Detection and Tracking of Golgi Outposts in Microscopy Data

Huei-Fang Yang, Chu-Song Chen, Xavier Descombes

▶ To cite this version:

DETECTION AND TRACKING OF GOLGI OUTPOSTS IN MICROSCOPY DATA

Huei-Fang Yang[⋆]   *Chu-Song Chen*[⋆†]   *Xavier Descombes*[‡]

[⋆]Research Center for Information Technology Innovation, Academia Sinica, Taiwan
[†]Institute of Information Science, Academia Sinica, Taiwan
[‡]INRIA, CRI Sophia Antipolis Méditerranée, France

ABSTRACT

Golgi outposts (GOPs) that transport proteins in both the anterograde and retrograde directions play an important role in determining the dendritic morphology in developing neurons. To obtain their heterogeneous motion patterns, we present a data association based framework that first detects the GOPs and then links the detection responses. In the GOP detection stage, we introduce a multi-scale Markov Point Process (MPP) based particle detector that uses multi-scale blobness images obtained by Laplace of Gaussian (LoG) for GOP appearances. This reduces the number of missed detections compared to the use of image intensity for GOP appearances. In the linking stage, we associate detection responses to form reliable tracklets and link the tracklets to form long, complete tracks. As such, high-level information (e.g., motion) is encoded in building the affinity model. We evaluate our approach on the microscopy data sets of dendritic arborization (da) sensory neurons in Drosophila larvae, and the results demonstrate the effectiveness of our method.

Index Terms— particle detection, particle tracking, microscopy

1. INTRODUCTION

Dendrites are the tree-like structures of diverse branching patterns, and their morphologies are highly correlated to neuron function and connectivity. It has been established that Golgi outposts (GOPs) play a central role in shaping the dendritic morphogenesis in developing neurons [1]. Understanding how the GOP dynamics affect the dendritic branching can provide new insights into dendritic growth.

The acquired great amounts of biological image data provide a great way for the study of dendritic morphogenesis. Figure 1(a) shows the maximum intensity projection (MIP) of a 3D confocal microscopy volume of the dendritic arborization (da) sensory neurons in the Drosophila larval peripheral nervous system (PNS). The close-ups of the rectangular area in Figure 1(a) at different time steps are shown in Figure 1(b) - (c). The GOPs, appearing as bright spots, are a component of the secretory pathway in dendrites and are engaged in delivering proteins to the cell surface or out of the cell. While most of the GOPs move in both the anterograde (from the cell body to the dendrites) and retrograde (from the dendrites to the cell body) directions to transport the proteins, some may stop moving and become stationary (see Figure 1(b) - (c) where the blue arrows
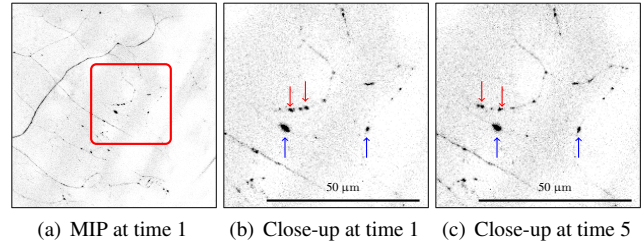
(a) MIP at time 1   (b) Close-up at time 1   (c) Close-up at time 5

Fig. 1: Maximum intensity projections (MIPs) of the confocal microscopy images of the Drosophila peripheral nervous system. (a) MIP of the original 3D volume at time 1. The size of the MIP is 512×512 pixels. (b) and (c) Close-ups of the rectangular area (200×200 pixels) in (a) at times 1 and 5, respectively. Red arrows point to the moving GOPs and blue arrows to the stalling ones. The GOPs differ not only in the size but also in the brightness. Image intensities are inverted and contrast is enhanced for visualization.

point to the stalling GOPs). This motion heterogeneity increases the difficulty in tracking the dendritic GOPs. Automated, sophisticated tracking methods that reduce the human effort are desirable.

A key step in tracking biological particles lies in establishment of particle correspondences in an image sequence. Theoretically, globally evaluating all possible associations throughout the whole image sequence can provide the most accurate results, but this is computationally intensive and impractical. Therefore, earlier methods resort to resolving the correspondences from frame to frame. The two-frame association approaches mainly rely on the spatial information (e.g., particle spatial positions) and do not deal well with the particle occlusion. Methods performing data association through multiple time steps can better cope with the particle occlusion because more temporal information is available [2, 3]. This strategy has also been incorporated in a Multiple Hypothesis Tracking (MHT) approach [4] that performs data association on a sliding window basis. These approaches have successfully tracked several types of biological particles, such as receptors and Golgi units. However, the GOPs exhibit different motion patterns from the biological objects they dealt with, and directly applying these methods to GOP tracking may not yield satisfactory results (see Section 4). This is mainly because high-level information, such as motion, may be required to resolve the correspondence ambiguities between GOPs.

In our GOP tracking framework, we first propose a particle detection method based on marked point processes (MPP) for identification of the GOPs on each frame in an image sequence (Section 2). We use multi-scale blobness images obtained by the Laplace of Gaussian filter to represent the GOP appearances; hence, our detector is effective for detecting GOPs of different sizes. Second, we propose a hierarchical data association method that first links detection
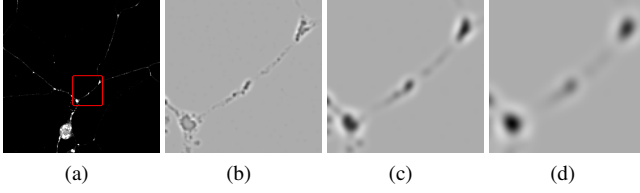
Fig. 2: Multi-scale LoG-filtered images. (a) One sample microscopy image. (b), (c) and (d) Close-ups of the LoG-filtered images of the rectangular area in (a) at scales 1, 3, and 5, respectively.

responses to form reliable tracklets (i.e., short tracks) and then links the tracklets to obtain complete trajectories (Section 3). The advantage of linking tracklets instead of single detection responses is that high-level information, such as motion, can be encoded in building the affinity model. We use time-lapse confocal microscopy images of the Drosophila larval peripheral nervous system (PNS) for evaluation (Section 4). The experimental results show that our method can detect spot-like particles and obtain their moving trajectories better than other competitive methods.

2. MULTI-SCALE MPP-BASED PARTICLE DETECTION

The first step of our multi-scale particle detection is to apply the Laplacian of a Gaussian (LoG) with different kernels to the image. This step aims to represent the objects to be detected at different scales. The second is to search for prominent features across all scales and determine the positions and scales of the objects. We use a Multiple Birth and Cut (MBC) algorithm based on marked point processes (MPP) [5] for the search task.

2.1. Scale-space Blobness Representation

Given an image $I(x, y)$, its scale space representation $L(x, y, \sigma)$ is computed by convolving it with a Gaussian $g(x, y, \sigma)$: $L(x, y, \sigma) = g(x, y, \sigma) * I(x, y)$, where $*$ denotes the convolution operator, and the Gaussian function g with standard deviation σ is $g(x, y, \sigma) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}}$. To highlight the blob-like objects in an image, we apply the scale-normalized Laplace operator to L:

$$\nabla^2_{\text{norm}} L(x, y, \sigma) = \sigma^2 \left(\frac{\partial^2 L}{\partial x^2} + \frac{\partial^2 L}{\partial y^2} \right), \quad (1)$$

where ∇ is the Nabla operator. Because of the associative property in the convolution, Equation (1) is equivalent to directly applying the LoG filter to the image I. The scale-normalized LoG is given by

$$\nabla^2_{\text{norm}} g = \left(x^2 + y^2 - \sigma^2 \right) e^{-\frac{x^2+y^2}{2\sigma^2}}. \quad (2)$$

By varying the scale parameter σ, we can obtain the multi-scale LoG-filtered images. Figure 2 shows an image and its LoG-filtered ones produced by varying the scales. The circular blob structures are highlighted in the LoG-filtered images.

2.2. MPP-based Scale Selection

We use circles to model particles. A circle is denoted by 3 parameters (x, y, r), where (x, y) are the coordinates of the center, and r is the radius. Using such a representation, our goal becomes to locate an unknown number of circle-shaped objects.

Let $\mathcal{C}$ denote the sets of configurations and $\omega = \{\omega_1, \ldots, \omega_n\}$ denote a configuration that contains n objects (circles, in our case). We aim at finding a configuration ω that minimizes the following:

$$E(\omega) = \sum_{\omega_i} U(\omega_i) + \sum_{\omega_i \sim \omega_j} V(\omega_i, \omega_j), \quad (3)$$

where $U(\omega_i)$ is the unary data term measuring the degree to which an object ω_i fits the image data, $\sim$ denotes the neighborhood relationship between two objects, and $V(\omega_i, \omega_j)$ is the piecewise interaction term (i.e. prior term) that takes into account the spatial interaction between two neighboring objects.

Data Term The data term measures how well a circle fits to the image data and plays a role in determining what scale should be selected. According to Equation (2), the zero-crossings of the LoG occur at $x^2 + y^2 = 2\sigma^2$, which corresponds to a circle of radius $\sqrt{2}\sigma$. In other words, for a given circle ω_i of radius r_i, we use the LoG-filtered image at scale $r_i/\sqrt{2}$ to compute the data term that compares the levels of blobness between the circle interior and its background. More specifically, letting R_{in} be the circle interior and R_{out} be its boundary, we compute the Bhattacharyya distance between them:

$$d_{\omega_i}(R_{\text{in}}, R_{\text{out}}) = \frac{1}{4} \frac{(\mu_{\text{in}} - \mu_{\text{out}})^2}{\sigma_{\text{in}}^2 + \sigma_{\text{out}}^2} + \frac{1}{2} \ln \frac{\sigma_{\text{in}}^2 + \sigma_{\text{out}}^2}{2\sigma_{\text{in}}\sigma_{\text{out}}}, \quad (4)$$

where the symbols μ and σ represent the mean levels of blobness and standard deviation of a region, respectively.

Piecewise Interaction Term Our target configuration is a set of non-overlapping objects so we penalize the objects that are spatially overlapped and give no penalty to the non-overlapping ones. This leads us to define the piecewise term as

$$V(\omega_i, \omega_j) = \begin{cases} \infty & \text{if } A_{\omega_i} \cap A_{\omega_j} = \varnothing \\ 0 & \text{otherwise} \end{cases}, \quad (5)$$

where A_{ω_i} and A_{ω_j} denote the areas of objects ω_i and ω_j, respectively.

Minimization of Equation (3) is performed via a Multiple Birth and Cut algorithm [5]. The algorithm starts with an initial configuration $\omega^{(0)}$ and then iterates between the birth and cut steps till convergence.

3. HIERARCHICAL DATA ASSOCIATION

Before elaborating on our method, we list the notations used through the paper:

- $o_i^t = (p_i^t, s_i^t, a_i^t)$: a detection response at time t consisting of position $p_i^t = (x_i^t, y_i^t)$, size s_i^t, and appearance a_i^t. We use the radius of a detected circle obtained by the proposed detector in Section 2 to represent the size of a GOP. The GOP appearances are represented by the mean intensity of the circle interior with a fixed radius (we set it to the mean radius of the detected GOPs in our experiments);

- $\mathcal{O} = \{o_i^t\}$: the whole detection set;

- $F_k = \left\{ o_{k_i}^{t_{k_i}} \mid \forall t_{k_i} < t_{k_{i+1}}, 1 \leq t_k^s \leq t_{k_i} \leq t_k^e \leq t_{\text{total}} \right\}$, where t_k^s and t_k^e are the start and end times of F_k, and t_{total} is the length of an image sequence: a tracklet is made up of detection responses in ascending order by their time stamps;

- $X = \{F_k\}$: a set of tracklets;
- $T_k = \{F_{k_1}, F_{k_2}, \cdots, F_{k_l}\}$: a trajectory hypothesis whose constituents are tracklets in ascending order;
- $\mathcal{T} = \{T_k\}$: a set of trajectory hypotheses.

3.1. Low-level Tracklet Generation

We tailor a dual threshold strategy [6] to generate the initial tracklets. The affinity of detection responses between consecutive frames is defined as:

$$A^0 \left(o_j^{t_j}|o_i^{t_i}\right) = \begin{cases} A_{ij}^{\mathrm{pos}} A_{ij}^{\mathrm{size}} A_{ij}^{\mathrm{app}} & t_j - t_i = 1 \\ 0 & \text{otherwise,} \end{cases} \quad (6)$$

where $A_{ij}^{\mathrm{pos}} = \exp(-\frac{||p_i^{t_i}-p_j^{t_j}||^2}{\sigma_d})$, $A_{ij}^{\mathrm{size}} = \exp(-\frac{|s_i^{t_i}-s_j^{t_j}|}{\sigma_s})$, and $A_{ij}^{\mathrm{app}} = \exp(-\frac{(a_i^{t_i}-a_j^{t_j})^2}{\sigma_a})$. To obtain reliable tracklets, the association of two detection responses is attained if the following equations are satisfied:

$$\begin{aligned} & A^0 \left(o_i|o_j\right) > \theta_1, \text{ and } \forall o_k \in \mathcal{O} - \{o_i, o_j\}, \\ & \min\left(A^0\left(o_i|o_j\right) - A^0\left(o_k|o_j\right), A^0\left(o_i|o_j\right) - A^0\left(o_i|o_k\right)\right) > \theta_2 \\ & \text{subject to } \theta_1 > \theta_2 > 0. \end{aligned}$$
$$(7)$$

3.2. High-level Tracklet Association

The goal of our tracklet association is to find the most probable trajectory set $\mathcal{T}$ given X:

$$\mathcal{T}^* = \arg\max_{\mathcal{T}} P\left(\mathcal{T}|X\right) = \arg\max_{\mathcal{T}} P\left(X|\mathcal{T}\right) P\left(\mathcal{T}\right) \quad (8)$$

$$= \arg\max_{\mathcal{T}} \prod_i P\left(F_i|\mathcal{T}\right) \prod_{T_k \in \mathcal{T}} P\left(T_k\right) \quad (9)$$

$$\text{s.t } T_k \cap T_l = \varnothing, \forall k \neq l. \quad (10)$$

Equation (9) assumes that the likelihood probabilities $P\left(F_i|\mathcal{T}\right)$ are conditionally independent given T and that the trajectory hypotheses $\{T_k\}$ are also independent of each other. Constraint (10) states the *non-overlapping constraint* that a tracklet belongs to at most one trajectory. We further assume that all the tracklets are reliable and set $P\left(F_i|\mathcal{T}\right) = 1$. The prior $P\left(T_k\right)$ in Equation (9) is modeled as a Markov chain:

$$\begin{aligned} P\left(T_k\right) &= P\left(\{F_{k_1}, F_{k_2}, ..., F_{k_l}\}\right) \\ &= P_s\left(F_{k_1}\right) \prod_{n=1}^{l-1} P\left(F_{k_{n+1}}|F_{k_n}\right) P_e\left(F_{k_l}\right), \end{aligned}$$
$$(11)$$

where P_s is the initialization probability, P_e the termination probability, and $P\left(F_{k_{n+1}}|F_{k_n}\right)$ the transition probability.

Maximizing Equation (8) to find the optimal $\mathcal{T}^*$ is equivalent to minimizing the cost of flow from source s to sink t in a cost-flow network. Following [7, 8], we can obtain:

$$c_i^s = -\log(P_s(F_i)), \ c_i^e = -\log(P_t(F_i)), \quad (12)$$

$$c_{ij} = -\log(P\left(F_j|F_i\right)), \quad (13)$$

where c_i^s is the cost from s to a node, c_i^e the cost from a node to t, and c_{ij} the cost from nodes i to j. For any two tracklets F_i and F_j,

the cost c_{ij} is defined based on the time, spatial, appearance, size, and motion information:

$$\begin{aligned} c_{ij} = &-\log(P_t\left(F_i, F_j\right) \cdot P_s\left(F_i, F_j\right) \cdot P_a\left(F_i, F_j\right) \\ &\cdot P_{size}\left(F_i, F_j\right) \cdot P_m\left(F_i, F_j\right)), \end{aligned}$$
$$(14)$$

where $P_t\left(\cdot\right)$, $P_s\left(\cdot\right)$, $P_a\left(\cdot\right)$, $P_{size}\left(\cdot\right)$, and $P_m\left(\cdot\right)$ denote the time, spatial, appearance, size, and motion affinities, respectively.

The affinities of the appearance and size are defined over the intensity and radius of the object, respectively. Assuming that F_i appears before F_j, the affinity based on time is defined by

$$P_t\left(F_i, F_j\right) = \begin{cases} 1 & \text{if } F_i \cap F_j = \varnothing \text{ and } 0 < |t_i^e - t_j^s| \leq t_{\max} \\ \infty & \text{otherwise} \end{cases}, $$
$$(15)$$

where $t_{\max}$ is the maximum time gap for linking two tracklets. The affinity based on the spatial positions is given by

$$P_s\left(F_i, F_j\right) = \begin{cases} 1 & \text{if } ||p_i^{t_i^e} - p_j^{t_j^s}||/ \triangle t \leq d_{\max} \\ \infty & \text{otherwise} \end{cases}, \quad (16)$$

where $\triangle t$ is the time gap between the end time of F_j and the start time of F_i, and $d_{\max}$ is the maximum moving distance of an object between two consecutive times. The motion affinity is defined based on the prediction errors of the spatial positions and the motion directions between two tracklets. Let $v_i^{t_i^e}$ be the estimated motion vector at the end time of tracklet F_i and $v_j^{t_j^s}$ be the estimated motion vector at the start time of tracklet F_j. We define the prediction error as $\triangle p = p_i^{t_i^e} + \mathrm{sign}(< v_i^{t_i^e}, v_j^{t_j^s} >)v_i^{t_i^e} \triangle t - p_j^{t_j^s}$, where $< \cdot, \cdot >$ denotes the dot product operator. Because a GOP may move in either anterograde or retrograde direction at a given time, we use the sign of the dot product of two motion vectors as an indicator of whether a GOP moves in a consistent direction. Taking into account the motion smoothness, we define the motion affinity as:

$$P_m(F_i, F_j) = \alpha(< \frac{v_i^{t_i^e}}{||v_i^{t_i^e}||}, \frac{v_j^{t_j^s}}{||v_j^{t_j^s}||} >) + (1 - \alpha)G\left(\triangle p, \sigma_p\right),$$
$$(17)$$

where α is a weighting parameter, and $G\left(\cdot, \sigma\right)$ is a zero-mean Gaussian function.
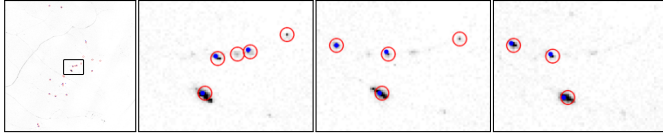
3.3. Optimization

We build a cost-flow network graph in which the nodes represent tracklets and the edges indicate association between the tracklets. To obtain the tracks, we run the successive shortest path algorithm.

4. EXPERIMENTAL RESULTS

4.1. Data

We used two confocal microscopy data sets of da neurons in the Drosophila PNS for evaluation. Sequence 1 was a 3D+t data set, containing 31 volumes, each of which was $512 \times 512 \times 5$ voxels, with a resolution of $0.35 \times 0.35 \times 1$ µm³/voxel. Sequence 2 was a 2D+t data set, including 100 images, each of which was 610×610 pixels, with a resolution of 0.167×0.167 µm²/pixel.

An experienced biologist manually tracked the GOPs on some selected regions to construct the reference tracks. Our data sets were acquired at a frequency of 1 Hz, and the GOPs usually moved very fast, leading to ambiguities in determining the GOP movement. Therefore, the reference tracks included only the ones annotated with a high level of confidence.

(a) res. at time 1 (b) close-up at time 1 (c) close-up at time 4 (d) close-up at time 16

Fig. 3: Selected detection results and manual annotations. (a) Detection results at time 1. (b)-(d) Close-ups of the rectangular area in (a) at times 1, 4, and 16, respectively. Red circles are the detection results and blue dots indicate the manual annotations. Image intensities are inverted. Best viewed in color.

Table 1: Comparison of the detection results by using our proposed multi-scale blobness and intensity in representation of GOP appearances. M, N, CD, MD and FA denote the numbers of manual annotations, detected objects, correct detections, missed detections and false alarms. Precision (prec) and recall (rec) are also given.

Data Term	M	N	CD	MD	FA	prec	rec
Intensity	523	587	416	107	171	0.7087	0.7954
Ours	523	598	435	88	163	0.7274	0.8317

4.2. Detection Results

Figure 3 shows the close-ups of the detection results (red circles), together with manual annotations (blue dots), on selected images of sequence 1. Our method successfully detected the GOPs that were annotated by the human, but produced some false positives due to other spot-like particles whose appearances were similar to those of the GOPs. Shown in Table 1 is the comparison of detection results obtained by using different features in computation of GOP appearances (Equation (4)) in the MPP framework: one used our proposed multi-scale blobness and the other the raw image intensity. Our model produced more correct detections, fewer missed detections, and fewer false alarms, performing better than the one using intensity for GOP appearances.

4.3. Linking Results

Figure 4 shows the obtained trajectories of GOPs of sequences 1 and 2. For quantitative evaluation, we used the evaluation metric described in [9] to compute the tracking accuracy P_{track} defined as $P_{\text{track}} = N_{\text{track,correct}}/N_{\text{track,total}}$, where $N_{\text{track,correct}}$ is the number of correctly computed tracks and $N_{\text{track,total}}$ is the number of reference tracks. This measurement penalized the computed tracks that were broken. The higher the P_{track} is, the better the performance is. $P_{\text{track}} = 1$ is the best. We compared our method with (1) the particle tracker plugin [2] in Fiji [10], (2) the multiple-hypothesis tracking (MHT) approach [4] implemented in the Icy software [11], and (3) a naive linking approach that used single detection responses as nodes in the cost-flow network graph and set the edge cost based on the positions of detection responses. Table 2 summarizes the performance of different linking strategies on the two image sequences. Our proposed method yielded better results than the others, suggesting that it is effective for tracking the GOPs.

Table 2: Performance of different linking methods measured by the tracking accuracy (P_{track}).

Method	Particle Tracker	MHT	Baseline	Ours
Sequence 1	0.6467	0.8542	0.7527	0.8734
Sequence 2	0.6092	0.7019	0.9100	0.9474

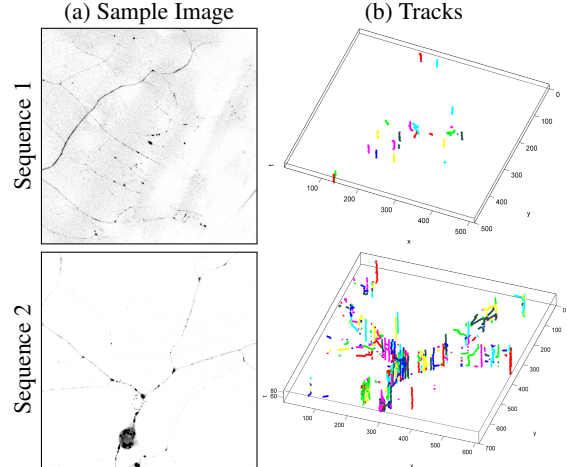(a) Sample Image (b) Tracks

Sequence 1

Sequence 2

Fig. 4: Tracking results of sequences 1 and 2. (a) Samples images. Image intensities are inverted and contrast is enhanced for visualization. (b) Tracking results. Each color represents a different track. Best viewed in color.

5. CONCLUSIONS AND FUTURE WORK

We presented a data association based framework for obtaining the motion patterns of dendritic GOPs. Our method used multi-scale blobness for GOP appearances, which reduced the number of missed detections in the detection stage. In the linking stage, the use of high-level information in the affinity model increased the association accuracy. The design of the proposed approach is versatile and can be adapted to various applications. We plan on obtaining statistical information from the GOP tracks for further analysis and simulating the dendritic growth based on the extracted motion patterns.

6. REFERENCES

[1] K. M. Ori-McKenney, L. Y. Jan, and Y.-N. Jan, "Golgi outposts shape dendrite morphology by functioning as sites of acentrosomal microtubule nucleation in neurons," *Neuron*, vol. 76, no. 5, pp. 921–930, 2012.

[2] I. F. Sbalzarini and P. Koumoutsakos, "Feature point tracking and trajectory analysis for video imaging in cell biology," *J. Structural Biology*, vol. 151, no. 2, pp. 182–195, 2005.

[3] K. Jaqaman, D. Loerke, M. Mettlen, H. Kuwata, S. Grinstein, S. L. Schmid, and G. Danuser, "Robust single particle tracking in live cell time-lapse sequences," *Nature methods*, vol. 5, no. 8, pp. 695–702, 2008.

[4] N. Chenouard, I. Bloch, and J.-C. Olivo-Marin, "Multiple hypothesis tracking for cluttered biological image sequences," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 35, no. 11, pp. 2736–2750, 2013.

[5] A. Gamal-Eldin, X. Descombes, G. Charpiat, and J. Zerubia, "Multiple birth and cut algorithm for multiple object detection," *J. Multimedia Processing and Technologies*, vol. 1, no. 4, pp. 260–276, 2010.

[6] C. Huang, Y. Li, and R. Nevatia, "Multiple target tracking by learning-based hierarchical association of detection responses," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 35, no. 4, pp. 898–910, 2013.

[7] B. Wang, G. Wang, K. L. Chan, and L. Wang, "Tracklet association with online target-specific metric learning," in *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, 2014, pp. 1234–1241.

[8] L. Zhang, Y. Li, and R. Nevatia, "Global data association for multi-object tracking using network flows," in *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, 2008, pp. 1–8.

[9] W. J. Godinez, M. Lampe, S. Wörz, B. Müller, R. Eils, and K. Rohr, "Deterministic and probabilistic approaches for tracking virus particles in time-lapse fluorescence microscopy image sequences," *Medical Image Analysis*, vol. 13, no. 2, pp. 325–342, 2009.

[10] J. Schindelin, I. Arganda-Carreras, E. Frise, V. Kaynig, M. Longair, T. Pietzsch, et al., "Fiji: an open-source platform for biological-image analysis," *Nature Methods*, vol. 9, no. 7, pp. 676–682, 2012.

[11] F. de Chaumont, S. Dallongeville, N. Chenouard, N. Hervé, S. Pop, T. Provoost, et al., "Icy: an open bioimage informatics platform for extended reproducible research," *Nature Methods*, vol. 9, no. 7, pp. 609–696, 2012.