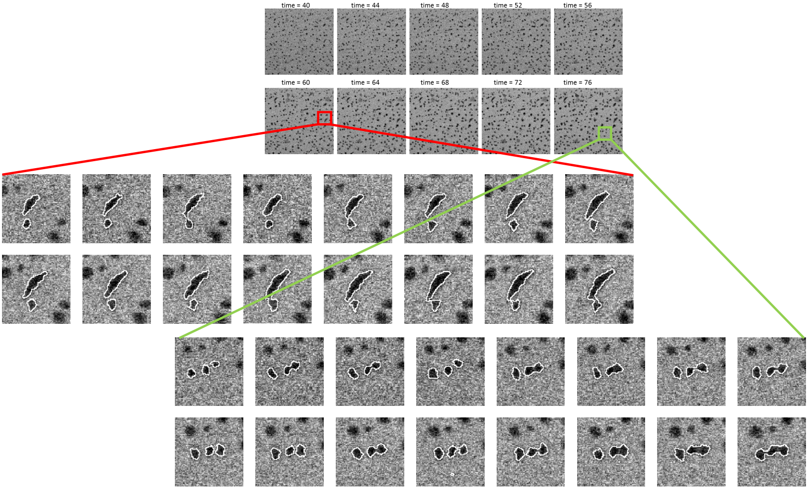
Analytics for High Frame-rate Image Streaming

Chiwoo Park, Florida State University

August 2017 at New York Scientific Data Summit

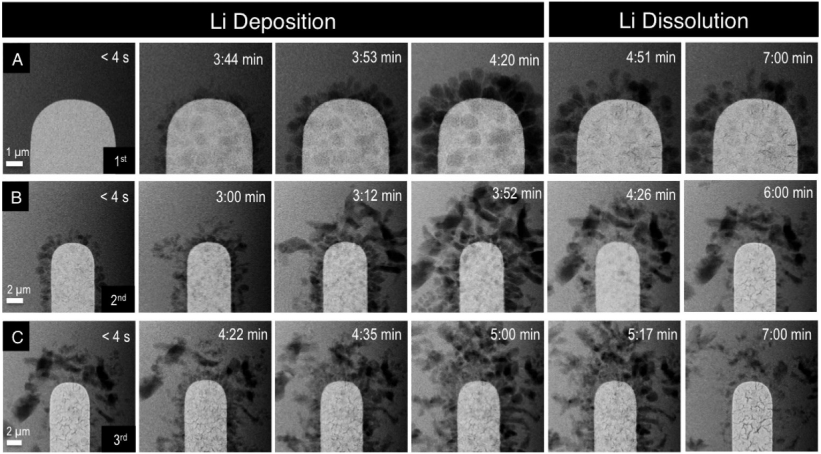
Many scientific studies have been relying on a stream of imagery observations.

- Ex 1. *in situ* microscopy of nanoparticle self-assembly



Many scientific studies have been relying on a stream of imagery observations.

- Ex 2. *Operando electrochemical* STEM of Li-ion battery



Many scientific studies have been relying on a stream of imagery observations.

- Ex 3. Remote sensing imagery (photo credit: Planet lab)

July 2 2017

July 4 2017

July 5 2017

Dec 20 2016

Jan 26 2017

Feb 22 2017

Amount of images collected each time is huge.

- *in situ* microscopy: small-scale changes occur in a short time scale. Capturing such fast changes would need high frame-rate measurements.

$$\begin{aligned}\text{Data rate} &= \text{16MBs per image} \times \text{1000 images per sec.}\\ &= \text{16GBs per sec}\end{aligned}$$

- Planet lab's constellation of 88 satellites: each collects over 2 million km^2 per day with a resolution of 3-5 meters.

$$\begin{aligned}\text{Data rate} &= 88 \times (2 \times 10^6 km^2 \text{ per day } * 40,000 \text{ pixels per } km^2)\\ &\approx \text{20 million GBs per day.}\\ &\approx \text{230 GBs per sec.}\end{aligned}$$

In situ analysis is typically preferred for high data rates.

- limited network bandwidth
  e.g. local disk writing $\approx$ 100 to 600 MBs per sec.
  e.g. satellite to ground station $\approx$ 200 MBs per sec.

- time-to-analysis requirement
  It takes too much time for data transfer, storage and batch processing.

  In situ analysis enables realtime or near realtime analysis of data.

Today we present an approach for *in situ* analysis of high frame-rate imagery observations.

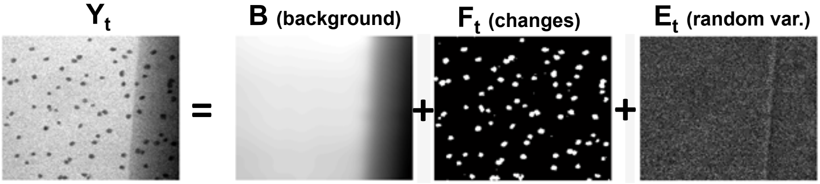
The new approach aims for near real-time analysis of

- Detect Changes: locate visual changes, e.g. appearance or disappearance of objects, morphology changes, color changes, texture changes,...
- Track Changes: associate visual changes obtained at various instances to form a track
- Find Longitudinal Patterns: find long-time range patterns in tracked changes.

Robust Change Detection

Formulating change detection

Let $\boldsymbol{Y}_t$ denote a $m \times n$ matrix representing an input image obtained at time t. The input matrix can be decomposed into three component matrices of the same size,



We want to estimate $\boldsymbol{F}_t$.

When $\boldsymbol{B}$ is assumed unchanged,

The likelihood maximization for $\boldsymbol{F}_t$ can be pursued.

$$\text{Minimize}_{\boldsymbol{B},\{\boldsymbol{F}_t\}} \quad \sum_{t=1}^{T} ||\boldsymbol{E}_t||_F^2$$

$$\boldsymbol{Y}_t = \boldsymbol{B} + \boldsymbol{F}_t + \boldsymbol{E}_t, t = 1, ..., T.$$

- $||\boldsymbol{F}_t||_1 \leq \mu$: $\boldsymbol{F}_t$ is sparse.

This is a batch processing to fit $\boldsymbol{F}_t$ all together. Since more data are used, this provides more robust estimates when $\boldsymbol{B}$ does not change in time, since less data are used.

When local changes of background is expected,

Local weights ω_t can be posed for local likelihood maximization. For each time t',

$$\text{Minimize}_{\boldsymbol{B},\boldsymbol{F}_{t'}} \quad \sum_{t=t'-\delta}^{t'+\delta} \omega_t ||\boldsymbol{E}_t||_F^2$$

$$\boldsymbol{Y}_t = \boldsymbol{B} + \boldsymbol{F}_t + \boldsymbol{E}_t, t = t' - \delta, ..., t + \delta.$$

- $\boldsymbol{B}$ might change in time.
- $||\boldsymbol{F}_t||_1 \leq \mu$.

δ dilemma: timeliness vs. robustness of estimation

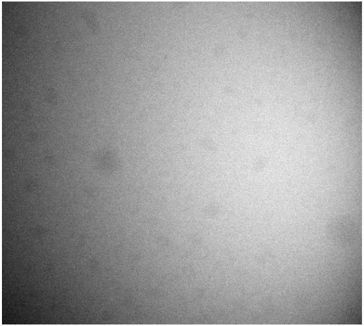
- $\delta = T$: batch processing, more robust
- $0 < \delta < T$: grouped processing, less robust
- $\delta = 0$: frame-by-frame processing, least robust

Can we maintain robustness of estimation for $\delta = 0$?

Degradation of robustness with small δ can be made up using prior knowledge on $\boldsymbol{B}$ in the form of a cost function, $\mathbb{J}(\boldsymbol{B})$.

- Example: Background is very simple and smooth for many microscope images. $\mathbb{J}(\boldsymbol{B})$ can be a smoothness measure.

(a) original



(b) background estimated

Can we maintain robustness for $\delta = 0$?

Use that prior knowledge on $\boldsymbol{B}$ to improve robustness of estimation

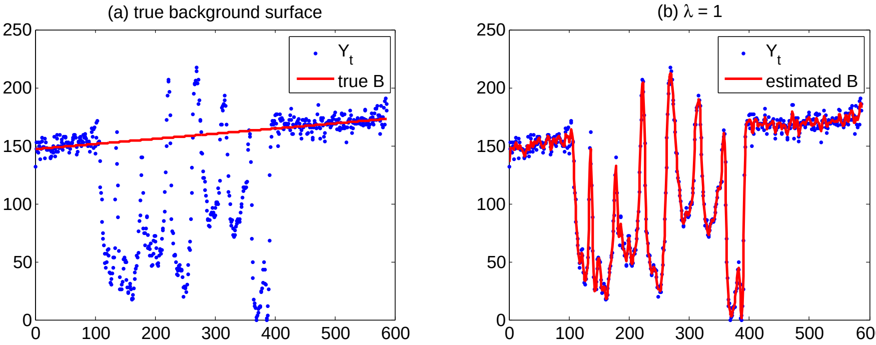
First trial: For each time t, optimize the regularized local likelihood

$$\begin{aligned} \text{Minimize}_{\boldsymbol{B},\boldsymbol{F}_t} \quad & ||\boldsymbol{E}_t||_F^2 + \lambda \mathbb{J}(\boldsymbol{B}) \\ & \boldsymbol{Y}_t = \boldsymbol{B} + \boldsymbol{F}_t + \boldsymbol{E}_t. \\ & ||\boldsymbol{F}_t||_1 \leq \mu. \end{aligned}$$

- $\boldsymbol{B}$ may be better guided by the prior cost function.

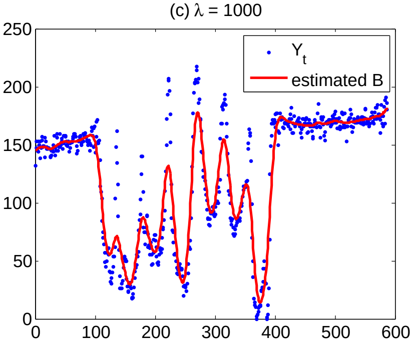
The trial gave a poor estimate.

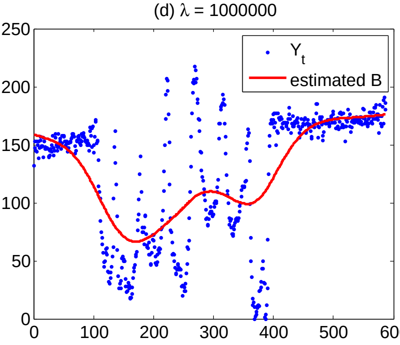
- $\delta = 0$ case is not robust yet. The estimation of $\boldsymbol{B}$ is quite affected by $\boldsymbol{F}_t$ and $\boldsymbol{E}_t$.



The trial gave a poor estimate.

- The increase of the weight on the prior cost (i.e. λ) can cause significant biases.





We borrow the concept of robust regression in statistics to increase the robustness.

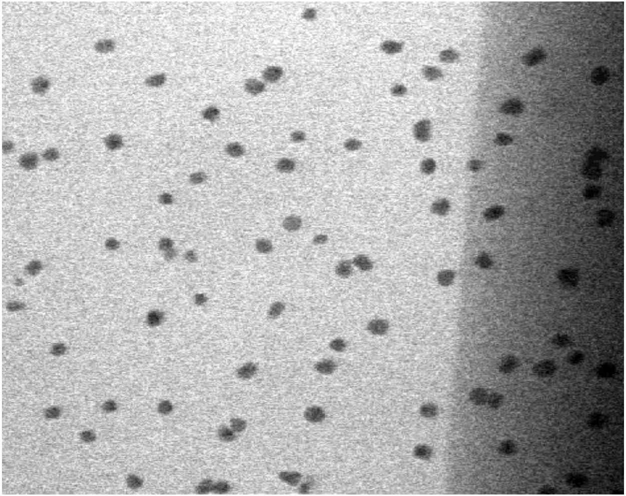
In statistics, the tendency of the square loss $||\boldsymbol{E}_t||_F^2$ being dominated by outliers (such as sudden changes) was discussed and addressed by changing it with the robust loss function, e.g. the Huber loss, $\mathbb{L}_H$,

$$
\begin{aligned}
\text{Minimize} \quad & \mathbb{L}_H(\boldsymbol{E}_t) + \lambda \mathbb{J}(\boldsymbol{B}) \\
& \boldsymbol{Y} = \boldsymbol{B} + \boldsymbol{F}_t + \boldsymbol{E}_t \\
& ||\boldsymbol{F}_t||_1 \leq \mu.
\end{aligned}
$$

The estimated $\boldsymbol{B}$ is less sensitive to the choice of $\boldsymbol{F}_t$ and $\boldsymbol{E}_t$. The solution approach of the estimation can be found at our paper Vo and Park (2016).

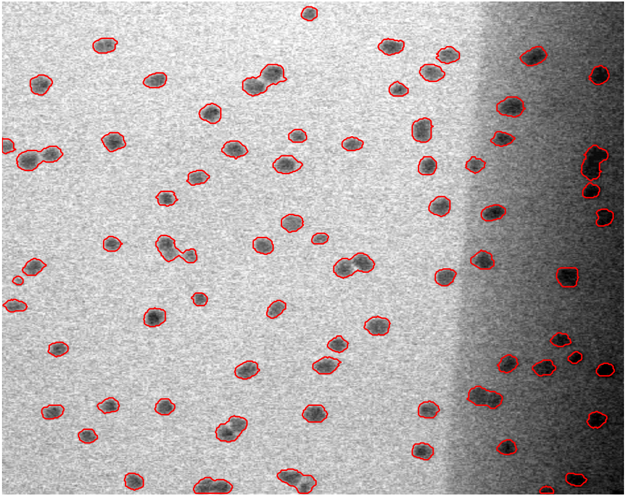
High Contrast Example (Gold NP)

(a) original

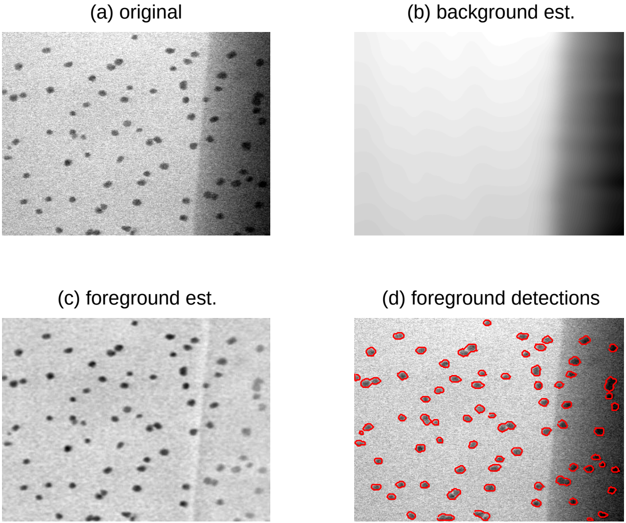


High Contrast Example (Gold NP): Output

(d) foreground detections

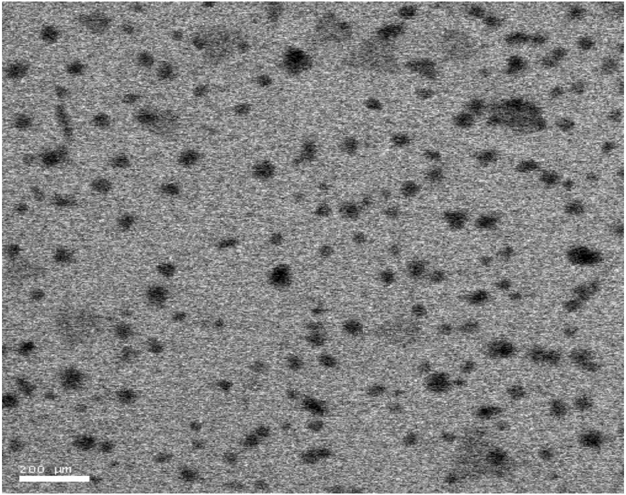


High Contrast Example (Gold NP): Output

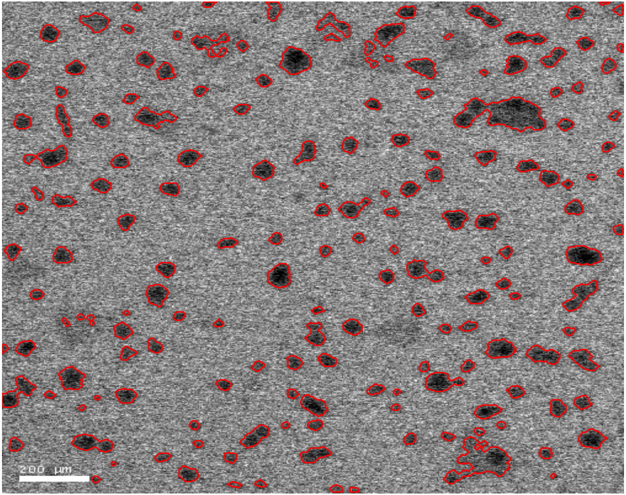


High Contrast Example (Silver NP)

(a) original

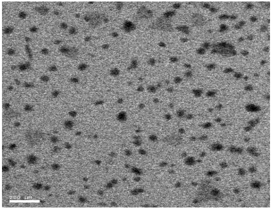


High Contrast Example (Silver NP): Output



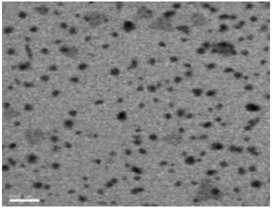
High Contrast Example (Silver NP): Output

(a) original

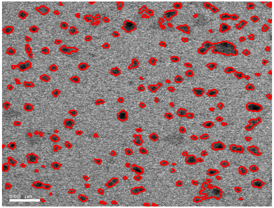


(b) background est

(c) foreground est

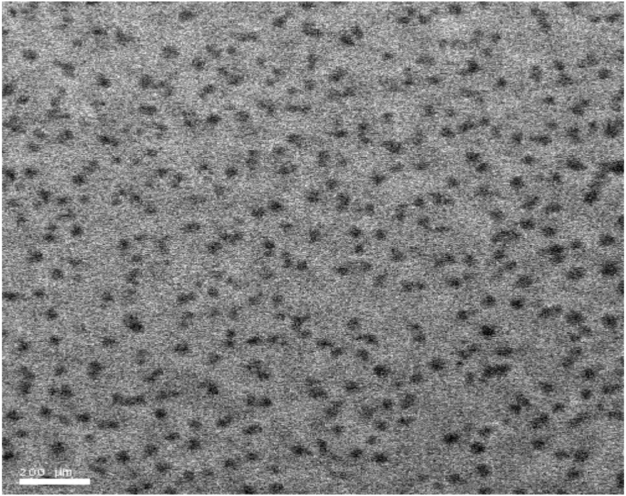


(d) particle detection



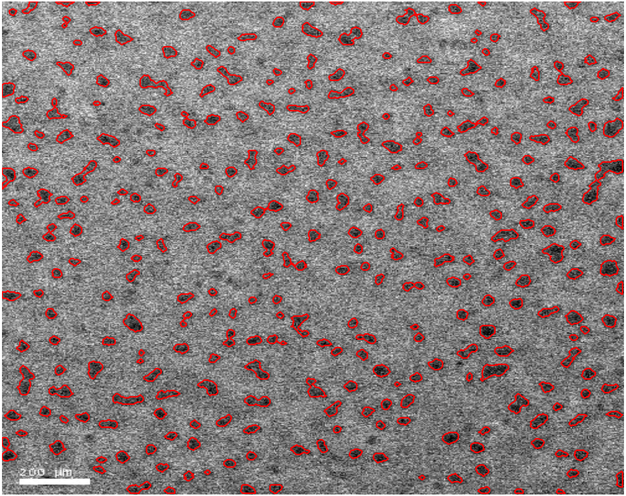
Medium Contrast Example (Silver NP)

(a) original



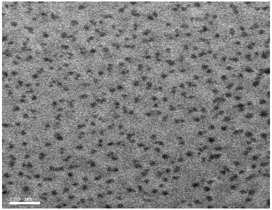
Medium Contrast Example (Silver NP): Output

(d) foreground detections



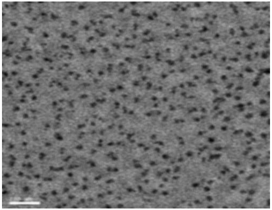
Medium Contrast Example (Silver NP): Output

(a) original

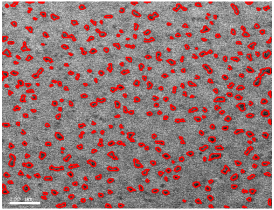


(b) background est.

(c) foreground est.

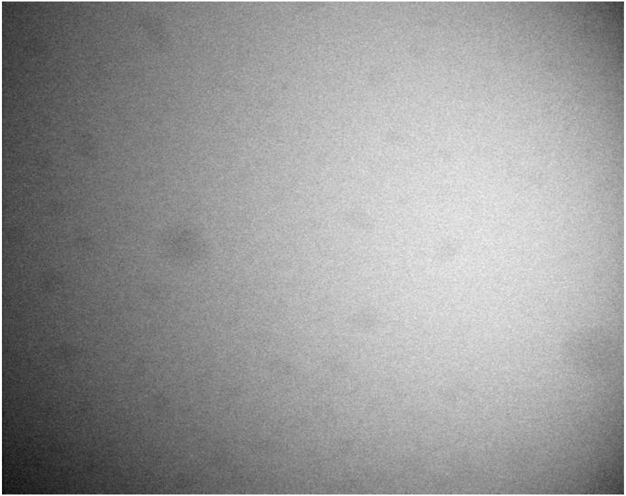


(d) foreground detections



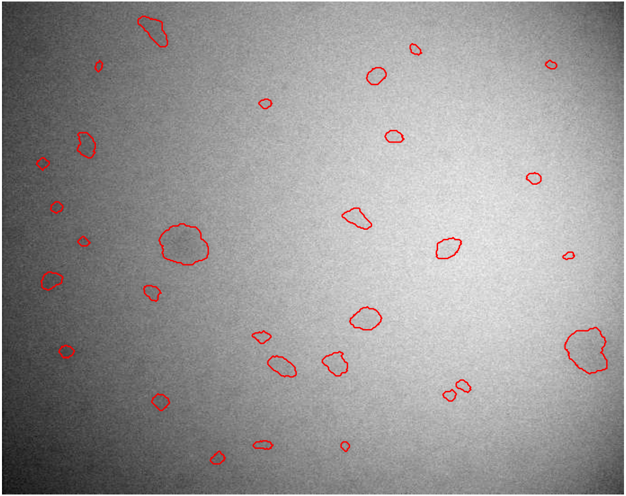
Low Contrast Example (Protein)

(a) original

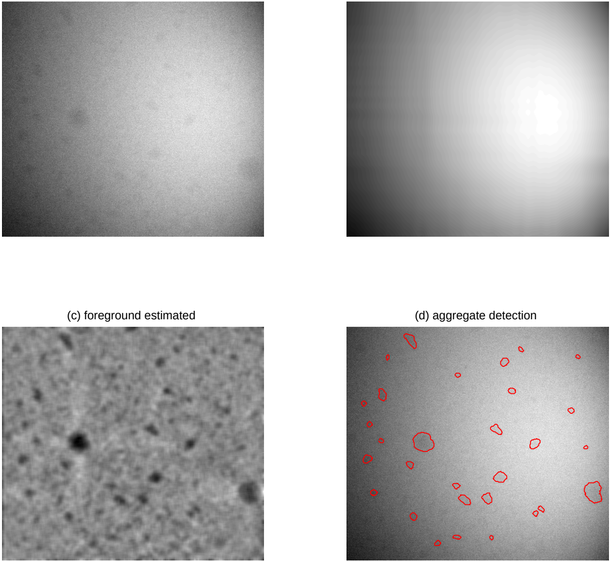


Low Contrast Example (Protein): Output

(d) aggregate detection



Low Contrast Example (Protein): Output



(a) original

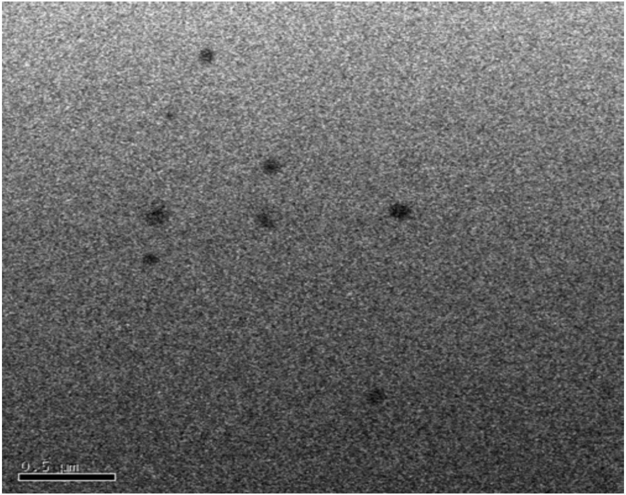
(b) background estimated

(c) foreground estimated

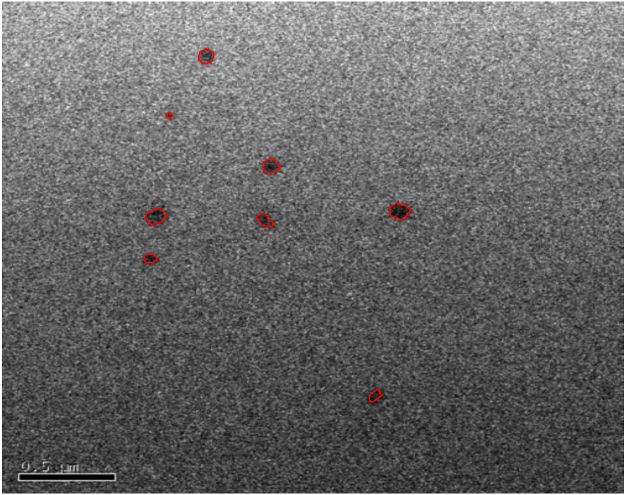
(d) aggregate detection

Low Contrast Example (Micelle)

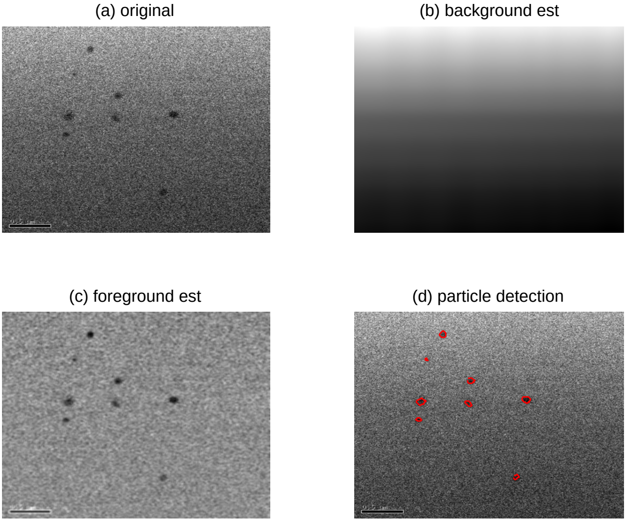
(a) original



Low Contrast Example (Micelle): Output

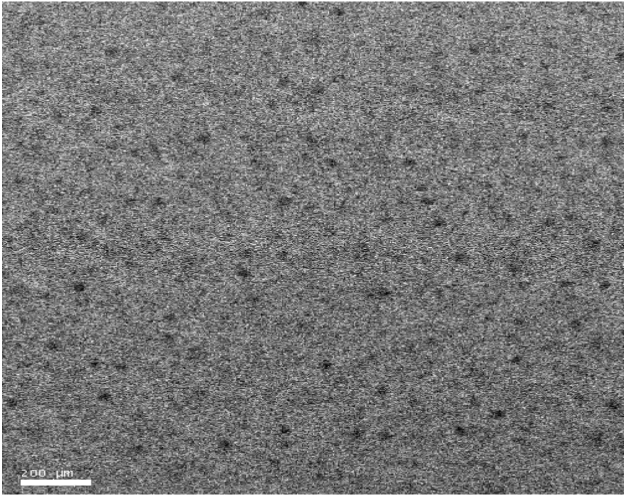


Low Contrast Example (Micelle): Output



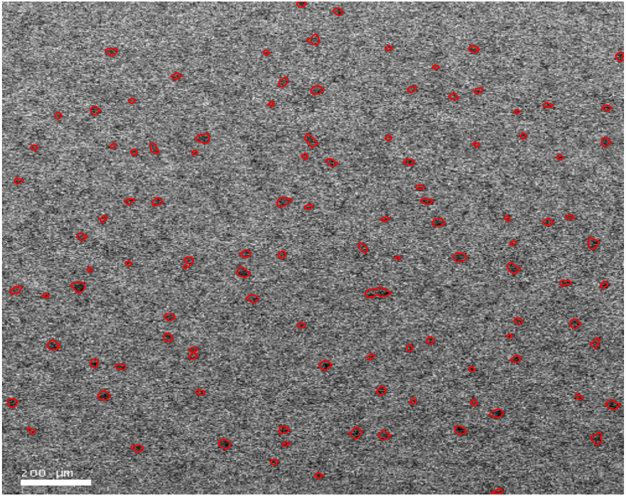
Low Contrast Example (NP)

(a) original



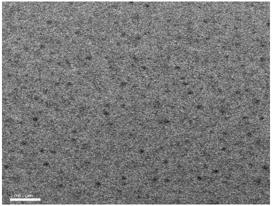
Low Contrast Example (NP): Output

(d) foreground detections



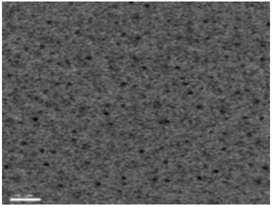
Low Contrast Example (NP): Output

(a) original

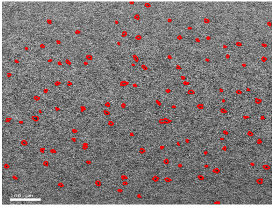


(b) background est.

(c) foreground est.



(d) foreground detections



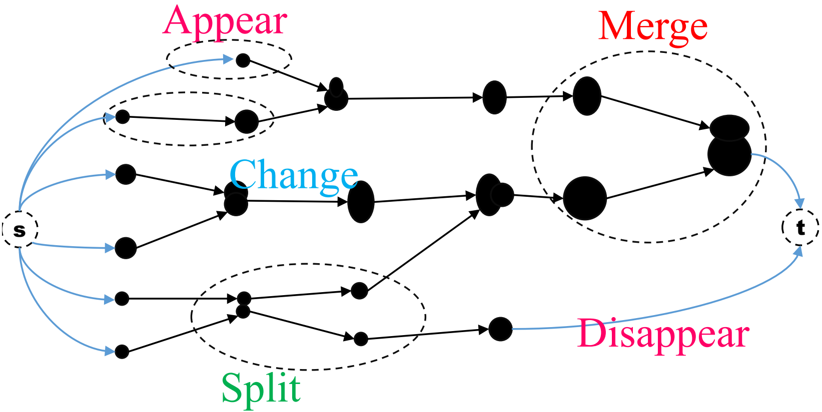
Some Other Examples

(a) input image (b) estimated background (c) estimated foreground

Track Changes

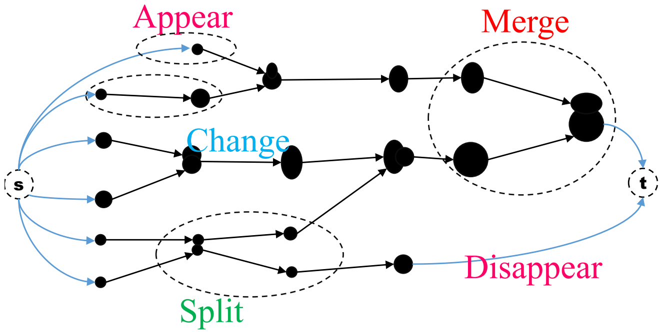
Associate visual changes obtained at various instances to form a track.

The association is represented as a digraph $G = (V, E)$, where $v \in V$ is a node representing a visual change, and $e \in E$ is an edge.



An association $a \in A$ is not only an edge but also a collection of edges,

1. Change by 1-to-1 association $e \in E$
2. Merge by m-to-1 associations $\{e \in E : sink(e) = v\}$
3. Split by 1-to-n associations $\{e \in E : source(e) = v\}$
4. Appear by an edge from a source node.
5. Disappear by an edge from a sink node.



Data association problem is a problem of finding G that minimizes the total association cost

$$\begin{aligned} \text{Minimize} \quad & \sum_{a \in A} c_a \cdot z_a \\ & z_a \in \{0, 1\} \\ & \{z_a; a \in A\} \in \mathcal{C}. \end{aligned}$$

- $z_a \in \{0, 1\}$ represents the activation of $a \in A$.
- c_a is the cost of the activation.

Only 1-to-1 associations were considered in literature. A few exceptions are

- Jaqaman et al. (2008) and Henriques et al. (2011) studied some linear optimization models to consider one-to-two or two-to-one associations.
- Khan et al. (2005a,b); Kreucher et al. (2005); Ng et al. (2007) studied some sequential Monte Carlo approaches. As the number of foreground objects increases, the state space becomes high dimensional, so the approaches are not scaling very well.

We formulate and solve a general data association problem

Model Assumptions

M-way association: The number of foreground objects involved in an association is at least 1 and at most M.

Imperfect detection: A foreground detection algorithm is not perfect. Some $v \in V$ can be faulty detections.

Binary integer programming problem can be formulated and solved.

The objective is to minimize the total cost of associations

$$\text{Min} \sum_{a \in A_{1,1}} z_a c_a + \sum_{a \in A_{m,1}} z_a c_a + \sum_{a \in A_{1,n}} z_a c_a + \sum_{a \in A_{m,n}} z_a c_a$$

subject to

- In-Degree Constraint for node v: $1 \leq \sum_{source(e)=v} z_e \leq M$
- Out-Degree Constraint for node v: $1 \leq \sum_{sink(e)=v} z_e \leq M$
- Relationship between z_e and z_a:

$$z_a \leq z_{e'} \text{ for } e' \in a$$

$$\sum_{e' \in a} (z_{e'} - 1) + 1 \leq z_a.$$

Using vector notations,

$$
\begin{aligned}
\text{Minimize} \quad & \boldsymbol{c}_1^T \boldsymbol{z}_1 + \sum_m \boldsymbol{c}_{m1}^T \boldsymbol{z}_{m1} + \sum_n \boldsymbol{c}_{n2}^T \boldsymbol{z}_{n2} + \sum_{m,n} \boldsymbol{d}_{mn}^T \boldsymbol{y}_{mn} \\
& \boldsymbol{A}_1 \boldsymbol{z}_1 \geq \boldsymbol{b}_1 && \text{(1a)} \\
& \boldsymbol{A}_{m1} \boldsymbol{z}_1 + \boldsymbol{B}_{m1} \boldsymbol{z}_{m1} \geq \boldsymbol{b}_{m1} && \text{(1b)} \\
& \boldsymbol{A}_{n2} \boldsymbol{z}_1 + \boldsymbol{C}_{n1} \boldsymbol{z}_{n2} \geq \boldsymbol{b}_{n2} && \text{(1c)} \\
& \boldsymbol{P}_{mn} \boldsymbol{z}_{m1} + \boldsymbol{Q}_{mn} \boldsymbol{z}_{n2} + \boldsymbol{y}_{mn} \geq \boldsymbol{1} && \text{(1d)} \\
& \boldsymbol{P}_{mn} \boldsymbol{z}_{m1} - \boldsymbol{y}_{mn} \geq \boldsymbol{0} && \text{(1e)} \\
& \boldsymbol{Q}_{mn} \boldsymbol{z}_{n2} - \boldsymbol{y}_{mn} \geq \boldsymbol{0} && \text{(1f)}
\end{aligned}
$$

$$
\boldsymbol{z}_1 \in B^{p_1}, \boldsymbol{z}_{m1} \in B^{p_{m1}}, \boldsymbol{z}_{n2} \in B^{p_{n2}}, \boldsymbol{y}_{mn} \in B^{q_{mn}}
$$

Batch Solution: We solve the Lagrange dual relaxation of the BIP.

Solving the binary optimization problem is NP-hard! We used the special structure of the problem to find an integer-valued suboptimal.

$$
\begin{array}{llllll}
\text{Minimize} & \mathbf{c}_1^T \mathbf{z}_1 & + \sum_m \mathbf{c}_{m1}^T \mathbf{z}_{m1} & + \sum_n \mathbf{c}_{n2}^T \mathbf{z}_{n2} & + \sum_{m,n} \mathbf{d}_{mn}^T \mathbf{y}_{mn} & \\
& \boxed{\mathbf{A}_1 \mathbf{z}_1} & & & & \geq \mathbf{b}_1 \quad (1a) \\
& \mathbf{A}_{m1} \mathbf{z}_1 & + \mathbf{B}_{m1} \mathbf{z}_{m1} & & & \geq \mathbf{b}_{m1} \quad (1b) \\
& \mathbf{A}_{n2} \mathbf{z}_1 & & + \mathbf{C}_{n1} \mathbf{z}_{n2} & & \geq \mathbf{b}_{n2} \quad (1c) \\
& & \mathbf{P}_{mn} \mathbf{z}_{m1} & + \mathbf{Q}_{mn} \mathbf{z}_{n2} & + \mathbf{y}_{mn} & \geq \mathbf{1} \quad (1d) \\
& & \mathbf{P}_{mn} \mathbf{z}_{m1} & & - \mathbf{y}_{mn} & \geq \mathbf{0} \quad (1e) \\
& & & \mathbf{Q}_{mn} \mathbf{z}_{n2} & - \mathbf{y}_{mn} & \geq \mathbf{0} \quad (1f)
\end{array}
$$

Totally Unimodular!

$$
\mathbf{z}_1 \in B^{p_1}, \mathbf{z}_{m1} \in B^{p_{m1}}, \mathbf{z}_{n2} \in B^{p_{n2}}, \mathbf{y}_{mn} \in B^{q_{mn}}
$$

Batch Solution: We solve the Lagrange dual relaxation of the BIP.

Repeat (SP) and (MP) until convergence.

(SP) Solve for $\boldsymbol{z}_1, \boldsymbol{z}_{m1}, \boldsymbol{z}_{n2}, \boldsymbol{y}_{mn}$ with fixed Lagrange multipliers.

$$
\begin{aligned}
\text{Min}\quad & \boldsymbol{c}_1^T \boldsymbol{z}_1 + \sum_m \boldsymbol{c}_{m1}^T \boldsymbol{z}_{m1} + \sum_n \boldsymbol{c}_{n2}^T \boldsymbol{z}_{n2} + \sum_{m,n} \boldsymbol{d}_{mn}^T \boldsymbol{y}_{mn} \\
& + \textstyle\sum_m \boldsymbol{\lambda}_{m1}^T (\boldsymbol{b}_{m1} - \boldsymbol{A}_{m1}\boldsymbol{z}_1 - \boldsymbol{B}_{m1}\boldsymbol{z}_{m1}) \\
& + \textstyle\sum_n \boldsymbol{\lambda}_{n2}^T (\boldsymbol{b}_{n2} - \boldsymbol{A}_{n2}\boldsymbol{z}_1 - \boldsymbol{B}_{n2}\boldsymbol{z}_{n2}) \\
& \boldsymbol{A}_1 \boldsymbol{z}_1 \geq \boldsymbol{b}_1 \\
& \boldsymbol{P}_{mn}\boldsymbol{z}_{m1} + \boldsymbol{Q}_{mn}\boldsymbol{z}_{n2} + \boldsymbol{y}_{mn} \geq \boldsymbol{1} \\
& \boldsymbol{P}_{mn}\boldsymbol{z}_{m1} - \boldsymbol{y}_{mn} \geq \boldsymbol{0} \\
& \boldsymbol{Q}_{mn}\boldsymbol{z}_{n2} - \boldsymbol{y}_{mn} \geq \boldsymbol{0} \\
& \boldsymbol{0} \leq \boldsymbol{z}_1, \boldsymbol{z}_{m1}, \boldsymbol{z}_{n2}, \boldsymbol{y}_{mn} \leq \boldsymbol{1}
\end{aligned}
$$

(MP) Improve the Lagrange multipliers $\boldsymbol{\lambda}_{m1}, \boldsymbol{\lambda}_{n2} \geq \boldsymbol{0}$:

$$
\text{Max}\quad \sum_m \boldsymbol{\lambda}_{m1}^T (\boldsymbol{b}_{m1} - \boldsymbol{A}_{m1}\boldsymbol{z}_1^* - \boldsymbol{B}_{m1}\boldsymbol{z}_{m1}^*) + \sum_n \boldsymbol{\lambda}_{n2}^T (\boldsymbol{b}_{n2} - \boldsymbol{A}_{n2}\boldsymbol{z}_1^* - \boldsymbol{B}_{n2}\boldsymbol{z}_{n2}^*)
$$

Near realtime solution

The previous solution approach associates all image frames in one step.

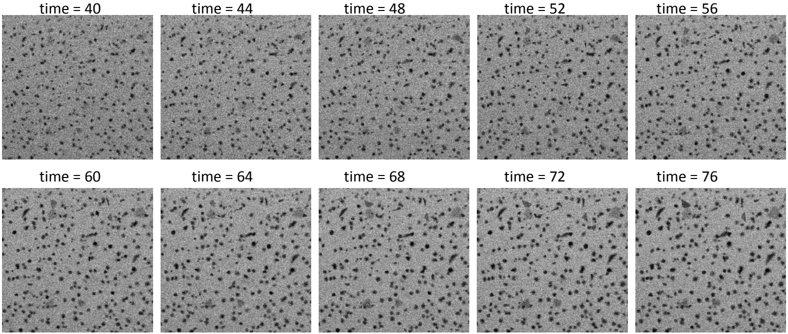
- Pros: It pursues for global optimality.
- Cons: This is a batch processing so far from realtime processing.

Near realtime solution can be sought by solving the BIP in a frame-by-frame fashion.

- Cons: When miss detections or faulty detections occur, the frame-by-frame association incurs significant fragmentations in traces.
- We combined the frame-by-frame data association with delayed data association strategy to fix this issue.

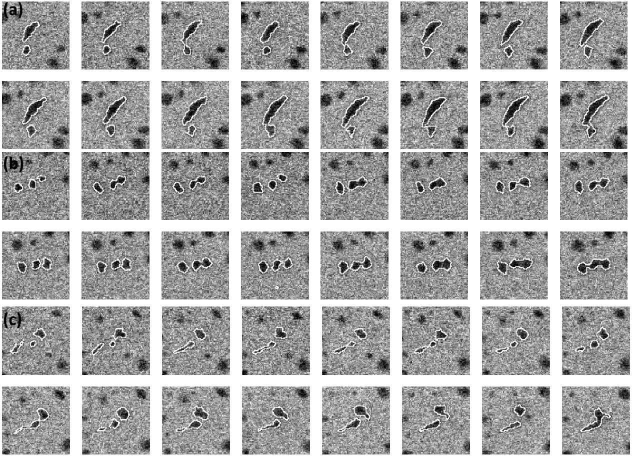
Demonstration (Real Case)

Solution phase silver nanoparticle growth was imaged by *in situ* transmission electron microscopy for 89 seconds.



Demonstration (Real Case)

We applied our method to track particle interactions; Evaluated the accuracy of the data association over the manually inspected 18 trajectories.



Demonstration (Real Case)

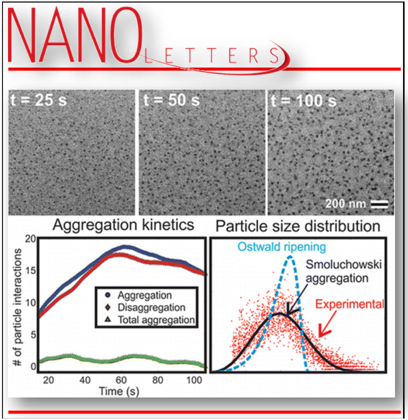
The data association errors were evaluated against the manually inspected 18 trajectories.

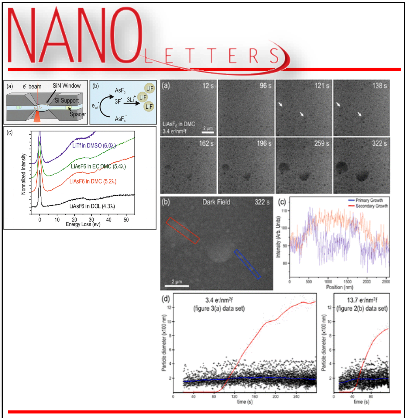
Type	Our method		Henrique		Jaqaman		Yu	
	FN	FP	FN	FP	FN	FP	FN	FP
1-to-1	0.033	0.038	0.086	0.061	0.491	0.261	0.507	0.286
1-to-m	0.020	0.109	0.100	0.167	0.960	0.800	1.000	1.000
n-to-1	0.035	0.098	0.114	0.137	0.895	0.586	0.991	0.909
Faulty	0.000	0.001	0.000	0.001	0.000	0.000	0.000	0.013
Birth	0.000	0.000	0.000	0.750	0.333	0.833	0.333	0.952
Death	0.000	1.000	0.000	1.000	0.000	1.000	0.000	1.000

Table: Real Microscope Data - Data association errors of our method with $M = 3$, Henriques et al. (2011), Jaqaman et al. (2008), and Yu and Medioni (2009).

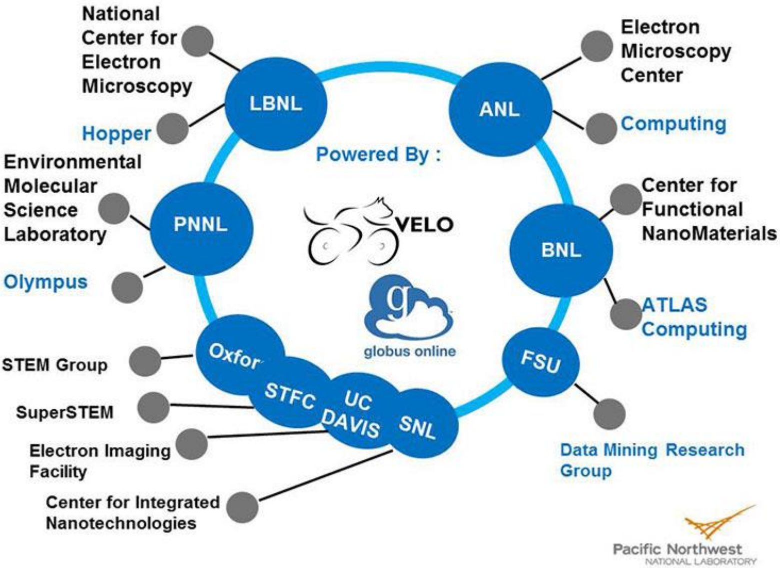
The proposed approach has been successfully applied to support high impact science research.

Examples of Applications





Broad use in microscopy



Closing Remarks

- Be able to analyze very low contrast images at the rate of ten images per second.
- This corresponds to processing rate of 160 MB per second.
- Be able to analyze moderate speed process in real-time.
- Burning a hardware logic for acceleration may help further increase the processing rate.

Thanks for general supports!

Boyd, S., N. Parikh, E. Chu, B. Peleato, and J. Eckstein (2011). Distributed optimization and statistical learning via the alternating direction method of multipliers. *Foundations and Trends® in Machine Learning 3*(1), 1–122.

Chen, Y., H. Tagare, S. Thiruvenkadam, F. Huang, D. Wilson, K. Gopinath, R. Briggs, and E. Geiser (2002). Using prior shapes in geometric active contours in a variational framework. *International Journal of Computer Vision 50*(3), 315–328.

Cheng, J. and J. Rajapakse (2009). Segmentation of clustered nuclei with shape markers and marking function. *IEEE Transactions on Biomedical Engineering 56*(2009), 741–748.

Daněk, O., P. Matula, C. Ortiz-De-Solórzano, A. Muñoz-Barrutia, M. Maška, and M. Kozubek (2009). Segmentation of touching cell nuclei using a two-stage graph cut model. In *Proceedings of the 16th Scandinavian Conference on Image Analysis*, Oslo, Norway, pp. 410–419.

Felzenszwalb, P. and D. Huttenlocher (2004). Efficient graph-based image segmentation. *International Journal of Computer Vision 59*(2), 167–181.

Foulonneau, A., P. Charbonnier, and F. Heitz (2009). Multi-reference shape priors for active contours. *International Journal of Computer Vision 81*(1), 68–81.

Henriques, J., R. Caseiro, and J. Batista (2011). Globally optimal solution to multi-object tracking with merged measurements. In *IEEE International Conference on Computer Vision*, pp. 2470–2477. IEEE.

Jaqaman, K., D. Loerke, M. Mettlen, H. Kuwata, S. Grinstein, S. L. Schmid, and G. Danuser (2008). Robust single-particle tracking in live-cell time-lapse sequences. *Nature methods 5*(8), 695–702.

Khan, Z., T. Balch, and F. Dellaert (2005a). MCMC-based particle filtering for tracking a variable number of interacting targets. *IEEE Transactions on Pattern Analysis and Machine Intelligence 27*(11), 1805–1819.

Khan, Z., T. Balch, and F. Dellaert (2005b). Multitarget tracking with split and merged measurements. In *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, Volume 1, pp. 605–610. IEEE.

Kreucher, C., M. Morelande, K. Kastella, and A. Hero (2005). Particle filtering for multitarget detection and tracking. In *IEEE Aerospace Conference*, pp. 2101–2116. IEEE.

Ng, W., J. Li, S. Godsill, and S. Pang (2007). Multitarget initiation, tracking and termination using Bayesian Monte Carlo methods. *The Computer Journal 50*(6), 674–693.

Niu, S., Q. Chen, L. de Sisternes, Z. Ji, Z. Zhou, and D. L. Rubin (2017). Robust noise region-based active contour model via local similarity factor for image segmentation. *Pattern Recognition 61*, 104–119.

Park, C. (2013). Estimating multiple pathways of object growth using non-longitudinal image data. *Technometrics, Accepted*.

Park, C. and A. Shrivastava (2013). Multimode geometric profile monitoring with correlated image data and its application to nanoparticle self-assembly processes. *Journal of Quality Technology, Accepted*.

Parvin, B., Q. Yang, J. Han, H. Chang, B. Rydberg, and M. H. Barcellos-Hoff (2007). Iterative voting for inference of structural saliency and characterization of subcellular events. *IEEE Transactions on Image Processing 16*(3), 615–623.

Quelhas, P., M. Marcuzzo, A. M. Mendonça, and A. Campilho (2010). Cell nuclei and cytoplasm joint segmentation using the sliding band filter. *IEEE Transactions on Medical Imaging 29*(8), 1463–1473.

Shi, J. and J. Malik (2000). Normalized cuts and image segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence 22*(8), 888–905.

Tek, F., A. Dempster, and I. Kale (2005). Blood cell segmentation using minimum area watershed and circle radon transformations. In *Proceedings of the 7th International Symposium on Mathematical Morphology*, Paris, France, pp. 441–454.

Vese, L. and T. Chan (2002). A multiphase level set framework for image segmentation using the Mumford and Shah model. *International Journal of Computer Vision 50*(3), 271–293.

Vo, G. and C. Park (2016). Robust matrix decomposition for image segmentation under heavy noises and uneven background intensities. *arXiv preprint arXiv:1609.08078*.

Yan, Q., J. Ye, and X. Shen (2015). Simultaneous pursuit of sparseness and rank structures for matrix decomposition. *Journal of Machine Learning Research 16*, 47–75.

Yu, Q. and G. Medioni (2009). Multiple-target tracking by spatiotemporal monte carlo markov chain data association. *Pattern Analysis and Machine Intelligence, IEEE Transactions on 31*(12), 2196–2210.

Zafari, S., T. Eerola, J. Sampo, H. Kälviäinen, and H. Haario (2015). Segmentation of partially overlapping nanoparticles using concave points. In *International Symposium on Visual Computing*, pp. 187–197. Springer.

Zhou, T. and D. Tao (2011). Godec: Randomized low-rank & sparse matrix decomposition in noisy case. In *Proceedings of the 28th International Conference on Machine Learning (ICML-11)*, pp. 33–40.