Computational Image Modeling for Characterization and Analysis of Intracellular Cargo Transport

Kuan-Chieh Chen[1,2], Minhua Qiu[1,2,4], Jelena Kovacevic[1,2,3], and Ge Yang[1,2]

[1] Department of Biomedical Engineering
[2] Center for Bioimage Informatics
[3] Department of Electrical Engineering, Carnegie Mellon University, Pittsburgh, PA 15213, USA
{kuanchic,jelenak,geyang}@andrew.cmu.edu
[4] Genomics Institute of the Novartis Research Foundation, San Diego, CA 92121, USA
minhua.qiu@gmail.com

Abstract. Active intracellular cargo transport is essential to survival and function of eukaryotic cells. How this process is controlled spatially and temporally so that the right cargo is delivered to the right destination at the right time remains poorly understood. To address this question, it is essential to characterize and analyze the molecular machinery and spatiotemporal behavior of intracellular transport. To this end, we developed related computational image models. Specifically, to study the molecular machinery of intracellular transport, we developed anisotropic spatial density kernels for reconstruction and segmentation of related super-resolution STORM (stochastic optical reconstruction microscopy) images. To study the spatiotemporal behavior of intracellular transport, we developed hidden Markov models and principal component analysis for representation and analysis of movement of individual transported cargoes. We validated and benchmarked the image models using simulated and actual experimental images. The models and related computational analysis methods developed in this study are general and can be used for studying molecular machinery and spatiotemporal dynamics of other cellular processes.

Keywords: image modeling, intracellular transport, spatiotemporal dynamics, super-resolution imaging, STORM imaging, spatial density estimation, hidden Markov model, principal component analysis.

1 Introduction

The inner environment of eukaryotic cells is highly dynamic and heterogeneous yet exquisitely organized. A basic principle of the organization is to compartment the intracellular environment into membrane-enclosed organelles. However, this creates a logistics problem because materials must be transferred between the organelles [1]. Cells solved this problem by utilizing an intracellular cargo transport system [2], which in many ways resembles the vehicular transport system of our cities.

Y.J. Zhang and J.M.R.S. Tavares (Eds.): CompIMAGE 2014, LNCS 8641, pp. 292–303, 2014.

Intracellular cargo transport is essential to survival and function of eukaryotic cells. This is especially evident in neurons, whose structure and function are highly polarized. A hallmark of the polarized structure of neurons is their long and thin axons, which in humans can extend for more than one meter at a diameter of a few micrometers. Because material synthesis and degradation are carried out mostly in the neuronal cell body, a wide variety of cargoes essential to the survival and function of neurons must be actively transported between the cell body and synaptic terminals [3]. Indeed, transport defects have been strongly implicated in the development of many human neurodegenerative diseases such as Alzheimer's disease [4].

A basic requirement for intracellular cargo transport is to deliver the right cargo to the right destination at the right time. This relies on spatial and temporal control of movement of individual cargoes. How this is achieved within cells remains largely unknown. To answer this question, it is essential to study the molecular machinery of intracellular transport as well as spatial and temporal behavior of transported cargoes. In related biological studies, the machinery and behavior of intracellular transport were often visualized using fluorescence microscopy techniques. It became evident from these studies that both are highly complex. Correspondingly, objects in images of intracellular transport exhibit highly complex structure and dynamics. Understanding the complex structure and dynamics of these image objects is essential to understanding the spatial and temporal control of intracellular transport. To achieve this goal, computational modeling of the image objects is essential.

In this paper we present our research results on using computational image modeling for characterization and analysis of intracellular transport. In the first part of the paper, we present results on using computational modeling of super-resolution STORM images for nanometer resolution characterization and analysis of molecular machinery of intracellular transport. In the second part of the paper, we present results on using computational modeling of time-lapse images for characterization and analysis of spatiotemporal behavior of intracellular transport.

1.1 Computational Image Modeling for Studying Molecular Machinery of Intracellular Transport

Active intracellular transport of cargoes is driven by molecular motors that walk along cytoskeletal filaments such as microtubules. Cargoes, molecular motors, and cytoskeletal filaments are therefore key components of the molecular machinery of intracellular transport. Dimensions of these components are often on the nanometer scale. Conventional fluorescence microscopy cannot fully resolve their structures because its resolution is limited to ~200nm by diffraction of visible light (Fig. 1A-B). This imposes a fundamental limitation on studying molecular machinery of intracellular transport. To overcome this limitation, we used STORM (stochastic optical reconstruction microscopy) [5], a super-resolution imaging technique that provides resolutions up to ~20nm. Because STORM relies on stochastic excitation of fluorophores [5], the resulting images of cellular structures are characterized by discontinuously distributed particles (Fig. 1C-D), which are significantly different

from the typically continuous region objects in conventional microscopy images (Fig. 1A-B). Computational image models are required for representing the spatial distribution of such particles so that the underlying molecular machinery of intracellular transport can be characterized and analyzed. In this study, we developed kernel-based spatial density models for reconstruction and segmentation of such particle images.

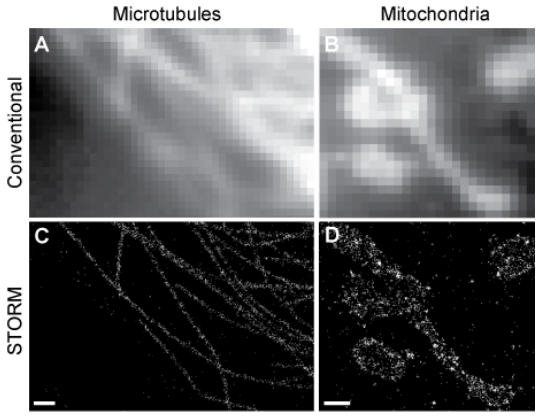

Fig. 1. Conventional fluorescence microscopy images versus super-resolution STORM images of microtubules and mitochondria, one among many different types of intracellular cargoes. (A, C) Conventional and STORM images of microtubules, respectively; (B, D) Conventional and STORM images of mitochondria, respectively. Scale bars: 500nm.

1.2 Computational Image Modeling for Studying Spatiotemporal Behavior of Intracellular Transport

In this study we focus on a specific model system of intracellular cargo transport (Fig 2A): the transport of amyloid precursor protein (APP) vesicles within axons of Drosophila third instar larvae. Vesicles are small intracellular cargoes that are enclosed by a lipid bilayer membrane. Drosophila has been used extensively to model human neurodegenerative diseases. The long and straight axons of Drosophila larvae (Fig. 2A) simplify imaging and data analysis and provide a powerful experimental system for studying intracellular cargo transport. To follow the movement of individual vesicles with high spatial resolution, we developed a single particle tracking method that recovers complete cargo trajectories at nanometer resolution (Fig. 2B-C) [6]. Recovered trajectories revealed complex spatiotemporal behavior of transported cargoes, such as pauses between movements, switches in movement directions, and changes in movement velocities [6]. In this study, we developed hidden Markov models (HMMs) [7] and principal component analysis (PCA) [8] for characterization and analysis of cargo behavior.

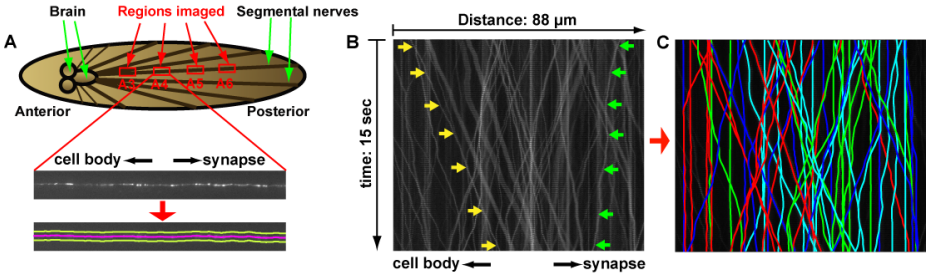

Fig. 2. Imaging and tracking axonal cargo transport in Drosophila third instar larvae. (A) Upper panel: regions selected for imaging. Middle panel: one frame from a time lapse video of APP vesicle transport. Lower panel: a band (5 pixels between green curves) following the axon (marked in cyan) is taken from each frame. (B) Bands from all frames were placed sequentially to generate a kymograph, a map of vesicle movement along the axon over time. Yellow and green arrows point to trajectories of two vesicles moving towards the synapse and the cell body, respectively. (C) Trajectories recovered by single particle tracking software were randomly colored and overlaid onto the kymograph for visual inspection.

1.3 Organization of the Paper

The rest of the paper is organized as follows: Section 2 presents results on using spatial density kernel based image models for reconstruction and segmentation of STORM images of molecular machinery of intracellular transport, specifically microtubules and mitochondria of mammalian BS-C-1 cells. Section 3 presents results on using hidden Markov models (HMMs) and principal component analysis (PCA) for representation and analysis of behavior of transported cargoes, specifically APP vesicles in neurons of Drosophila third instar larvae. Section 4 concludes with a summary as well as an outlook on ongoing work.

2 Computational Image Modeling for Studying Molecular Machinery of Intracellular Transport

2.1 Image Data Collection

Fluorescently labeled microtubules and mitochondria in fixed BS-C-1 cells were imaged using a Nikon N-STORM microscope with a 100×/1.41NA oil objective lens. Images were captured using an Andor Ultra 897 EMCCD camera, with a pixel size of 16 µm. Image collection was controlled using Nikon Element software. Image reconstruction was performed using custom software.

2.2 Spatial Density Kernel Models for Reconstruction of STORM Images

Resolution of conventional fluorescence microscopy is limited by diffraction of visible light to ~200nm. This resolution limit results from simultaneous excitation of

closely spaced fluorophores. When their point spread functions overlap substantially, it becomes infeasible to resolve them individually. STORM overcomes this limitation by randomly activating a small fraction of the total population of fluorophores so that closely spaced fluorophores are activated at different time points. This allows them to be resolved separately over time. Since different fluorophores are activated randomly over time while fluorophore labeling of cellular structure may not be complete, reconstructed STORM images are usually characterized by discontinuously distributed particles. By themselves, these particles only partially represent the underlying cellular structure. Consequently, conventional image analysis techniques such as segmentation techniques often give poor results if applied directly to STORM images due to lack of information (see Fig. 4 & 5 for examples). Solving this problem requires computational image models to fully extrapolate or reconstruct the underlying cellular structure. This extrapolation or reconstruction can also be considered as smoothing of STORM images.

In this study, we consider the distributed particles as random samples of an unknown continuous spatial density function that represents the underlying cellular structure. We propose to estimate this spatial density function using kernel based multivariate density estimation [9]. Specifically, we estimated the 2D spatial density function at location x using the following equation

$$\hat{p}(x) = \frac{1}{N}\sum_{i=1}^{N} K(x, x_i) \tag{1}$$

where $x_1, x_2, \ldots, x_N$ are the N nearest neighbors in the STORM image, and $K(\cdot)$ is a kernel function.

We started with a commonly adopted isotropic Gaussian kernel

$$K_i(x, x_i) = \frac{1}{2\pi\sigma^2}\exp\left(-\frac{\|x - x_i\|^2}{2\sigma^2}\right) \tag{2}$$

where σ is the standard deviation and a measure of kernel size. In this study, we chose σ empirically to be within the range of 10~50 nm to account for uncertainty of fluorophore position detection as well as incompleteness of cell structure labeling. Although this isotropic kernel based estimation approach is simple and straightforward to implement, it has several limitations. First, the estimation becomes unreliable when the distribution of particles becomes highly directional, as is for example the case near boundaries of objects. Second, even when the spatial distribution of particles is omnidirectional, the estimation becomes inaccurate when the distribution is spatially inhomogeneous. Third, since the same kernel size σ needs to be used over the image, empirical tuning is often required to reconcile variations in particle distribution within different image regions.

To overcome these limitations, we proposed an adaptive anisotropic kernel based approach in which we use the following anisotropic Gaussian kernel:

$$K_a(x, x_i) = \frac{1}{2\pi\det\left(\sum_{x_i}\right)}\exp\left(-\frac{1}{2}(x - x_i)^T \sum_{x_i}^{-1}(x - x_i)\right) \tag{3}$$

where the local covariance matrix $\sum_{x_i}$ at x_i is estimated using the following equation:

$$\sum_{x_i} = \frac{1}{N}\sum_{j=1}^{N}\left(x_j - x_i\right)\left(x_j - x_i\right)^T \quad (4)$$

where $x_1, x_2, \ldots x_N$ are the N nearest neighbors. N is typically chosen within 10~30 to balance the localization and reliability of estimation. We first used a simple example to compare the performance of isotropic kernels versus adaptive anisotropic kernels for spatial density estimation (Fig. 3A-D). A qualitative visual assessment confirmed that the anisotropic kernels outperform the isotropic kernels substantially.

To characterize the molecular machinery of intracellular transport, the corresponding image objects in reconstructed images must be segmented. Here, we used STORM images reconstructed from the estimation of spatial density function for segmentation. This also allowed us to further compare quantitatively performance of isotropic kernels versus anisotropic kernels in reconstruction of STORM images. As discussed previously, the kernel-based estimation of spatial density function also provides a way for smoothing raw STORM images. This allowed us to use a simple density thresholding scheme for image segmentation. We empirically modeled the density distribution of actual image objects by finding the non-zero density bins from the spatial histogram. The density threshold was set by finding the pth percentile of the spatial densities. We found p to be within the range of 1~15. We tested our segmentation method on both simulated and actual STORM images.

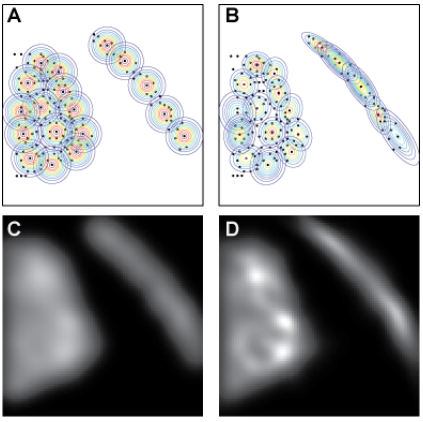

Fig. 3. Spatial density estimation using isotropic kernels versus adaptive anisotropic kernels. (A, B) Isotropic kernels and adaptive anisotropic kernels, respectively, overlaid on data points sampled from a curvilinear feature and a blob feature. (C, D) Estimated spatial density function images using isotropic kernels and adaptive anisotropic kernels, respectively.

2.3 Segmentation of Simulated STORM Images

Simulated STORM images were generated for microtubules and mitochondria, two key components of the molecular machinery of intracellular transport. Assuming a uniform fluorophore labeling density in object regions, we simulated random

activation of fluorophores at two densities: 1,000 and 2,000 activations in an image region of 2560nm×2560nm. Random background noise was added under different signal to noise ratios (SNRs), which was set to 20 dB in this paper. To provide a reference for comparison, we used an active contour algorithm for segmentation of simulated STORM images [10]. We found that boundaries of image objects segmented using this algorithm were very sensitive to missing particles due to random fluorophore activation (Fig. 4).

We then applied the previously described density thresholding algorithm for segmentation of images reconstructed using isotropic kernels and adaptive anisotropic kernels, respectively. Compared to the active contour segmentation, thresholding of images reconstructed using isotropic kernels provided better performance but still suffered from over-segmentation as well as false segmentation of background noise (Fig. 4; Table 1). Thresholding of images reconstructed using anisotropic kernels provided overall the best result. In particular, it performed robustly against random variations of activated fluorophores and provided more accurate segmentation for both sample structures.

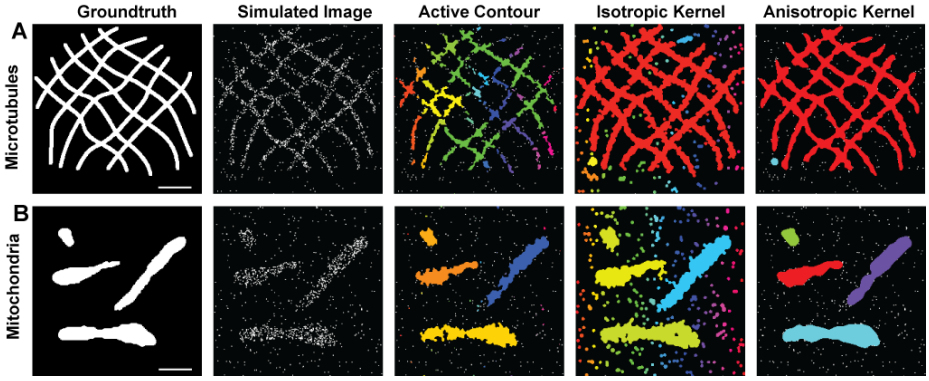

Fig. 4. Comparison of different image segmentation algorithms on simulated STORM images. Segmentation results are shown in random colors and overlaid onto simulated images. Scale bars: 500nm.

For quantitative comparison, we adopted performance metrics used in [11]: area similarity (AS), precision (P), and recall (R). AS measures the overlap of segmentation results relative to the ground truth. P is the ratio between the total number of correctly segmented objects to the total number of segmented objects. R is the ratio between the total number of correctly segmented objects to the total number of objects in the ground truth. Calculated metrics were averaged over all objects and all images and listed in Table 1. It is clear that thresholding of images reconstructed using anisotropic kernels provided overall the best result, especially in minimizing over-segmentation, as measured by precision P.

Table 1. Performance metrics of segmentation on simulated STORM images

	Microtubules			Mitochondria		
	AC	IKDES	AKDES	AC	IKDES	AKDES
AS	0.64	0.75	**0.79**	0.90	0.88	**0.90**
P	0.01	0.02	**0.37**	0.30	0.44	**0.96**
R	1.00	1.00	**1.00**	1.00	1.00	**1.00**

AC: active contour segmentation

IKDES: isotropic kernel based density estimation segmentation

AKDES: anisotropic kernel based density estimation segmentation

2.4 Segmentation of Actual STORM Images

We also compared the three segmentation algorithms on actual STORM images of microtubules and mitochondria in fixed BS-C-1 cells. Fig. 5 shows results in a region of 4000nm×4000nm. Similar to the case of simulated images, the active contour segmentation produced fragmented clusters of particles while the isotropic kernel based method suffered from over-segmentation as well as false segmentation of background noise. The adaptive anisotropic kernel based segmentation reliably identified image object boundaries and performed robustly against background noise.

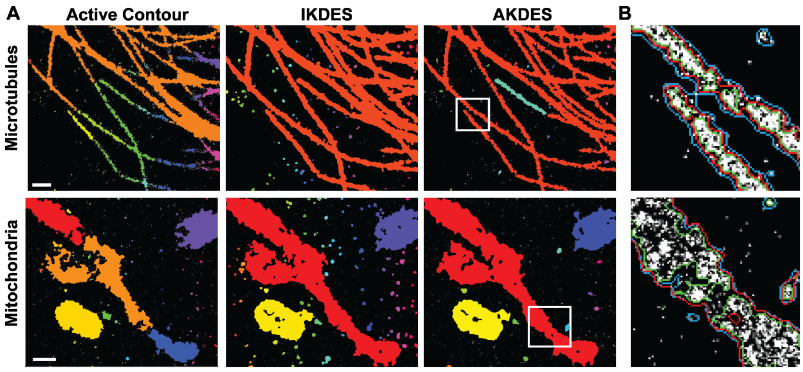

Fig. 5. Comparison of different image segmentation algorithms on actual STORM images. (A) Comparison of different image segmentation algorithms. (B) Comparison of segmented image object boundaries in regions corresponding to boxed regions in column 3 of panel A, where green, blue, and red contours show active contour, IKDES, and AKDES results. Scale bars: 500nm.

3 Computational Image Modeling for Studying Spatiotemporal Behavior of Intracellular Transport

3.1 Image Data Collection

Axonal transport of vesicles carrying amyloid precursor protein tagged with YFP (APP-YFP) was imaged in axons within segmental nerves of dissected Drosophila third instar larvae under different genetic modifications of kinesin and dynein

subunits. Each time-lapse movie was collected for 20 seconds at 11 frames per second using a Nikon Ti-E inverted microscope with a 100×/1.41NA oil objective lens and a Photometric CoolSnap HQ2 camera. Movie collection was controlled using Nikon Element software.

3.2 Computational Modeling of Vesicle Movement

Complete trajectories of vesicles were recovered at nanometer resolution using custom software [6]. At any time, individual vesicles may reside in one of the following states: moving towards a distal synaptic terminal (anterograde movement; denoted by *A*), moving towards the neuronal cell body (retrograde movement; denoted by *R*), or pausing (denoted by *P*). A vesicle may switch between movement and pause and between anterograde and retrograde movement.

To characterize vesicle movement, we first calculated and analyzed segmental velocities of individual vesicles. A *segment* is a part of the trajectory of a vesicle in which it moves consistently in one direction. It is usually a part of the trajectory between two pauses, two direction reversals, or a pause and a directional reversal. The segmental velocity is the average velocity of a vesicle within the given segment. Using model-based clustering [12], we determined that anterograde as well as retrograde segmental velocity always follow three different modes under the different genetic conditions analyzed [13] (Fig. 6A).

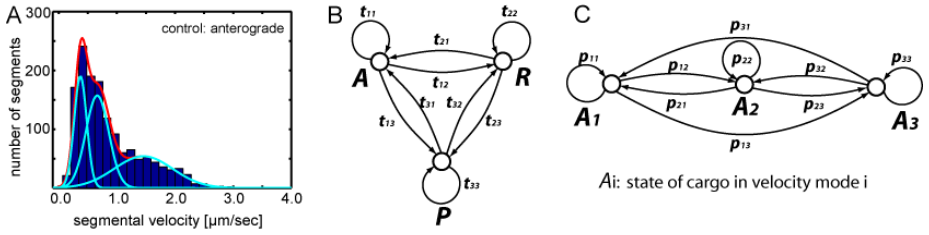

Fig. 6. Computational modeling of vesicle behavior. (A) Velocities of cargoes moving towards the synaptic terminal follow three modes (red: total distribution. cyan: individual modes). (B) A coarse scale hidden Markov model (HMM) of cargo behavior, which characterizes transitions between anterograde movement *A*, retrograde movement *R*, and pause *P*. (C) A fine scale hidden Markov model (HMM) of cargo behavior, which characterizes transitions between different anterograde velocity modes.

To characterize and analyze vesicle behavior, we used hidden Markov models [7, 14]. Depending on the research questions to be addressed, we modeled vesicle behavior at two levels. The coarse scale model (Fig. 6B) characterizes transition of cargo between pause (*P*) and anterograde movement (*A*) and retrograde movement (*R*). A fine scale model (Fig. 6C) allows us to determine the transition probabilities between different anterograde (or retrograde) velocity modes.

Here we give an example of estimating probabilities of transition between different anterograde velocities modes in wild-type (*wt*) animals versus animals with genetic modification of kinesin (kinesin heavy chain, khc^{8}/+) and dynein (dynein intermediate

chain, *GEN-DIC/+*), where $P_1 = p_{11} + p_{22} + p_{33}$ (Fig. 6C) is the probability of no mode change, $P_2 = p_{12} + p_{21} + p_{23} + p_{32}$ is the probability of shifting to neighboring states (i.e. $I \leftrightarrow II$, $II \leftrightarrow III$) and $P_3 = p_{13} + p_{31}$ is the probability of mode jump (i.e. $I \leftrightarrow III$). All computations were performed using the PMTK (Probabilistic Modeling Toolkit for MATLAB/Octave) package. The results, listed in Table 2, indicate that genetic modifications of molecular motors can influence transitions between different velocity modes. In this specific case, genetic reductions of kinesin and dynein reduced probabilities of switching between velocity mode I and mode II.

Table 2. Estimated transition probabilities between different anterograde velocity modes

	P_1	P_2	P_3
wt	0.59	0.38	0.03
***khc*^{8}/+**	0.79	0.20	0.01
GEN-DIC/+	0.79	0.18	0.03

3.3 Principal Component Analysis for Comprehensive Characterization of Cargo Behavior

In the previous section, we analyzed segmental velocities of transported vesicles. However, comprehensive characterization of transport behavior requires additional descriptors. Table 3 shows representative descriptors of vesicle behavior under three categories: movement, pause, and switch (i.e. reversal in movement direction). Descriptors under each category characterize vesicle behavior from different perspectives. Detailed definitions of these descriptors are given in [13].

Table 3. Representative descriptors of vesicle behavior

Anterograde movement descriptor	**Abbreviation**	**Retrograde movement descriptors**	**Abbreviation**
anterograde net velocity	aNetV	retrograde net velocity	rNetV
anterograde instantaneous velocity	aInstV	retrograde instantaneous velocity	rInstV
anterograde segmental velocity	aSegV	retrograde segmental velocity	rSegV
anterograde duration weighted segmental velocity	aDWSegV	retrograde duration weighted segmental velocity	rDWSegV
anterograde net run-length	aNet RL	retrograde net run-length	rNet RL
anterograde total run-length	aTotRL	retrograde total run-length	rTotRL
anterograde segmental pause frequency	aSegPF	retrograde segmental pause frequency	rSegPF
anterograde total pause duration	aTotPD	retrograde total pause duration	rTotPD
switch frequency*	SF	switch frequency**	SF

*,**: switch frequency are used as descriptors for both anterograde and retrograde movement.

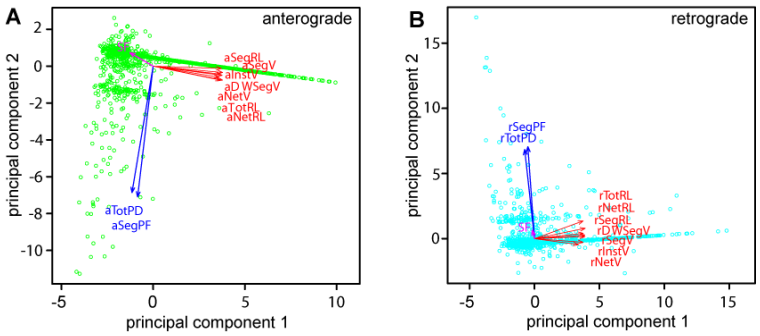

Fig. 7. Principal component analysis of cargo behavior descriptors. (A) anterograde descriptors and principal component, calculated based on n= 1745 measurements from 18 animals. (B) retrograde descriptors and principal components, calculated based on n= 1745 measurements from 18 animals.

Here two basic questions should be addressed. First, do descriptors under the same category provide the same information? Second, what are the relations between descriptors under different categories? To answer these questions, we used principal component analysis (Fig. 7 A-B). It showed that different descriptors under the same category, represented in Fig.7 using the same color, are highly correlated and thus provide similar information. For example, for both anterograde and retrograde movement, velocity descriptors and run-length descriptors are highly correlated and provide similar information, indicating that fast moving vesicles tend to travel long distances. On the other hand, descriptors under different categories are nearly orthogonal and thus uncorrelated. For example, pause descriptors are orthogonal and thus uncorrelated with movement descriptors, suggesting that pauses are likely caused by mechanisms different from those driving movement. Similarly, switch frequency are largely orthogonal to pause and motion descriptors, indicating that switches in movement are likely mediated by different mechanisms.

4 Summary and Outlook

In this study we developed and applied computational image models for studying the molecular machinery and spatiotemporal behavior of intracellular transport. Our results demonstrated the power of these models in representing complex cellular structure and dynamic behavior. The models and related computational analysis methods are general and applicable to studies of other cellular processes. We are addressing several limitations of this study in ongoing work. First, the anisotropic kernel cannot be used to model structures that spatially overlap with each other. We are addressing this limitation using Gaussian mixture models. Second, the hidden Markov models only describe ensemble cargo behavior. It does not directly represent spatial behavior of individual cargoes. We are addressing this limitation by developing statistical models that explicitly describe cargo spatial behavior. Third, the

axonal transport process is essential one dimensional whereas intracellular transport in non-polarized cells is two dimensional. We are extending our current one-dimensional cargo behavior models for two-dimensional intracellular transport.

Acknowledgments. We gratefully acknowledge support from the US National Science Foundation through grants 1017278 (J.K.), and DBI-1149494, DBI-1052925, MCB-1052660 (G.Y).

References

1. Wickner, W., Schekman, R.: Protein translocation across biological membranes. Science 310, 1452–1456 (2005)
2. Vale, R.D.: The molecular motor toolbox for intracellular transport. Cell 112, 467–480 (2003)
3. Brown, A.: Axonal transport of membranous and nonmembranous cargoes: a unified perspective. J. Cell Biol. 160, 817–821 (2003)
4. De Vos, K.J., Grierson, A.J., Ackerley, S., Miller, C.C.J.: Role of axonal transport in neurodegenerative diseases. Annu. Rev. Neurosci. 31, 151–173 (2008)
5. Rust, M.J., Bates, M., Zhuang, X.: Sub-diffraction-limit imaging by stochastic optical reconstruction microscopy (STORM). Nat. Meth. 3, 793–796 (2006)
6. Qiu, M., Yang, G.: Nanometer resolution tracking and modeling of bidirectional axonal cargo transport. In: Proc. IEEE Int. Symp. Biomedical Imaging (ISBI), Barcelona, Spain, pp. 992–995 (2012)
7. Rabiner, L.R.: A tutorial on hidden Markov models and selected applications in speech recognition. Proc. IEEE 77, 257–286 (1989)
8. Jolliffe, I.T.: Principal Component Analysis. Springer (2002)
9. Scott, D.W.: Multivariate Density Estimation. John Wiley & Sons (1992)
10. Chan, T.F., Vese, L.A.: Active contours without edges. IEEE Trans. Image Processing 10, 266–277 (2001)
11. Chen, K.C.J., Yu, Y., Li, R., Lee, H.-C., Yang, G., Kovacevic, J.: Adaptive active-mask image segmentation for quantitative characterization of mitochondrial morphology. In: 2012 19th IEEE Int. Conf. Image Processing (ICIP), pp. 2033–2036 (2012)
12. Fraley, C., Raftery, A.E.: Model-based clustering, discriminant analysis and density estimation. J. Am. Stat. Assoc. 97, 611–631 (2002)
13. Reis, G.F., Yang, G., Szpankowski, L., Weaver, C., Shah, S.B., Robinson, J.T., Hays, T.S., Danuser, G., Goldstein, L.S.B.: Molecular motor function in axonal transport in vivo probed by genetic and computational analysis in Drosophila. Mol. Biol. Cell 23, 1700–1714 (2012)