Feature-Aided Multiple-Hypothesis Tracking and Classification of Biological Cells

Stefano Coraluppi and Craig Carthel
Compunetix Inc.
2420 Mosside Boulevard, Monroeville, PA 15146 USA
stefano.coraluppi@ieee.org, craig@carthel.com

Samuel J. Dickerson
Nanophoretics LLC
300 Seco Road, Monroeville PA 15146
sam@nanophoretics.com

Donald Chiarulli and Steven Levitan
University of Pittsburgh
Sennott Square Building, Pittsburgh, PA 15260 USA
don@cs.pitt.edu, levitan@pitt.edu

***Abstract*—Our work is motivated by emerging technology for dielectrophoresis-based amplification of cell kinematic responses, allowing for cell classification based on motion in addition to static cell features. Accordingly, the approach requires high-performance multi-target tracking. This paper discusses our approach to the problem, which involves suitable enhancements to track-oriented multiple-hypothesis tracking algorithms. Additionally, we apply a generalized likelihood ratio test for cell classification. Preliminary simulation-based testing and laboratory experimentation with cell and bacteria cultures show encouraging results.**

Keywords—multi-target tracking; multiple-hypothesis tracking; generalized likelihood ratio test; dielectrophoresis; cell tracking and classification.

I. INTRODUCTION

Much of the research on object detection and *multi-target tracking* (MTT) has traditionally been conducted with a focus on security and defense applications. A number of references on multi-object tracking have appeared over the years; see [1-3] and further references therein. Nonetheless, these techniques are applicable to challenges outside the traditional domains. An overview of progress and challenges in cell tracking is provided in [4]. To our knowledge, most of tracking researchers in the security and defense communities have not examined cell-tracking applications, with a rare exception in [5].

Recently, we have applied segmentation-based detection processing and track-oriented *multiple-hypothesis tracking* (MHT) to international tracking challenges organized as part of the IEEE ISBI 2012-2013 conferences [6-7]. Our approach performed competitively against a number of solution schemes tailored to the cell-imaging domain [8-9].

In this paper, we extend our MHT solution to address dielectrophoresis-based cell classification. This involves a number of enhancements. We introduce a generalized *Integrated Ornstein-Uhlenbeck* (IOU) kinematic model for cells excited at a fixed frequency. The parameters of the model depend on the cell type and the excitation frequency; these spectral characteristics are available for many cell types of interest [10-12].

In practice, the *Kalman filter* (KF) based on the generalized IOU model requires knowledge of cell type. This is not available when we have a heterogeneous mixture of cells. Thus, a first approach is to base the KF on a *nearly-constant velocity* (NCV) motion model. This turns out to be problematic due to a non-trivial model mismatch, but it is easily remedied with a time-varying NCV model with higher object process noise at the times of frequency-excitation transition.

MHT is generally acknowledged as the most powerful currently-known paradigm for multi-target tracking. It was first formalized in what is now referred to as *hypothesis-oriented* form in [13]. Unfortunately, the hypothesis-oriented methodology typically leads to an unmanageable number of hypotheses even for small problems. The track-oriented MHT approach was developed in the 1980s by researchers at ALPHATECH and is well-described in [14]. Our key enhancements to (single-stage) MHT are discussed in [15-17].

At the output of the KF-MHT, we reprocess each track (with knowledge of sensor measurements) with a forward-backward Kalman smoother [18]. This provides less erratic trajectories for subsequent track classification, based on the *generalized likelihood ratio test* (GLRT) [19].

II. DIELECTROPHORESIS AND THE GENERALIZED IOU MODEL

Dielectrophoresis is the use of non-uniform AC electric fields to control the motion of particles in a fluid [20-21]. System calibration and subsequent screening, identification, assay, and characterization of micro-scale particle mixtures requires multi-target technology.

The real and imaginary dielectrophoresis force components can be measured indirectly by configuring a known electric field and observing the particle's velocity. If these observations are made over a wide range of frequencies, the unique responses of the particles can be extracted. Figure 1 illustrates the velocity spectrum, i.e. velocity as a function of excitation frequency, for some particles of interest.

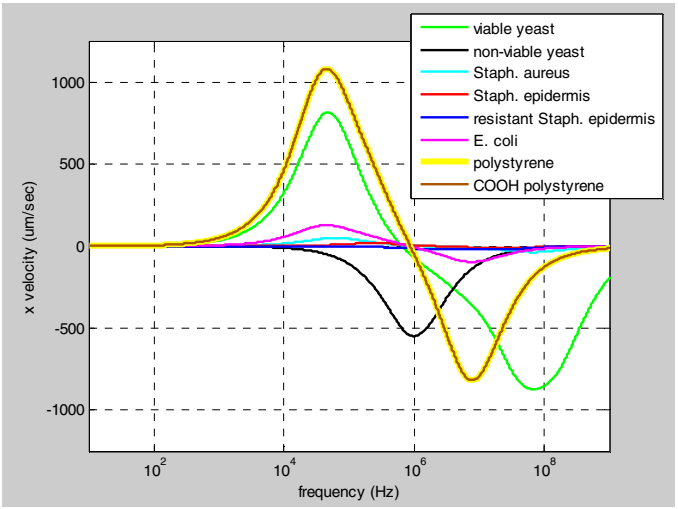

Figure 1. Some steady-state velocity spectra.

The nominal velocity spectrum depends on the geometry and other electrode characteristics, the fluid medium, and about a dozen particle characteristics including number and thickness of shells, permittivities, conductances, etc. [10-12, 20-21]. Variability in these physical parameters induces variability in the spectrum. For example, a 5% variability in all physical parameters (as a multiplicative scaling of the nominal) induces a sample mean that is very close to the nominal velocity spectrum, and a sample standard deviation of 13% (again, as a multiplicative scaling of the nominal). A more pronounced variability of 20% leads to the responses shown in Figure 2.

$$\xi_i = \frac{1}{N}\sum_i \frac{v_i - \bar{v}}{\bar{v}}, \tag{1}$$

$$\hat{\xi} = \frac{1}{N}\sum_i \xi_i, \tag{2}$$

$$\hat{\sigma}_\xi = \left(\frac{1}{N-1}\sum_i \left(\xi_i - \hat{\xi}\right)^2\right)^{1/2} = 0.127. \tag{3}$$

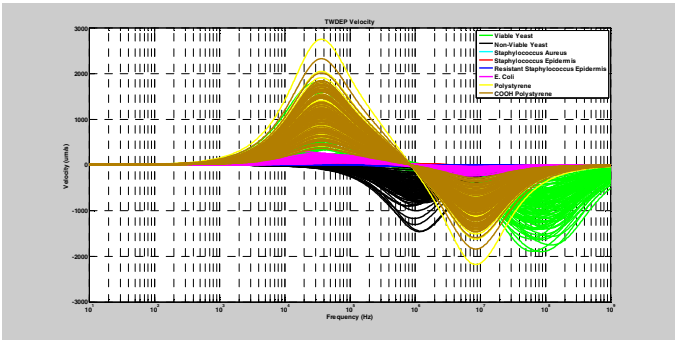

Figure 2. Sample realizations of spectra with varying particle physical characteristics.

As illustrated in Figure 3, we introduce a straightforward generalization to the IOU process [2] to model cell dynamics. This model allows us to address the impact of transitions in excitation frequency. The same model applies in each of two physical dimensions; the drift velocity v will be different in each dimension. The feedback gain and process noise intensity are set to model realistic cell trajectories.

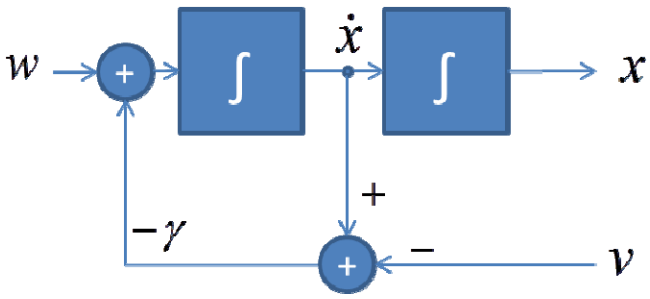

Figure 3. Generalized IOU process.

In addition to the kinematic state, we include as well a static cell feature state that includes *rgb* color and major and minor axes in the detection ellipsoidal approximation. See Figure 4. As with the kinematic model, we consider a distinct nominal feature state for each cell type, and allow for variability for each individual cell.

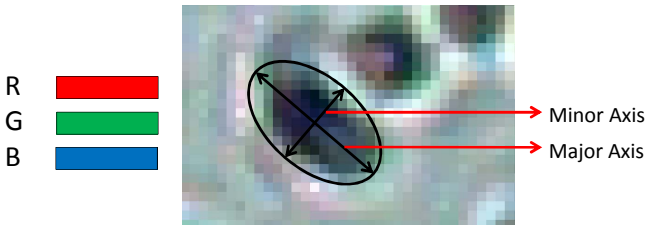

Figure 4. Cell feature state.

III. Kalman Filtering, MHT Processing, and Trajectory Smoothing

Using standard state-space notation, we have the following discretization of the generalized IOU model.

$$X_{k+1} = A_k X_k + \int_{t_k}^{t_{k+1}} \exp\left(F\left(t_{k+1}-t\right)\right)\begin{bmatrix}0\\ \gamma v + w(t)\end{bmatrix} dt, \tag{4}$$

$$\begin{aligned} X_{k+1} &= A_k X_k + w_k + \int_{t_k}^{t_{k+1}} \begin{bmatrix} \frac{1-\exp\left(-\gamma\left(t_{k+1}-t\right)\right)}{\gamma}\gamma v \\ \exp\left(-\gamma\left(t_{k+1}-t\right)\right)\gamma v \end{bmatrix} dt \\ &= A_k X_k + w_k + \begin{bmatrix} v\Delta t_k - \frac{v}{\gamma}\left(1-\exp\left(-\gamma\Delta t_k\right)\right) \\ v\left(1-\exp\left(-\gamma\Delta t_k\right)\right) \end{bmatrix}, \end{aligned} \tag{5}$$

$$A_k = \begin{bmatrix} 1 & \frac{1-\exp\left(-\gamma_2\Delta t_k\right)}{\gamma_2} \\ 0 & \exp\left(-\gamma_2\Delta t_k\right) \end{bmatrix}. \tag{6}$$

Track management and data association are the key technical challenges in MTT. One is given a sequence of sets of measurements, and must determine which measurements to associate and which to discard. The track-oriented MHT paradigm is generally acknowledged as the most powerful currently-known paradigm for MTT [1, 3, 14-17]. Our feature-

aided MHT is based on noisy positional and cell-feature measurements.

We have a sequence of sets of contacts $Z^k = (Z_1, ..., Z_k)$, and we wish to estimate the state history X^k for all objects present in the surveillance region. X^k is compact notation that represents the state trajectories of targets that exist over the time sequence $(t_1, ..., t_k)$. Note that each target may exist for a subset of these times, with a single birth and a single death occurrence, i.e. targets do not reappear. We introduce the auxiliary discrete state history q^k that represents a full interpretation of all contact data: which contacts are false, how the object-originated ones are to be associated, and when objects are born and die. There are two fundamental assumptions of note. The first is that there are no target births in the absence of a corresponding detection, i.e. we do not reason over new, undetected objects. The second is that there is *at most* one contact per object per scan.

We are interested in the probability distribution $p(X^k \mid Z^k)$ for object state histories given data. This quantity can be obtained by conditioning over all possible auxiliary states histories q^k.

$$p(X^k \mid Z^k) = \sum_{q^k} p(X^k \mid Z^k, q^k) p(q^k \mid Z^k). \tag{7}$$

The MHT approach seeks to identify the MAP estimate for the auxiliary state history q^k, and identify the corresponding MMSE estimate for the object state history X^k conditioned on the estimate for q^k.

$$\hat{q}^k = \arg\max_{q^k} p(q^k \mid Z^k), \tag{8}$$

$$\hat{X} = \hat{X}_{MMSE}(Z^k, \hat{q}^k). \tag{9}$$

Track-oriented MHT avoids enumeration of all global hypotheses q^k, though these are implicitly defined in the set of track hypotheses trees. We illustrate this with a simple example. Assume that the number of contacts received in three scans of data are given by $|Z_1| = 1$, $|Z_2| = 1$, and $|Z_1| = 1$. Figure 5 illustrates the set of global hypotheses, while Figure 6 illustrates the set of track trees that identify all *local* or *track* hypotheses. In the figures, ⊗ denotes object death while O denotes a missed detection event. With the track-oriented approach, we are to select a set of leaves that identify the MAP solution (8), with the feasibility constraint that all measurements are utilized at most once.

In practice, computational requirements preclude optimal, batch MHT solutions. Practical MHT solutions adopt a number of computation simplifications that include the following.

1. We do not reason over undetected object birth hypotheses (indeed, no such hypotheses are represented in the example above). In some settings, a broader set of hypotheses is necessary [17].
2. We limit hypothesis growth by hypothesis gating techniques that disallow sufficiently unlikely associations. Further, we do not spawn both missed-detection and death hypotheses as in the example above. Generally, only a missed detection hypothesis is spawned, and after a sufficient number of missed detections only a target death hypothesis is spawned.
3. The track hypothesis trees are pruned so as to result in a single global hypothesis with a fixed delay. Hypothesis pruning generally relies on relaxation techniques for an otherwise challenging integer optimization problem. Well-known techniques include Lagrangian relaxation and linear programming [15, 22].
4. Track extraction is performed sequentially using logic-based or statistical tests including the *sequential probabilistic ratio test* (SPRT) [1]. We opt for the former approach as in practice the SPRT generally is quite similar to an *M-of-N* confirmation test.

$$\left\{\begin{array}{l}
\{(Z_1,Z_2,Z_3),(Z_{22}O),\{(Z_1,Z_2,Z_3),(Z_{22}\otimes)\},\\
\{(Z_1,Z_{22},Z_3),(Z_2O),\{(Z_1,Z_{22},Z_3),(Z_2\otimes)\},\\
\{(Z_1,Z_2,O),(Z_{22},O),(Z_3)\},\{(Z_1,Z_2,O),(Z_{22}\otimes),(Z_3)\},\{(Z_1,Z_2,\otimes),(Z_{22}O),(Z_3)\},\{(Z_1,Z_2,\otimes),(Z_{22}\otimes),(Z_3)\},\\
\{(Z_1,Z_{22},O),(Z_2,O),(Z_3)\},\{(Z_1,Z_{22},O),(Z_2\otimes),(Z_3)\},\{(Z_1,Z_{22}\otimes),(Z_2,O),(Z_3)\},\{(Z_1,Z_{22}\otimes),(Z_2\otimes),(Z_3)\},\\
\{(Z_1,O,Z_3),(Z_2,O),(Z_{22}O)\},\{(Z_1,O,Z_3),(Z_2,O),(Z_{22}\otimes)\},\{(Z_1,O,Z_3),(Z_2\otimes),(Z_{22}O)\},\{(Z_1,O,Z_3),(Z_2\otimes),(Z_{22}\otimes)\}\\
\{(Z_1,O,O),(Z_2,Z_3),(Z_{22}O)\},\{(Z_1,O\otimes),(Z_2,Z_3),(Z_{22}O)\},\{(Z_1\otimes),(Z_2,Z_3),(Z_{22}O)\},\\
\{(Z_1,O,O),(Z_2,Z_3),(Z_{22}\otimes)\},\{(Z_1,O\otimes),(Z_2,Z_3),(Z_{22}\otimes)\},\{(Z_1\otimes),(Z_2,Z_3),(Z_{22}\otimes)\},\\
\{(Z_1,O,O),(Z_{22},Z_3),(Z_2,O)\},\{(Z_1,O\otimes),(Z_{22},Z_3),(Z_2,O)\},\{(Z_1\otimes),(Z_{22},Z_3),(Z_2,O)\},\\
\{(Z_1,O,O),(Z_{22},Z_3),(Z_2,\otimes)\},\{(Z_1,O\otimes),(Z_{22},Z_3),(Z_2,\otimes)\},\{(Z_1\otimes),(Z_{22},Z_3),(Z_2,\otimes)\},\\
\{(Z_1,O,O),(Z_2,O),(Z_{22},O),(Z_3)\},\{(Z_1,O\otimes),(Z_2,O),(Z_{22},O),(Z_3)\},\{(Z_1\otimes),(Z_2,O),(Z_{22},O),(Z_3)\},\\
\{(Z_1,O,O),(Z_2,O),(Z_{22}\otimes),(Z_3)\},\{(Z_1,O\otimes),(Z_2,O),(Z_{22}\otimes),(Z_3)\},\{(Z_1\otimes),(Z_2,O),(Z_{22}\otimes),(Z_3)\},\\
\{(Z_1,O,O),(Z_2\otimes),(Z_{22},O),(Z_3)\},\{(Z_1,O\otimes),(Z_2\otimes),(Z_{22},O),(Z_3)\},\{(Z_1\otimes),(Z_2\otimes),(Z_{22},O),(Z_3)\},\\
\{(Z_1,O,O),(Z_2\otimes),(Z_{22}\otimes),(Z_3)\},\{(Z_1,O\otimes),(Z_2\otimes),(Z_{22}\otimes),(Z_3)\},\{(Z_1\otimes),(Z_2\otimes),(Z_{22}\otimes),(Z_3)\},
\end{array}\right\}$$

Figure 5. The set of global hypotheses. This explicit enumeration is needed under hypothesis-oriented MHT.

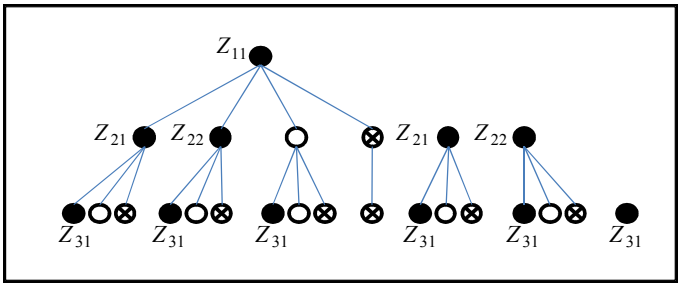

Figure 6. The set of track hypothesis trees that implicitly define global hypotheses. This formulation is adopted under track-oriented MHT.

The key enabler for track-oriented MHT is the recursive formulation for the posterior probability of a global hypothesis given below. This expression relies on the assumption that in each scan the number of target births is Poisson distributed with mean λ_b, the number of false returns is Poisson distributed with mean λ_{fa}, targets die with probability p_χ, and targets are detected with probability p_d. Further, r is the number of measurements, τ is the number of tentative tracks under q^{k-1} and, with q^k, d is the number of detections, χ is the number of track deaths, and b is the number of new tracks.

Correspondingly, $r-d-b$ is the number of false returns, where r is the number of contacts in the current scan. Finally, J_d, J_b, and J_{fa} denote the sets of detections, births, and false alarms, respectively, with $|J_d|+|J_b|+|J_{fa}|=r$.

$$p(q^k \mid Z^k) = p_\chi^\chi\left((1-p_\chi)(1-p_d)\right)^{\tau-\chi-d} \cdot \prod_{j\in J_d}\left[\frac{(1-p_\chi)p_d f_d(z_j \mid Z^{k-1},q^k)}{\lambda_{fa} f_{fa}(z_j \mid Z^{k-1},q^k)}\right] \cdot \prod_{j\in J_b}\left[\frac{p_d \lambda_b f_b(z_j \mid Z^{k-1},q^k)}{\lambda_{fa} f_{fa}(z_j \mid Z^{k-1},q^k)}\right]\frac{p(q^{k-1} \mid Z^{k-1})}{\bar{c}_k}, \quad (10)$$

$$\bar{c}_k = \frac{c_k}{\left\{\frac{\exp(-\lambda_b-\lambda_{fa})}{r!}\lambda_{fa}^r\right\}\prod_{j\in J_d\cup J_b\cup J_{fa}} f_{fa}(z_j \mid Z^{k-1},q^k)}. \quad (11)$$

In the linear Gaussian case, $f_d(z_j \mid Z^{k-1},q^k)$ is a Gaussian residual, i.e. it is the probability of observing z_j given a sequence of preceding measurements. If there is no prior information on objects, $f_b(z_j \mid Z^{k-1},q^k)$ is generally the value of the uniform density function over measurement space. Similarly, $f_{fa}(z_j \mid Z^{k-1},q^k)$ is as well usually taken to be the value of the uniform density function over measurement space, under the assumption of uniformly distributed false returns. Note that the expressions given here are general and allow for quite general target and sensor models.

Figure 7 illustrates hypothesis generation and resolution for a small example, where update decisions for two confirmed tracks are made by reasoning over two scans of data. Decisions are made with some latency (here, a one-scan delay).

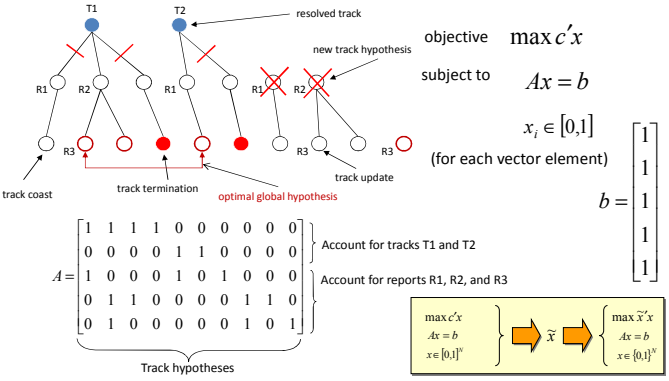

Figure 7. Linear-programming approach to hypothesis pruning in track-oriented MHT.

An illustration of resulting positional and x-velocity trajectories are in Figures 8-9. (Our electrode configuration is such that there is no drift velocity in the y dimension.)

While recursive filtering is sufficient for data-association purposes, track-smoothing aids subsequent track classification. There are several possible implementations of the Kalman smoother; we use the forward-backward filter [23]. Figure 10 illustrates the benefit of trajectory smoothing.

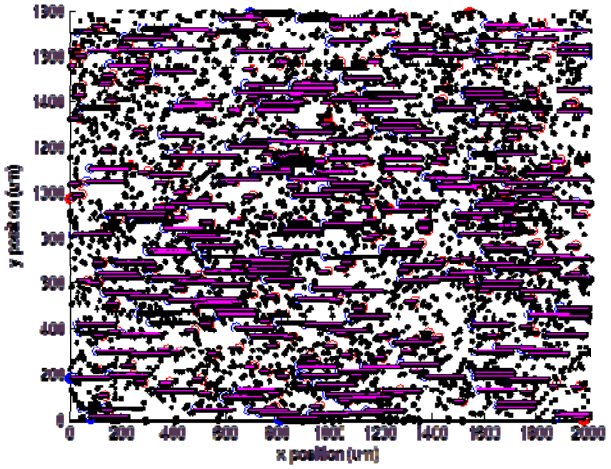

Figure 8. Positional trajectories of live and dead yeast cells.

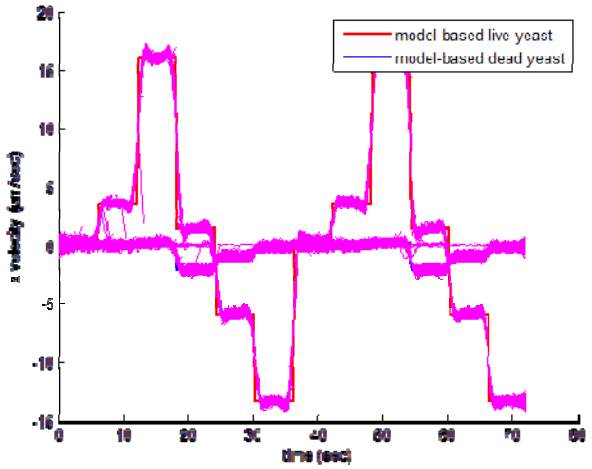

Figure 9. Velocity trajectories (note imaging-boundary sticking).

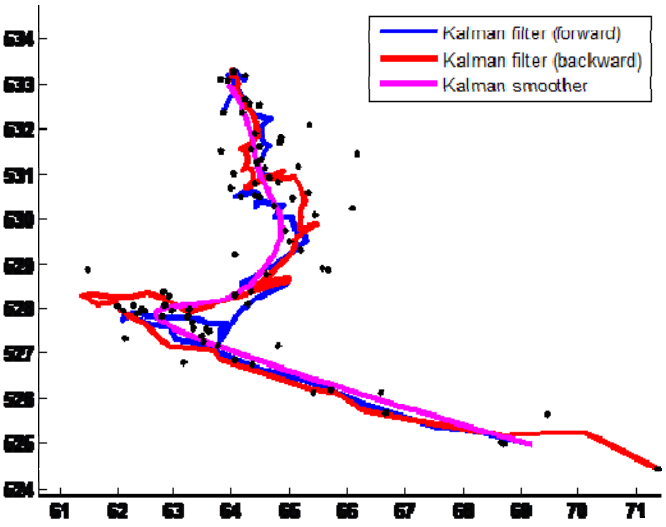

Figure 10. An improved cell track after Kalman smoothing.

As part of our MHT processing, Kalman-based filtering and smoothing based on the generalized IOU model is in principle better than with a NCV motion model that introduces a modeling mismatch (see Figure 11). Smoothing requires back-prediction equations, which relies on the following:

$$X_k = A_k^{-1}\left(X_{k+1} - B_k v - w_k\right) + \begin{bmatrix} 0 \\ v \end{bmatrix}, \tag{12}$$

$$A_k^{-1} = \begin{bmatrix} 1 & \dfrac{1-\exp(-\gamma\Delta t_k)}{\gamma} \\ 0 & \exp(-\gamma\Delta t_k) \end{bmatrix}^{-1} = \begin{bmatrix} 1 & \dfrac{1-\exp(\gamma\Delta t_k)}{\gamma} \\ 0 & \exp(\gamma\Delta t_k) \end{bmatrix}. \tag{13}$$

In practice, the NCV model is more robust due to imaging-boundary effects. Also, with the generalized IOU model, the drift velocity is particle-type dependent. Parameters of the NCV model are type-dependent as well, but can be selected for adequate filtering performance over a range of cell types.

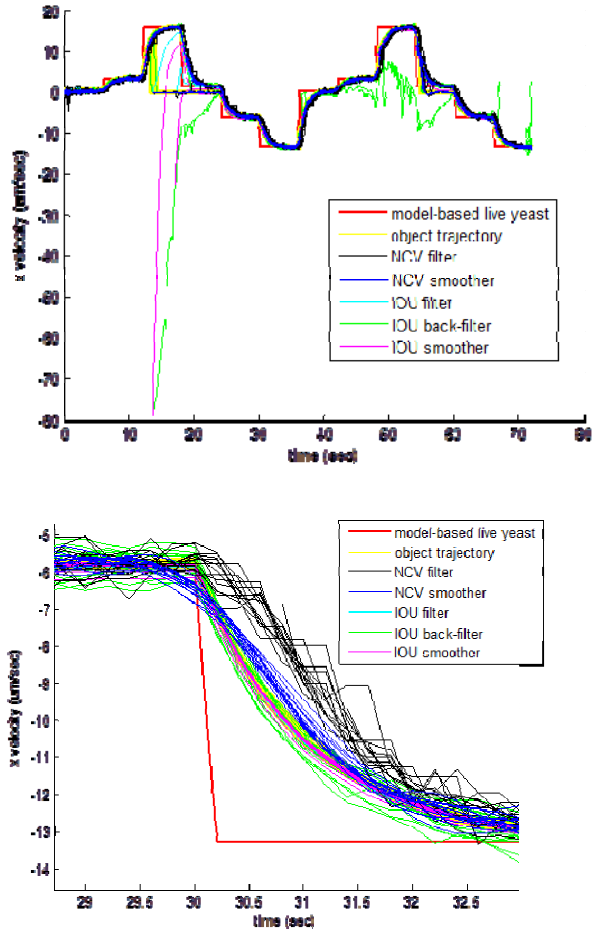

Figure 11. Better performance (but less robustness) with IOU smoother.

A key aspect to our application is that frequency-excitation transition times are known. Accordingly, we find that there is significant benefit to use of time-varying process noise in the NCV-based KF. In particular, we use a higher process noise intensity for a time interval following the transition, and lower process noise intensity thereafter, until the next transition time. Trajectory smoothing is based on the lower process noise. This approach proves to be robust and effective (see Figures 12-14).

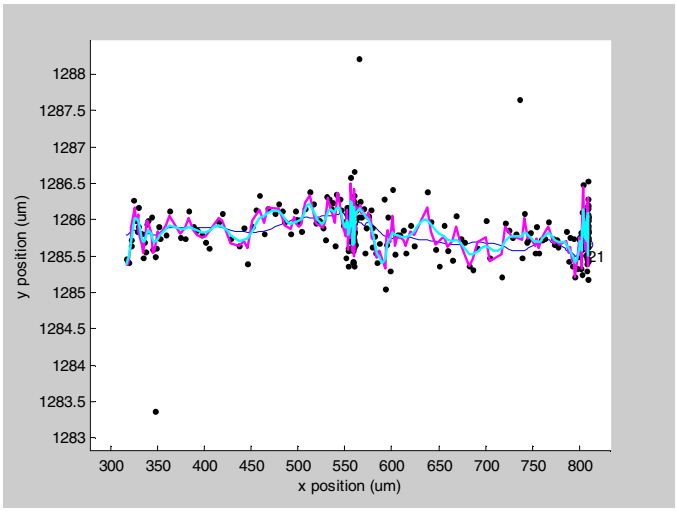

Figure 12. Drug-resistant Staph. epidermis: positional trajectories (blue – ground truth, magenta – filtering, cyan – smoothing).

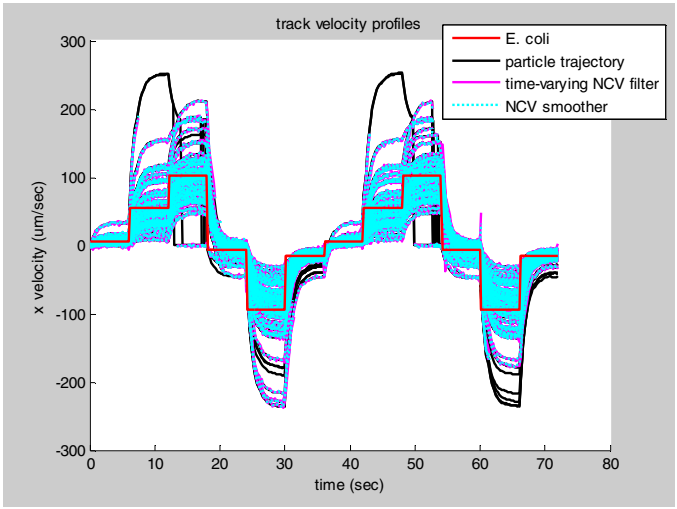

Figure 13. E. coli: velocity trajectories.

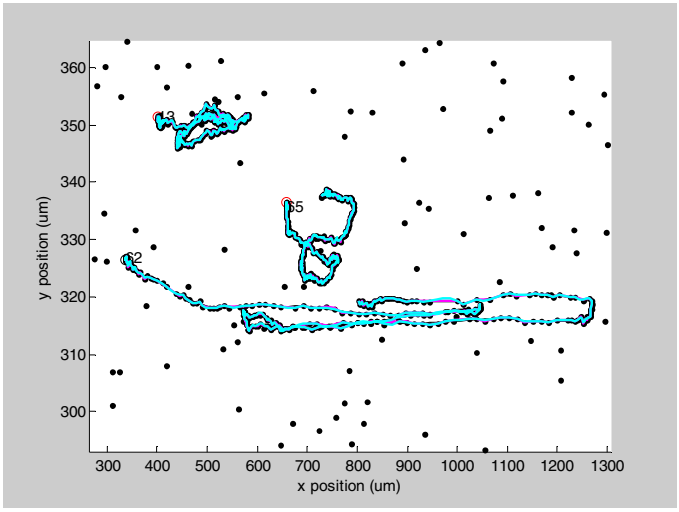

Figure 14. Mixed-type trajectories.

IV. Track Classification

The smoothed single-object trajectories constitute the input the GLRT-based track classification [24]. The classifier requires calibration (training) for each cell type. We determine the first and second moments of a statistic that includes cell x-velocity at a sequence of excitation frequencies, as well as the five static features states (cell color and dimensions). Subsequently, trajectories from an arbitrary mixture of cell types are scored with respect to each calibrated type. We consider as well type *other*, within the standard GLRT paradigm.

The essential calibration equations are the following:

$$\mu_{i,j} = \frac{1}{N}\sum_{N\,tracks} X_s\,, \tag{14}$$

$$\Sigma^2_{i,j} = \frac{1}{N-1}\sum_{N\,tracks}\left(X_s - \mu_{i,j}\right)^2\,. \tag{15}$$

Our GLRT classifier is simplified in that we first determine the best-scoring cell type for each cell track. Then, we compare against the best-scoring *other*, by considering all profiles outside an exclusion region around the best-scoring cell type. In turn, this evaluation is simplified due to the properties of the assumed Gaussian statistics, given by eqn. (18) and illustrated in Figure 15.

$$p(z \mid H_i) = N(z;\mu_i,\Sigma_i), \tag{16}$$

$$p(z \mid H_0) = \max_{\mu \notin \bigcup_i [\mu_i-\Delta\ \ \mu_i+\Delta],\,\Sigma>\Sigma_{\min}} N(z;\mu,\Sigma)\,. \tag{17}$$

$$\frac{d}{d\Sigma}\log N(z;\mu,\Sigma) = \frac{d}{d\Sigma}\left(-\frac{1}{2}\log(2\pi\Sigma) - \frac{(z-\mu)^2}{2\Sigma}\right) \tag{18}$$
$$= -\frac{1}{2\Sigma} + \frac{(z-\mu)^2}{2\Sigma^2} = 0 \Rightarrow \Sigma = (z-\mu)^2$$

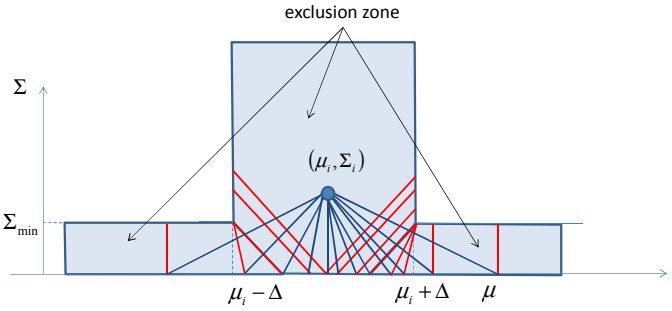

Figure 15. Illustration of comparison of data to ML type as well as to best-scoring "other".

Misclassification of cell tracks may well result, particularly for cells with similar velocity spectrum or as a result of track fragmentation. Figure 16 illustrates this phenomenon when the track fragment may reasonably result from one of two cell types. All cells in the example are dead yeast; one track is classified nonetheless as live yeast.

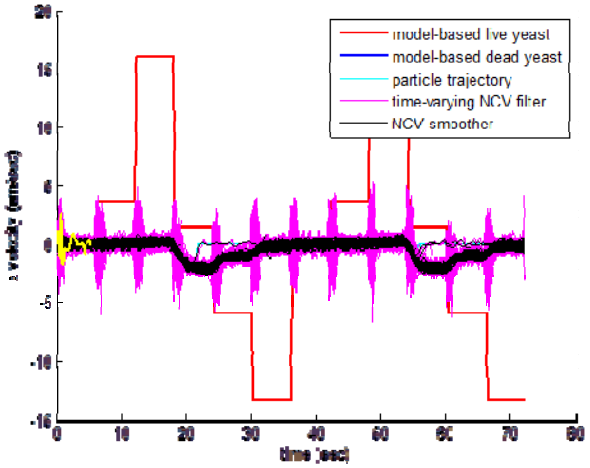

Figure 16. Misclassification example due to track fragmentation.

The GLRT classifier does not require a prior distribution on object type. This is critical, since estimating the number of objects and their distribution over type is precisely the objective of our diagnostic system.

V. Track Aggregation and Statistical Classification Performance

It is difficult to assess directly the performance of the GLRT-based classifier, since in general tracks may originate from measurements due to more than one object as well as false returns. Thus, common metrics of interest in the diagnostic community, like specificity and sensitivity, cannot be measured directly [25]. We choose instead to assess overall system performance with an aggregate assessment of classification performance.

The fundamental question is how to quantify the discrepancy between two vectors of numbers, with each representing the number of objects of a given type, and the last representing type "other". We introduce a Jaccard metric J that is motivated by the Jaccard index that is defined on sets [26], as well as a modified (non-symmetric) Jaccard metric $\bar{J}$ that focuses on discrepancies with respect to the true number of objects of each type. We define these as follows: consider as well type *other*, within the standard GLRT paradigm.

$$J(n,m) = \frac{1}{N}\sum_i \frac{\max\{n_i,m_i\} - \min\{n_i,m_i\}}{\max\{n_i,m_i\}}\,, \tag{19}$$

$$\bar{J}(n,m) = \frac{1}{N}\sum_i \frac{\max\{n_i,m_i\} - \min\{n_i,m_i\}}{n_i}\,. \tag{20}$$

Automatic tracking incurs track fragmentation errors. Thus, we do not apply the metrics (19-20) to the raw classifier output. Rather, we consider two *track aggregation* schemes that reduce the number of declared tracks of each type prior to classification assessment. The first scheme – *lifetime-based aggregation* – computes, for each type, the ratio of the sum of track lifetimes to total scenario duration. The second scheme – *concurrency-based aggregation* – computes the maximum number of tracks of each type in existence at any time.

VI. Simulation Results

Thus far, we have conducted extensive testing of simulation-based testing with the following elements: ground truth generation according to the generalized IOU model with calibrated velocity spectra while allowing for variation in underlying physical parameters; image-processing simulation with a fixed detection probability, location and feature measurements with additive Gaussian noise, and uniformly distributed false alarms in imaging space (and with corresponding feature measurement based on a randomly-selected cell type); MHT tracking and subsequent Kalman smoothing; GLRT-based track classification; track aggregation and performance evaluation according to (19-20).

Experimentation requires a *training* phase that allows for system calibration, followed by a *testing* phase with an arbitrary mix of cell types. Tables 1-2 illustrate representative results from a simulation run with 100 cells that include Staph. epidermis, drug-resistant Staph. epidermis, E. coli, and Staph. aureus. Training is based only on the first three types, while Staph. aureus represents an (unknown) "other". Position and velocity trajectories for the testing phase are shown in Figures 17-18. To date, our simulation-based testing has included 100-1000 particles per run, in an imaging region of 1.8mm by 2mm that matches our actual experimental setup.

Table 1. Classifier output.

Type	Staph. epidermis	drug-resistant Staph.	E. coli	Other (Staph. aureus)
ground truth	19	31	24	26
raw output	27	35	10	65
lifetime-based aggregation	18	25	6	42
concurrency-based aggregation	21	31	5	40

Table 2. System performance.

metric	raw output	lifetime aggregation	concurrency aggr.
Jaccard	0.398	0.344	0.309
mod. Jaccard	0.658	0.403	0.359

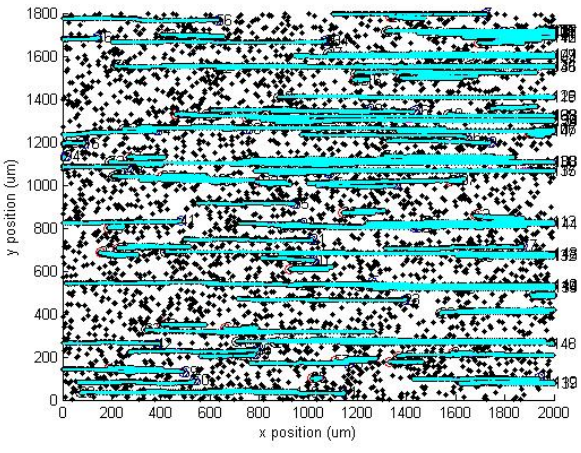

Figure 17. Measurement data and track positional trajectories.

It is worth emphasizing that the simulation-based capability described here is a valuable tool for system testing. The numbers reported in Tables 1-2 are a function of many processing parameters that have not been optimized to date. We see improved performance with track aggregation, evaluated with both metrics. The metrics are zero for ideal system performance, and the Jaccard metric is bounded by one.

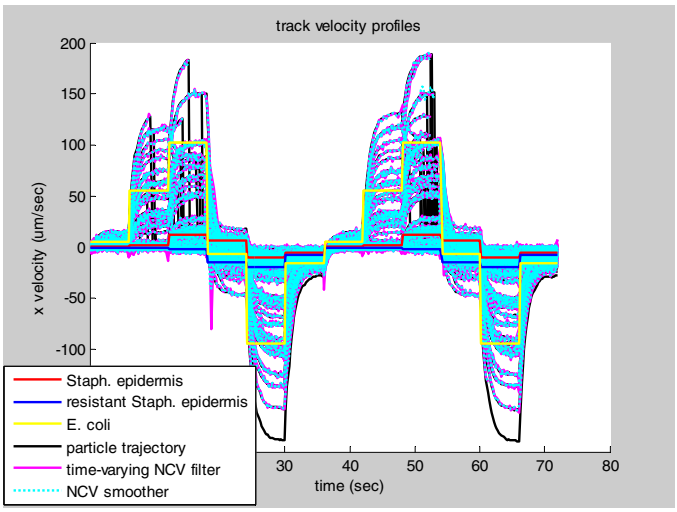

Figure 18. Nominal and actual track velocity profiles.

VII. Laboratory Experimentation

In experimental settings, particle-detection processing is required. Thus, we apply image-segmentation algorithms from which only those elements are extracted for which dimensions were roughly consistent with yeast cell sizes. Our early experience in this area stems from the cell tracking challenges mentioned previously [8-9]. One of these challenge datasets is illustrated in Figure 19, where object trajectories (blue) and the tracking output (red) are overlaid. The challenge problems include a mix of real and synthetic data of varying complexity.

Our initial laboratory experimentation has focused on mixtures of live and dead yeast cells. A tracking output is illustrated in Figure 20. More recently, we have begun experimentation with Staph. epidermis and E. coli. Detection processing is illustrated in Figure 21. Overall system testing with Staph. epidermis and E. coli will be the focus of the coming months of research and experimentation.

VIII. Conclusions

Dielectrophoresis is the use of non-uniform AC electric fields to control the motion of particles in a fluid. Our dielectrophoresis-based technology for the classification of bacteriological pathogens requires high-quality detection, tracking and classification of biological cells from microscope imagery. We employ mature track-oriented MHT technology, trajectory smoothing, and GLRT-based track classification. The Jaccard metric provides aggregate system performance.

We have significant control of scenario and sensor parameters. This includes the selection of the density and motion of particles via the choice of excitation frequencies. The frequency vs. terminal velocity profile for particles of interest is known, and may be exploited for high-performance filtering. We step through a sequence of frequencies for a fixed duration.

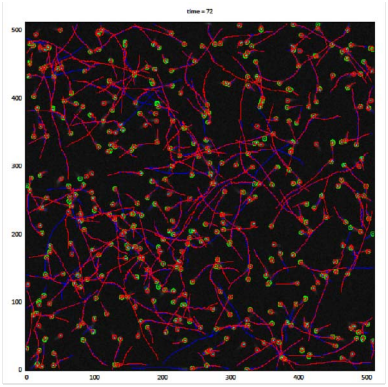

Figure 19. Wide-area view of an IEEE ISBI tracking scenario.

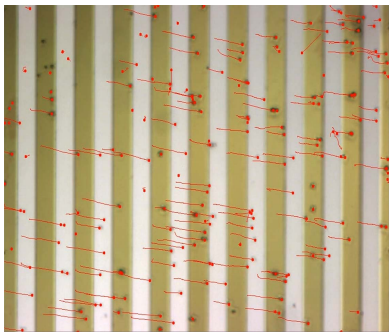

Figure 20. Tracking yeast excited via dielectrophoresis.

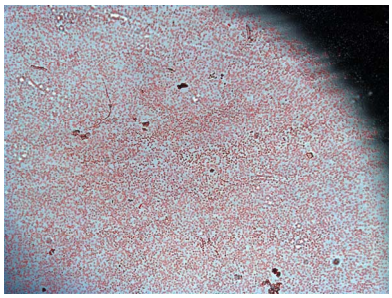

Figure 21. Detection processing in bacteria images.

Our approach to particle classification employs a *nearly-constant velocity* (NCV) motion model and decoupled tracking and classification modules with track smoothing prior to GLRT-based classification using track kinematic and feature information. Since the frequency-excitation transition times are known, this information may be exploited for improved performance. Thus, we employ time-varying process noise statistics in the NCV filter.

An enhanced approach is to model particle motion via a generalization IOU process. This ultimately requires a multiple-model filtering solution and a joint tracking and classification solution. While this approach holds the promise of improved performance, there are some robustness concerns that are the subject of ongoing research.

REFERENCES

[1] S. Blackman and R. Popoli, *Design and Analysis of Modern Tracking Systems*, Artech House, 1999.

[2] L. Stone, C. Barlow, and T. Corwin, *Bayesian Multiple Target Tracking*, Artech House, 1999.

[3] Y. Bar-Shalom, P. Willett, and X. Tian, *Tracking and Data Fusion: A Handbook of Algorithms*, YBS Publishing, 2011.

[4] E. Meijering, I. Smal, and G. Danuser, Tracking in Molecular Bioimaging, *IEEE Signal Processing Magazine*, vol. 23(3), May 2006.

[5] T. Kirubarajan, Y. Bar-Shalom, and K. Pattipati, Multiassignment for Tracking a Large Number of Overlapping Objects, *IEEE Transactions on Aerospace and Electronic Systems*, vol. 37(1), Jan. 2001.

[6] http://bioimageanalysis.org/track/

[7] http://www.codesolorzano.com/celltrackingchallenge

[8] N. Chenouard, I. Smal, F. de Chaumont, M. Maška, I. Sbalzarini, Y. Gong, J. Cardinale, C. Carthel, S. Coraluppi, M. Winter, A. Cohen, W. Godinez, K. Rohr, Y. Kalaidzidis, L. Liang, J. Duncan, H. Shen, Y. Xu, K. Magnusson, J. Jaldén, H. Blau, P. Paul-Gilloteaux, P. Roudot, C. Kervrann, F. Waharte, J.-Y. Tinevez, S. Shorte, J. Willemse, K. Celler, G. Wezel, H.-W. Dan, Y.-S. Tsai, C. Ortiz de Solórzano, J.-C. Olivo-Marin, and E. Meijering, Objective Comparison of Particle Tracking Methods, *Nature Methods*, pp. 1-10, January 2014.

[9] M. Maška, V. Ulman, D. Svoboda, P. Matula, P. Matula, C. Ederra, A. Urbiola, T. España, S. Venkatesan, D. Balak, P. Karas, T. Bolcková, M. Štreitová, C. Carthel, S. Coraluppi, N. Harder, K. Rohr, K. Magnusson, J. Jaldén, H. Blau, O. Dzyubachyk, P. Krížek, G. Hagen, D. Pastor-Escuredo, D. Jimenez-Carretero, M. Ledesma-Carbayo, A. Muñoz-Barrutia, E. Meijering, M. Kozubek, C. Ortiz-de-Solorzano, A Benchmark for Comparison of Cell Tracking Algorithms, to appear in *Bioinformatics*.

[10] Y. Huang, R. Hölzel, R. Pethig, X.-B. Wang, Differences in the AC electrodynamics of viable and non-viable yeast cells determined through combined dielectrophoresis and electrorotation studies, *Physics in Medicine and Biology*, vol. 37(7), pp. 1499-1517, 1992.

[11] J. Johari, Y. Hübner, J. Hull, J. Dale, and M. Hughes, Dielectrophoretic assay of bacterial resistance to antibiotics, *Physics in Medicine and Biology*, vol. 48, pp. 193-198, 2003.

[12] A. Sanchis, A. Brown, M. Sancho, G. Martínez, J. L. Sebastián. S. Muñoz, and J. Miranda, Dielectric Characterization of Bacterial Cells Using Dielectrophoresis, *Bioelectromagnetics*, vol. 28, pp. 393-401, 2007.

[13] D. Reid, An Algorithm for Tracking Multiple Targets, *IEEE Transactions on Automatic Control*, vol. 24(6), December 1979.

[14] T. Kurien, "Issues in the Design of Practical Multitarget Tracking Algorithms," Chapter 3, in *Multitarget-Multisensor Tracking: Advanced Applications*, Y. Bar-Shalom (Ed.), Artech House, 1990.

[15] S. Coraluppi and C. Carthel, Recursive Track Fusion for Multi-Sensor Surveillance, *Information Fusion*, vol. 5(1), pp. 23-33, March 2004.

[16] S. Coraluppi and C. Carthel, Modified Scoring in Multiple-Hypothesis Tracking, *ISIF Journal of Advances in Information Fusion*, vol. 7(2), pp. 153-164, December 2012.

[17] S. Coraluppi and C. Carthel, If a Tree Falls in the Woods, it Does Make a Sound: Multiple-Hypothesis Tracking with Undetected Target Births, to appear in *IEEE Transactions on Aerospace and Electronic Systems*.

[18] A. Gelb (Ed.), *Applied Optimal Estimation*, The MIT Press, Cambridge, MA, 1974.

[19] H. V. Poor, *An Introduction to Signal Detection and Estimation*, Springer-Verlag, 1988.

[20] Pohl, Herbert A. The Motion and Precipitation of Suspensoids in Divergent Electric Fields, *J. Appl. Phys.*, vol. 22(7), 1951.

[21] S. Dickerson, D. Chiarulli, and S. Levitan, 3D Integrated Circuits for Lab-on-Chip Applications, in *Proceedings of the IEEE International 3D Systems Integration Conference*, San Francisco CA, September 2009.

[22] A. Poore and N. Rijavec, A Lagrangian Relaxation Algorithm for Multidimensional Assignment Problems Arising from Multitarget Tracking, *SIAM J. Optimization*, vol. 3(3), August 1993.

[23] A. Gelb, *Applied Optimal Estimation*, MIT Press, 1974.

[24] H. V. Poor, *An Introduction to Signal Detection and Estimation*, Springer-Verlag, 1988.

[25] D. Altman and J. Bland, Diagnostic tests 1: sensitivity and specificity, *British Medical Journal*, vol. 308, June 1994.

[26] http://en.wikipedia.org/wiki/Jaccard_index.