthetic movies of 101 frames each, and stored the locations of the mitochondrial tips, to serve as the ground truth data. We quantified the tracking accuracy on the synthetic data using a similar approach to [4]: we paired each ground truth track with a *hypothesis track* generated by the tracking method. This was done via solving a Linear Assignment Problem (LAP), where the cost c of each pairing is defined as

$$c(g,h) = \sum_{i \in G \cap H} \min(||g_i - h_i||_2, d_{max}) + |G \ominus H| \cdot d_{max} \tag{3}$$

Here, we used the following definitions: $|\cdot|$ denotes set cardinality, and $\ominus$ is the symmetric difference between sets. We defined $d_{max} = 2$ pixels to be the maximum distance between a ground truth point and a hypothesis point at which we can consider a point detected. G and H are sets that contain the indices of the movie frames in which a ground truth track and the paired hypothesis track are present (the former corresponding to frames where the object is visible, i.e. in focus), g and h are ground truth and hypothesis tracks, and the track points in frame i are denoted by g_i and h_i.

The effect of the latter summand in Eq. 3 is to increment c by d_{max} for all ground truth points that do not have a matching hypothesis point (misses), and all the hypothesis points that do not have a matching ground truth point (false detections). In the solution of the LAP, a pairing was not allowed between tracks that had no points less than d_{max} pixels apart; unpaired tracks were handled by assigning them to dummy elements in the LAP.

From the paired tracks, we wanted to answer the following questions: How likely is a hypothesis track to be paired with a ground truth track? Also, how likely is a true object track to be detected by the tracking method? To address these questions, we defined the following error measures, which were calculated over all the tracks in the data set:

$$E_1 = 1 - \frac{\#\{\text{hypothesis tracks paired with a ground truth track}\}}{\#\{\text{hypothesis tracks}\}} \tag{4}$$

$$E_2 = 1 - \frac{\#\{\text{ground truth tracks paired with a hypothesis track}\}}{\#\{\text{ground truth tracks}\}}. \tag{5}$$

In addition, we wanted to quantify the quality of the detected tracks. To this end, we asked: in a correctly detected true track, how likely is a point to be missed or assigned to the wrong object? For this, we calculated the following:

$$E_3 = \frac{\sum_{i \in G} \mathcal{I}(i \notin H)}{|G|} \tag{6}$$

$$E_4 = \frac{|H \setminus G| + \sum_{i \in G \cap H} \mathcal{I}(||g_i - h_i||_2 > d_{max})}{|H|}. \tag{7}$$

Here, $\mathcal{I}$ denotes the indicator function. The quantity E_3 measures the fraction of ground truth points that were missed, and E_4 measures the fraction of hypothesis points that were assigned to a wrong object.

The results from the experiment were as follows: E_1, 0.09; E_2, 0.29; E_3, 0.4 and E_4, 0.09. Thus, 9% of the hypothesis tracks were not paired; 29% of the true tracks were not detected; 40% of the ground truth points were missed; and 9% of the hypothesis points were assigned to a wrong object. We confirmed by visual inspection that the high values for E_2 and E_3 were largely due to failures of the detection method. Another source of missed points was that some track segments were not linked by the tracking method.

3.3 Quantification of Mitochondrial Motion in U2OS Cells

After validating the method on synthetic data, we tested whether it can quantify changes in mitochondrial movement caused by Nocodazole, which causes the depolymerization of microtubules (see e.g. [14]). Since intracellular transport of mitochondria occurs, at least to some extent, along microtubules (see [7]), the application of Nocodazole should result in reduced mobility of mitochondria. Such a reduction in mobility has been observed in various cell types [14,15]. We can quantify this effect with the proposed method.

To this end, we computed the mean speed for each track as the mean displacement between movie frames. Figure 4 shows the distributions of the mean speeds in each cell, for the first and the last movie, and the medians of these distributions in all movies. Between the first and the last movie, the median speed has decreased by 14–44%, with the difference being statistically significant for each cell ($P < 0.01$, Wilcoxon rank-sum test, N=416–755).

4 Discussion

With the ongoing development of live, single-cell time-lapse imaging techniques, the objective analysis of the kinetics of mitochondria is expected to play a significant role in the detection of mitochondria-related diseases, among other. This will require the use of image analysis tools capable of automatic detection and tracking mitochondria.

We demonstrated that our method can detect the tips of fluorescently labeled mitochondria with reasonable accuracy. In addition, we showed that the direct detection of tips using the new method results in fewer false positives than a segmentation-based approach, as indicated by the higher PPV.

In our tests with synthetic data, we found that the detection and tracking approach proposed here tends to result in incomplete tracks, as well as completely missed tracks. This is in part due to some tips being missed by the detection method, and also due to the tracking method sometimes failing to link track segments. Still, as long as the application does not require complete tracks, or unique tracks for each object, the method produces reliable results in this regard.

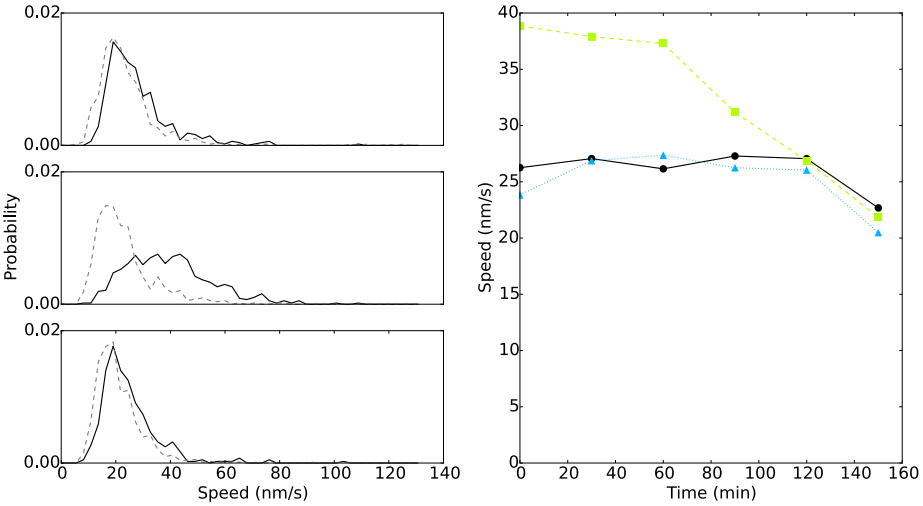

Fig. 4. Left: normalized histograms of mean speeds for cells 1-3 (top to bottom). The solid and dashed lines correspond to the movies captured 0 minutes and 180 minutes after application of Nocodazole, respectively. Right: medians of mean speeds for cells 1 (circles), 2 (squares) and 3 (triangles); each data point corresponds to one movie.

The results also showed that, for a correctly identified track, although many points will be missed (relating to the aforementioned issues), the tracks consist mainly of correctly identified points. This, along with the previous result, suggests that the method's results can be trusted.

Finally, by applying the method to images of U2OS cells treated with Nocodazole, we detected a decrease in mitochondrial motility in all cells; this result is consistent with previous studies [14,15].

For a more complete characterization of mitochondrial dynamics, it would be useful to keep track of not only the tips, but the whole mitochondrial filaments. This should be feasible by, for example, coupling the present method with an active contour segmentation method.

Finally, we expect our approach to be applicable to tracking other subcellular structures with similar shapes, such as microtubules and other cytoskeletal filaments, as well.

References

1. Beraud, N., Pelloux, S., Usson, Y., Kuznetsov, A.V., Ronot, X., Tourneur, Y., Saks, V.: Mitochondrial dynamics in heart cells: very low amplitude high frequency fluctuations in adult cardiomyocytes and flow motion in non beating HL-1 cells. Journal of bioenergetics and biomembranes **41**(2), 195–214 (2009)
2. Chen, H., Chan, D.C.: Mitochondrial dynamics-fusion, fission, movement, and mitophagy-in neurodegenerative diseases. Human Molecular Genetics **18**(R2), R169–76 (2009)

3. Chen, H., Detmer, S., Ewald, A.J., Griffin, E.E., Fraser, S.E., Chan, D.C.: Mitofusins Mfn1 and Mfn2 coordinately regulate mitochondrial fusion and are essential for embryonic development. The Journal of cell biology **160**(2), 189–200 (2003)
4. Chenouard, N., Smal, I., Chaumont, F.D.: Objective comparison of particle tracking methods. Nature Methods **11**(3), 281–289 (2014)
5. Demchouk, A.O., Gardner, M.K., Odde, D.J.: Microtubule Tip Tracking and Tip Structures at the Nanometer Scale Using Digital Fluorescence Microscopy. Cellular and molecular bioengineering **4**(2), 192–204 (2011)
6. Hadjidemetriou, S., Toomre, D., Duncan, J.: Motion tracking of the outer tips of microtubules. Medical image analysis **12**(6), 689–702 (2008)
7. Hales, K.G.: The machinery of mitochondrial fusion, division, and distribution, and emerging connections to apoptosis. Mitochondrion **4**(4), 285–308 (2004)
8. Jaqaman, K., Loerke, D., Mettlen, M.: Robust single-particle tracking in live-cell time-lapse sequences. Nature methods **5**(8), 695–702 (2008)
9. Lihavainen, E., Mäkelä, J., Spelbrink, J.N., Ribeiro, A.S.: Mytoe: automatic analysis of mitochondrial dynamics. Bioinformatics **28**(7), 1050–1 (2012)
10. Saban, M., Altinok, A., Peck, A., Kenney, C., Feinstein, S., Wilson, L., Rose, K., Manjunath, B.S.: Automated tracking and modeling of microtubule dynamics. In: 3rd IEEE International Symposium on Biomedical Imaging: Nano to Macro, pp. 1032–1035 (2006)
11. Saunter, C.D., Perng, M.D., Love, G.D., Quinlan, R.A.: Stochastically determined directed movement explains the dominant small-scale mitochondrial movements within non-neuronal tissue culture cells. FEBS letters **583**(8), 1267–73 (2009)
12. Silberberg, Y.R., Pelling, A.E., Yakubov, G.E., Crum, W.R., Hawkes, D.J., Horton, M.A.: Tracking displacements of intracellular organelles in response to nanomechanical forces. In: 2008 5th IEEE International Symposium on Biomedical Imaging: From Nano to Macro, pp. 1335–1338, May 2008
13. Smith, M.B., Li, H., Shen, T., Huang, X., Yusuf, E., Vavylonis, D.: Segmentation and tracking of cytoskeletal filaments using open active contours. Cytoskeleton **67**(11), 693–705 (2010)
14. Steinberg, G., Schliwa, M.: Organelle movements in the wild type and wall-less fz; sg; os-1 mutants of Neurospora crassa are mediated by cytoplasmic microtubules **564**, 555–564 (1993)
15. Yi, M., Weaver, D., Hajnóczky, G.: Control of mitochondrial motility and distribution by the calcium signal: a homeostatic circuit. The Journal of cell biology **167**(4), 661–72 (2004)