Large-Scale Fiber Tracking Through Sparsely Sampled Image Sequences of Composite Materials

Youjie Zhou, *Student Member, IEEE*, Hongkai Yu, *Student Member, IEEE*, Jeff Simmons, *Member, IEEE*, Craig P. Przybyla, and Song Wang, *Senior Member, IEEE*

***Abstract*— Fast and accurate characterization of fiber micro-structures plays a central role for material scientists to analyze physical properties of continuous fiber reinforced composite materials. In materials science, this is usually achieved by continuously cross-sectioning a 3D material sample for a sequence of 2D microscopic images, followed by a fiber detection/tracking algorithm through the obtained image sequence. To speed up this process and be able to handle larger size material samples, this paper proposes sparse sampling with larger inter-slice distance in cross sectioning and develops a new algorithm that can robustly track large-scale fibers from such a sparsely sampled image sequence. In particular, the problem is formulated as multi-target tracking, and the Kalman filters are applied to track each fiber along the image sequence. One main challenge in this tracking process is to correctly associate each fiber to its observation given that: 1) fiber observations are of large scale, crowded, and show very similar appearances in a 2D slice and 2) there may be a large gap between the predicted location of a fiber and its observation in the sparse sampling. To address this challenge, a novel group-wise association algorithm is developed by leveraging the fact that fibers are implanted in bundles and the fibers in the same bundle are highly correlated through the image sequence. In experiments, the proposed algorithm is tested on three tiles of 100-slice S200 material samples and the tracking performance is evaluated using 1136 human annotated ground-truth fiber tracks. Both quantitative and qualitative results show that the proposed algorithm clearly outperforms the state-of-the-art multiple-target tracking algorithms on sparsely sampled image sequences.**

***Index Terms*— Large-scale fiber tracking, Kalman filters, thin-plate splines, sparse imaging.**

I. Introduction

CONTINUOUS fiber reinforced composite (FRC) materials have been playing a very important role in modern industry [1], [2] because their strength and stiffness are far better than those of traditional materials [2]. These superior properties of FRC materials are largely dependent on the micro-structure of the reinforced fibers. For example, the strength of a FRC material is generally much higher along the direction of reinforced fibers than the direction that is perpendicular to the fibers. Fast and accurate characterization of the underlying fiber micro-structure can substantially speed up the design and development of new composite materials.

To unveil the underlying large-scale fiber micro-structure, material scientists often continuously cross-section the material sample for a sequence of high-resolution 2D microscopic images. Then they manually annotate the large-scale fibers on each 2D slice and match the annotated fibers through the 2D image sequence for reconstructing the 3D fiber structures. However, this process is usually very time consuming and requires superior quality equipments. First, to facilitate the inter-slice fiber matching and 3D reconstruction, extremely dense cross-sectioning, i.e., sampling with very small inter-slice distance, is usually required and this may take a long time even for small-size material samples. Second, dense sampling leads to a longer image sequence, which also increases the load and time of manual annotation. In practice, it may take material scientists several weeks or months to obtain the fiber micro-structure of a $1cm^3$ material sample, not to mention that the manual annotation of large-scale fibers is very tedious and prone to error.

To quickly and accurately characterize the fiber micro-structures of larger-size material samples, this paper proposes sparse sampling in cross-sectioning and then develops a new algorithm to automatically track large-scale fibers through the obtained sparsely sampled image sequence. This can be formulated as a multi-target tracking problem, where the main challenge lies in the step of association: large-scale of crowded fibers usually show very similar appearance in a 2D slice (as shown in the leftmost panel of Fig. 2 [3], making it difficult to associate each fiber to its corresponding observation when tracking moves into a new slice. This issue gets much worse with the increase of inter-slice distance in sparse sampling due to the possible larger gap between the prediction and observation of each fiber.

In materials science, large-scale 3D fibers are always implanted in bundles and the fibers in the same bundle are usually highly correlated by showing good proximity and parallelism. This paper explores such fiber correlation information to facilitate the large-scale fiber association over a large inter-slice distance. Unfortunately, the composition of fiber bundles is not priorly known. This paper develops a new group-wise approach that can simultaneously explore the fiber-bundle composition and perform accurate fiber association and tracking through sparsely sampled image sequences.

Manuscript received March 25, 2016; revised July 18, 2016; accepted July 19, 2016. Date of publication August 8, 2016; date of current version August 29, 2016. This work was supported in part by AFOSR under Grant FA9550-11-1-0327 and in part by UES Inc./AFRL under Grant S-901-486-002. This paper was presented at the IEEE Conference on Computer Vision and Pattern Recognition, 2016. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Martin Kleinsteuber. *(Youjie Zhou and Hongkai Yu contributed equally to this work.)*

Y. Zhou, H. Yu, and S. Wang are with the Department of Computer Science& Engineering, University of South Carolina, Columbia, SC 29208 USA (e-mail: zhou42@email.sc.edu; yu55@email.sc.edu; songwang@cec.sc.edu).

J. Simmons and C. P. Przybyla are with the Materials and Manufacturing Directorate, Air Force Research Laboratory, Dayton, OH 45433 USA (e-mail: jeff.simmons.3@us.af.mil; craig.przybyla@us.af.mil).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TIP.2016.2598640

Specifically, this paper uses Kalman filter to track each fiber through the image sequence for its location and velocity (i.e., location change) from slice to slice. For group-wise association, the nonrigid Thin-Plate Spline (TPS) is used to model the mapping between predictions and observations of all the fibers within a same bundle. To identify the correlated fibers, a new algorithm is developed by dividing fibers into groups, followed by three steps of refinement: group shrinking, group growing and group merging. The proposed association considers the possible false positives and false negatives in the fiber observations, i.e., the numbers of fiber predictions and observations may be different when the tracking moves into a new slice.

To verify the performance of the proposed algorithm, it is tested on three tiles of 100-slice S200 material samples and the fiber tracking performance is evaluated using 1,136 human annotated ground-truth fiber tracks. In particular, the fiber tracking performance is evaluated by increasing the sampling sparsity, i.e., inter-slice distance, in cross sectioning. The experiment results show that the proposed algorithm outperforms several other Kalman-filter based baseline algorithms and state-of-the-art multi-target tracking algorithms.

The remainder of the paper is organized as follows. Section II briefly reviews the related work. Section III describes the proposed large-scale fiber tracking method. Quantitative and qualitative evaluation results and discussions are presented in Section IV, followed by brief conclusions in Section V.

II. Related Work

Existing multi-target tracking algorithms are mainly developed for tracking multiple persons or vehicles from videos, or multiple cells from biomedical image sequences. They can be generally categorized as recursive and non-recursive methods.

Recursive tracking methods estimate the state of the target in a new slice (frame for video tracking) only using the information from previous slices that have been processed. Typical recursive tracking methods include the classical Kalman filter [4]–[6], Particle filter [7], [8] and non-parametric mixture Particle filters [9]. When tracking moves to a new slice, these recursive methods first make a prediction of the state for each target, such as target location and velocity, using the information from previous slices. They then detect targets in this new slice as observations, followed by an association step that correspond the predictions and observations of multiple targets. Finally, the posterior distribution of each tracker is corrected using the associated observation. These three steps of prediction, association, and correction are recursively performed on each slice along the image sequence. The recursive methods are particularly applicable to online tracking tasks.

Non-recursive tracking methods assume the availability of the whole image sequence before tracking multiple targets over this sequence. In these methods, observations of multiple targets are first detected on all the slices and then linked across slices along the image sequence for the final tracks. Graph model and algorithms are usually employed for non-recursive tracking – each graph node represents an observation, each graph edge indicates a possible linking of the observations across two slices, and the tracking through the sequence can be reduced to find optimal paths in the constructed graph, given appropriately defined cost functions, such as maximum a posteriori (MAP) [10]–[13]. In [14], motion dynamics similarity is incorporated into the cost function, resulting in a non-recursive SMOT tracking method for multi-target video tracking. In [15], a KTH tracking method is developed by searching shortest paths in the constructed graph model and this method was successfully used to track living cells through microscopy biomedical image sequences. In [16], a CEM tracking method is proposed by optimizing a continuous cost function that considers the detection, appearance, motion priors, and physical constraints of the targets.

As mentioned in Section I, one important component in the multi-target tracking methods is the association. For recursive methods, association is usually formulated as an explicit step which finds the correspondence between the predictions and the observations. For non-recursive methods, the association is implied in the cost function and the optimization algorithm – the extracted paths find the associations between the observations across neighboring slices. In many multi-target tracking tasks, such as human/vehicle tracking, where above-mentioned recursive and non-recursive methods have applied, the targets to be tracked are of small number, and they are spatially scattered or in different appearance [10]–[14], [16]. Moreover, the sampling along the image sequence, such as videos, can be very dense, then the predictions or the observations to be associated between neighboring slices show high-level of spatial continuity. In such cases, a simple nearest neighbor technique or an appearance-based matching algorithm may be sufficient for association of multiple targets.

Differently, the goal of this work is to track large-scale (hundreds or thousands of) fibers through sparsely sampled image sequences. In addition, these fibers are crowdedly distributed and show very similar appearance and shape in each slice (See Fig. 2 for an example). As a result, the association becomes very challenging and many of the above mentioned existing methods, neither recursive nor non-recursive methods, would have difficulty in accurately tracking these fibers. This motivates us to develop a new group-wise association algorithm that leverages the fiber bundles for fiber tracking. In particular, the framework of the classical Kalman filter is chosen for developing the new association algorithm and the entire tracking method based on three considerations. First, the Kalman filter algorithm is simple, well established and very efficient. Second, as a recursive method, it supports the online tracking and provides slice-by-slice predictions, which can potentially be used to adaptively select the serial-sectioning distance in imaging. Third, most non-recurisve methods optimize global cost functions involving the entire image sequence and further consideration of the fiber bundle information will lead to overly complex cost functions that cannot be efficiently optimized.

III. Proposed Method

As illustrated in Fig. 1, the proposed method uses Kalman filter to track each fiber, by recursively performing prediction,

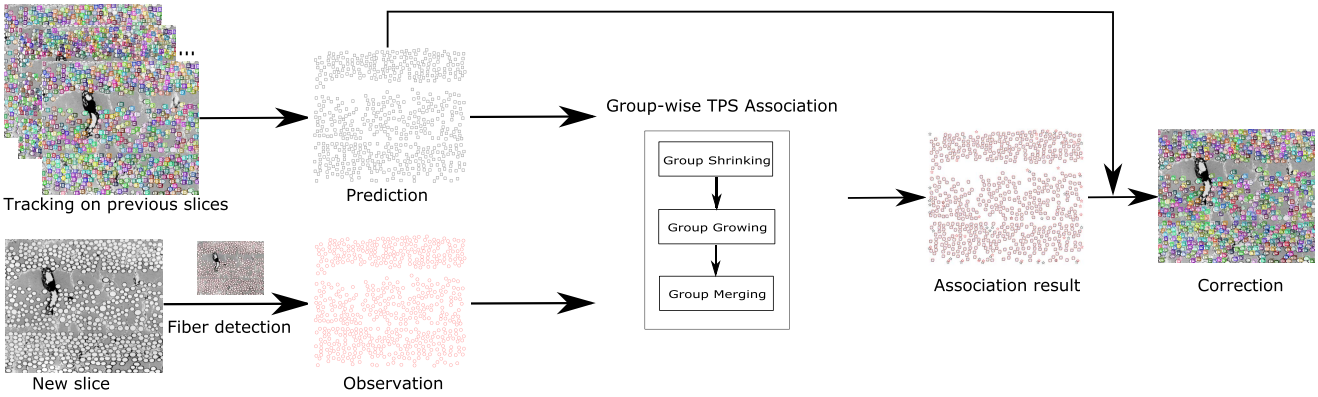


Fig. 1. The pipeline of the proposed large-scale fiber tracking based on Kalman filter, which sequentially performs prediction, association and correction when moving into a new slice.

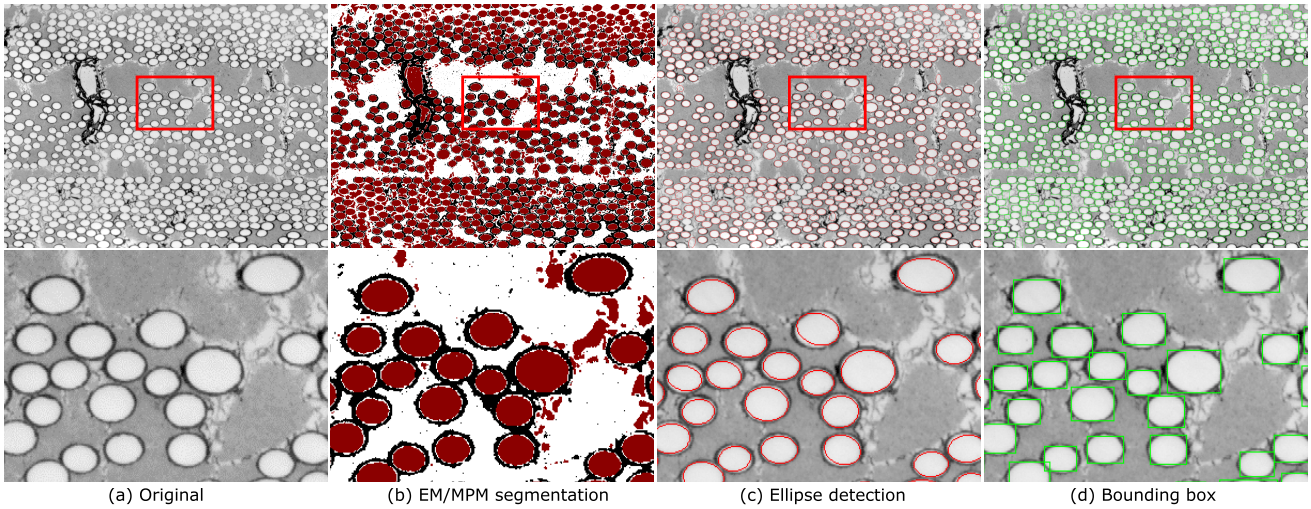


Fig. 2. An illustration of fiber detection. Second row shows the zoomed images of the cropped regions (red box) in the first row.

association and correction along the image sequence. The key contribution of this paper is the development of a new group-wise algorithm for association, which enables the developed method to track through sparsely sampled image sequences.

A. 2D Fiber Detection

When Kalman filter based tracking moves into a new slice, the first step is to detect fibers on this slice as the observations. Different from multi-target tracking in the video surveillance where each target may show different appearance, large-scale fibers usually show very similar appearance in 2D microscopic image slices. As shown in Fig. 1, most of the fibers are of an ellipse shape in the 2D slices and we can use an ellipse detection algorithms to detect them.

To detect fibers, the EM/MPM algorithm [17] is first applied to segment the image slice into fiber (foreground) and non-fiber (background) regions. As shown in Fig. 2(b), three segments are obtained using EM/MPM, where the red segment indicates the fiber region and the black/white segments indicate the non-fiber regions, including void, coating and SiC matrix. EM/MPM is a Markov Random Field (MRF) based pixel labeling technique that minimizes the expected number of mis-segmented pixels. The pixel labeling problem is formularized as the maximization of the posterior marginal (MPM) problem, and model parameters are estimated by the EM algorithm. EM/MPM has been found to be very effective in segmenting various materials-science images [17], [18]. From the fiber region, a set of connected components can be extracted and each of them is treated as a candidate of a fiber.

Sobel operator is then applied to detect the boundary of each connected component, which is then fitted by an ellipse using a set of ellipse geometry constraints [19]. The locations (the center coordinates) of the fitted ellipses are taken as the fiber observations on this slice and used for the proposed association and tracking. In addition to the locations, the tight bounding box around each fitted ellipse are also recorded, as shown in Fig. 2(d), which are used by several previous state-of-the-art tracking methods that are chosen for performance comparison in the later experiments.

Note that fiber detection is not perfect. Given the image noise and blurs, both false positives and false negatives may occur in the fiber detection, as indicated in Fig. 2(c). In addition, 3D fiber may be cropped by the image boundary when moving from one slice to another. They make the fiber association not an exact one-on-one mapping – a fiber tracked from previous slices may not have an associated observation when moving to a new slice and vice versa. Clearly these complications further increase the difficulty of association.

B. Fiber Tracking Using Kalman Filter

In this paper, for each fiber, a Kalman filter is applied to track it. In such a Kalman filter, state $\mathbf{s} = \left(x, y, v_x, v_y\right)^T$ is defined to describe the corresponding fiber in 2D slices, where $\mathbf{z} = (x, y)^T$ is the fiber location (e.g., ellipse center) and $\left(v_x, v_y\right)$ is the fiber velocity (e.g., fiber location change between neighboring slices), in horizontal and vertical directions respectively. The state is evolved linearly from slice to slice based on a predefined model. In this paper, fibers are highly smooth in 3D space and a constant velocity is assumed for each fiber. This way, the state transition from slice $i-1$ to slice i is modeled as

$$\mathbf{s}^i = A\mathbf{s}^{i-1} + \mathbf{w}^{i-1} \tag{1}$$

where the transition matrix

$$A = \begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

and $\mathbf{w} \sim \mathcal{N}(\mathbf{0}, Q)$ is the Gaussian noise for state transition. The observation model is defined by

$$\mathbf{o}^i = H\mathbf{s}^i + \mathbf{r}^i \tag{2}$$

where the observation matrix

$$H = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix},$$

the observation $\mathbf{o}^i = [x_{\mathbf{o}}, y_{\mathbf{o}}]^T$ is the location of the fiber detected on slice i, and $\mathbf{r} \sim \mathcal{N}(\mathbf{0}, R)$ is the Gaussian noise for observation.

The step of prediction makes a prior state estimate of the fiber on a new slice i as

$$\hat{\mathbf{s}}^i = A\mathbf{s}^{i-1} \tag{3}$$

and calculates the prior estimate error covariance as

$$\hat{P}^i = AP^{i-1}A^T + Q, \tag{4}$$

where $\mathbf{s}^{i-1}$ and P^{i-1} are the posterior state estimate and the posterior estimate error covariance at slice $i-1$.

The step of correction considers the observation $\mathbf{o}^i$ on the slice i and computes the posterior state estimate and the posterior estimate error covariance at slice i by

$$\mathbf{s}^i = \hat{\mathbf{s}}^i + K^i(\mathbf{o}^i - H\hat{\mathbf{s}}^i) \tag{5}$$

and

$$P^i = (I - K^i H)\hat{P}^i, \tag{6}$$

where I is an identity matrix and K^i is the Kalman gain at slice i:

$$K^i = \hat{P}^i H^T \left(H\hat{P}^i H^T + R\right)^{-1}. \tag{7}$$

However, there are large-scale crowded fibers to track and before running the step of correction in slice i, the correct observation must be found for each fiber, from the large set of observations detected in Section III-A.

C. Thin-Plate Spline Robust Point Matching (TPS-RPM) [20]

This section briefly reviews the Thin-Plate Spline Robust Point Matching (TPS-RPM) algorithm [20], which will be used to develop the proposed group-wise association algorithm. TPS-RPM can robustly match two sets of 2D points by exploring the correlations among these points. Specifically, let $U = \left\{\mathbf{u}_p\right\}_{p=1}^N$ and $V = \left\{\mathbf{v}_q\right\}_{q=1}^M$ be two sets of 2D points, i.e., $\mathbf{u}_p = (u_{px}, u_{py})$, $p = 1, 2, \cdots, N$ and $\mathbf{v}_q = (v_{qx}, v_{qy})$, $q = 1, 2, \cdots, M$. The matching between these two sets of points is represented by a matrix $\mathcal{H} = \left[h_{p,q}\right]_{N\times M}$, where $h_{p,q} = [0, 1]$ indicates the probability of matching $\mathbf{u}_p$ and $\mathbf{v}_q$. TPS-RPM can jointly determine a non-rigid 2D transform $\mathbf{f} = (f_x, f_y) : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ and the matrix $\mathcal{H}$ to minimize a cost function

$$\begin{aligned} &E_{TPS-RPM}(\mathcal{H}, \mathbf{f}) \\ &\quad = \sum_{p=1}^{N}\sum_{q=1}^{M} h_{pq} \parallel \mathbf{f}\left(\mathbf{u}_p\right) - \mathbf{v}_q \parallel^2 \\ &\qquad + \alpha\phi(\mathbf{f}) + \beta\sum_{p=1}^{N}\sum_{q=1}^{M} h_{pq}\log h_{pq} - \gamma\sum_{p=1}^{N}\sum_{q=1}^{M} h_{pq}, \end{aligned} \tag{8}$$

where $\phi(\mathbf{f}) = \iint \left[L(f_x) + L(f_y)\right] dxdy$ is the TPS bending energy [20], [21] with $L(\cdot) = \left(\frac{\partial^2}{\partial x^2}\right)^2 + 2\left(\frac{\partial^2}{\partial x \partial y}\right)^2 + \left(\frac{\partial^2}{\partial y^2}\right)^2$ and it reflects the smoothness of the 2D mapping $\mathbf{f}$ – the smaller the $\phi(\mathbf{f})$, the smoother the mapping $\mathbf{f}$. The last two terms in this cost function control the fuzziness of $\mathcal{H}$ and the percentage of points without matchings, respectively [20]. This cost function is then alternately minimized in terms of $\mathcal{H}$ and $\mathbf{f}$ respectively until convergence. Finally the obtained matrix $\mathcal{H}$ is thresholded to build the point matching between U and V. The TPS-RPM algorithm shows two useful properties:

1) All the matched points can be approximately described by a nonrigid TPS mapping $\mathbf{f}$, which is invariant to affine transforms, such as scaling, rotation and translation. By minimizing the TPS bending energy $\phi(\mathbf{f})$, TPS-RPM can identify the within-set point correlation, e.g., the subset of points that largely follow the same smooth TPS transform in the matching.
2) By introducing additional terms in the cost function, TPS-RPM can handle the noise and identify points without matchings.

In the following, a new group-wise fiber association algorithm is developed based on TPS-RPM.

D. Group-Wise TPS Association - Initialization

When tracking moves into a new slice i, there are a set of N fiber predictions

$$\left\{\hat{\mathbf{s}}_p^i\right\}_{p=1}^N$$

derived from the previous slices and a set of M fiber observations

$$\left\{\mathbf{o}_q^i\right\}_{q=1}^M$$

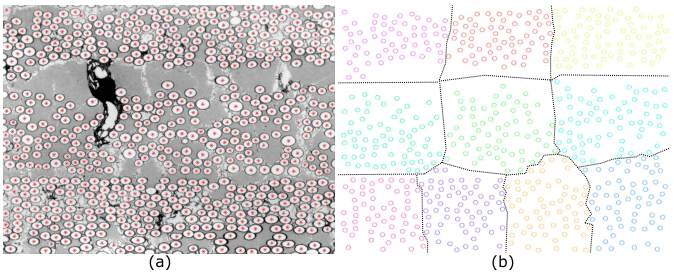


Fig. 3. An illustration of clustering predictions into a set of compact groups. (a) Predictions in red and (b) clustered groups in different colors, separated by dashed lines.

detected on the new slice. For simplicity, the subscript i is dropped and the predictions and observations are denoted as $\{\hat{\mathbf{s}}_p\}_{p=1}^{N}$ and $\{\mathbf{o}_q\}_{q=1}^{M}$, respectively, when it does not introduce ambiguity. The goal of association is to build a correspondence between them, i.e., for each prediction, find its corresponding observation. This is very similar to the point matching problem solved by TPS-RPM as described in Section III-C.

However, using a single global TPS transform $\mathbf{f}$ to describe the mapping between the predictions and the observations is problematic – not all the 3D fibers are highly correlated by showing good parallelism and the TPS bending energy for the true association between $\{\hat{\mathbf{s}}_p\}_{p=1}^{N}$ and $\{\mathbf{o}_q\}_{q=1}^{M}$ may be quite large, especially when the inter-slice distance is large in the sparse sampling. As a result, minimizing a single global TPS-RPM cost function may not produce the desired association.

To address this issue, this paper leverages the fact that in composite materials, fibers are usually implanted in bundles and the fibers within a bundle usually show good proximity and parallelism in 3D space. Therefore, even with highly sparse sampling in serial sectioning, a smooth TPS transform can well describe the fiber association within a bundle. If the bundle assignments are known for each prediction and observation, TPS-RPM can be used as described in Section III-C for finding the association in each bundle. However, bundle assignments are unknown priorly. In this section, a new approach is developed by simultaneously exploring the fiber-bundle composition and the fiber association.

Without knowing the bundle assignments of the fibers, all the predictions are first divided into smaller groups, as shown in Fig. 3. In this paper, this goal is achieved by applying the K-means clustering to the predictions $\{\hat{\mathbf{s}}_p\}_{p=1}^{N}$. While each fiber prediction $\hat{\mathbf{s}}_p$ is a 4D vector made up of a 2D location and a 2D velocity, clustering is only in terms of the locations $\{\hat{\mathbf{z}}_p = (\hat{x}_p, \hat{y}_p)\}_{p=1}^{N}$. As a result, the fibers in the same group are more likely to be from a same fiber bundle and TPS-RPM can then be used to match each group of the predictions to the observations.

However, the observations are not divided into corresponding groups as for the predictions. A sliding-window strategy is used to address this issue. As shown in Fig. 4, for each group of predictions, shown by red circles, its bounding box can be derived as shown by a blue box. Then this bounding box is dilated a little (5 pixels along each direction in our

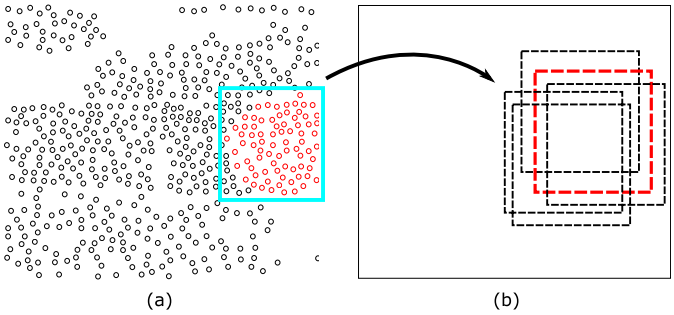


Fig. 4. An illustration of finding the initial matching for each group using a sliding-window strategy. (a) One group of predictions (in red) and (b) sliding windows across the slice with observations: the optimal window with the best matching is shown in red.

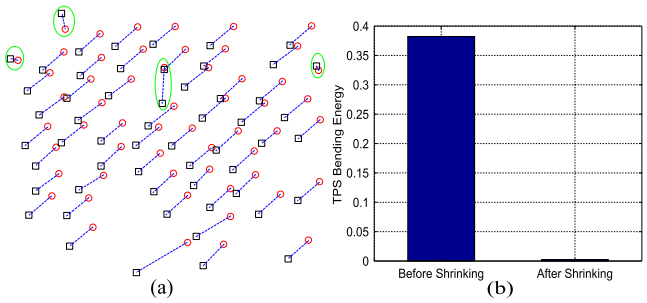


Fig. 5. An illustration of the group shrinking. (a) Initial matching of one group of the predictions (black boxes) and observations (red circles), with matching pairs linked by dashed lines. Outlier matchings that are removed in group shrinking are highlighted in green vertical ellipses. (b) TPS bending energies before and after removing the outlier matchings.

experiments) as the size of the sliding window and the sliding window is applied to the slice with the observations, shown by black boxes – TPS-RPM is performed between the group of prediction against the observations in each of the sliding window. Window sliding is performed around the center of the bounding box (shown as blue box), with range $\Delta x \in [-10, 10]$ pixels, $\Delta y \in [-10, 10]$ pixels and the step length of 10 pixels. Limiting the range of the sliding windows helps reduce the computational cost. The sliding window that leads to the minimum cost $E_{TPS-RPM}$ is taken as the optimal window, shown by the red box in Fig. 4, and the matched observations in this window are taken as the matching to the considered group of predictions. Note that each fiber prediction $\hat{\mathbf{s}}_p$ is a 4D vector made up of a 2D location and a 2D velocity and only the 2D location $\hat{\mathbf{z}}_p$ is used when applying the TPS-RPM against the observations, because the observations are 2D locations. The other steps of Kalman tracking, including the prediction and correction, still use the 4D state vector consisting of both the location and the velocity.

Applying such a sliding-window based matching for each group of predictions leads to its corresponding matched group of the observations and constructs an initial association between the predictions and the observations. However, K-means clustering cannot guarantee the predictions from a same group are all from a same bundle and the initial association from TPS-RPM in such a group may not be reliable. In the following, an algorithm is proposed to refine this initial association result.

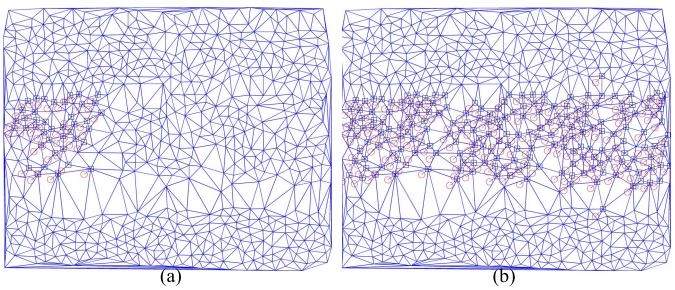


Fig. 6. An illustration of the group growing. (a) Delaunay triangulation of the predictions and a matching group before group growing. (b) The same matching group after the group growing. Matched pairs in the group are shown by the dashed-line linked boxes (predictions) and circles (observations).

E. Group-Wise TPS Association - Refinement

The proposed group refinement algorithm consists of three steps – *group shrinking*, *group growing* and *group merging*.

Group shrinking further removes the outlier matching in each group. Without loss of generality, let $\left(\hat{\mathbf{s}}_p, \mathbf{o}_p\right)$, $p = 1, 2, \cdots, m$ be one matched group of predictions and observations, as shown in Fig. 5(a). To identify and remove the outlier matching pairs from this group, the TPS bending energy is calculated as in Eq.(8) for this matching, in a matrix form [22]

$$\phi\left(\left(\hat{\mathbf{s}}_p, \mathbf{o}_p\right) | p = 1, 2, \cdots, m\right) = \frac{1}{8\pi}\left(\mathbf{x_o}^T \mathbf{L} \mathbf{x_o} + \mathbf{y_o}^T \mathbf{L} \mathbf{y_o}\right), \tag{9}$$

where $\mathbf{L}$ is the $m \times m$ upper-left block of the matrix

$$\begin{pmatrix} \mathbf{K} & \mathbf{P} \\ \mathbf{P}^T & \mathbf{0} \end{pmatrix}^{-1},$$

$\mathbf{K}$ is the $m \times m$ TPS kernel matrix with element $k_{pq} = k(\hat{\mathbf{z}}_p, \hat{\mathbf{z}}_q) = \|\hat{\mathbf{z}}_p - \hat{\mathbf{z}}_q\|^2 \log \|\hat{\mathbf{z}}_p - \hat{\mathbf{z}}_q\|$, $\mathbf{P} = \left(\mathbf{1}, \hat{\mathbf{x}}, \hat{\mathbf{y}}\right)$ with $\hat{\mathbf{x}}$ and $\hat{\mathbf{y}}$ being the concatenated vectors for all the x and y coordinates of the predictions $\left\{\hat{\mathbf{s}}_p\right\}_{p=1}^m$ (i.e., $\left\{\hat{\mathbf{z}}_p\right\}_{p=1}^m$) respectively, and $\mathbf{x_o}$ and $\mathbf{y_o}$ are the concatenated vectors for all the x and y coordinates of the observations $\left\{\mathbf{o}_p\right\}_{p=1}^m$, respectively. By calculating the leave-one-pair-out TPS bending energy $\phi_j = \phi\left(\left(\hat{\mathbf{s}}_p, \mathbf{o}_p\right) | p = 1, 2, \cdots, m; p \neq j\right)$, $j = 1, 2, \cdots, m$, the j^*-th pair that leads to the largest decrease of bending energy is removed, i.e., $j^* = \arg\min_j \phi_j$. This process is repeated until a specified percentage (δ) of pairs are removed from each group, as shown in Fig. 5. Note that, by pre-specifying the percentage δ, the step of group shrinking may remove true positive matchings from a group. This will be handled in the later steps of group growing and group merging.

Based on the group matching after the group shrinking, a TPS mapping can be constructed for the matching in each group. Without loss of generality, let $\left(\hat{\mathbf{s}}_q, \mathbf{o}_q\right)$, $q = 1, 2, \cdots, n$ be one matched group of predictions and observations after group shrinking. The TPS transform $\mathbf{f} = (f_x, f_y)$ such that $\mathbf{o}_q = \mathbf{f}(\hat{\mathbf{z}}_q)$, $q = 1, 2, \cdots, n$ is in the form of [22]

$$\begin{cases} f_x(\hat{\mathbf{z}}) = & a_1 + a_2\hat{x} + a_3\hat{y} + \sum_{q=1}^n c_q k(\hat{\mathbf{z}}, \hat{\mathbf{z}}_q) \\ f_y(\hat{\mathbf{z}}) = & b_1 + b_2\hat{x} + b_3\hat{y} + \sum_{q=1}^n d_q k(\hat{\mathbf{z}}, \hat{\mathbf{z}}_q), \end{cases} \tag{10}$$

where $\hat{\mathbf{z}} = (\hat{x}, \hat{y})^T$ is any location in the domain of prediction and $\hat{\mathbf{z}}_q$ is the location of the prediction $\hat{\mathbf{s}}_q$. The parameters $\mathbf{a} = (a_1, a_2, a_3)^T$, $\mathbf{b} = (b_1, b_2, b_3)^T$, $\mathbf{c} = (c_1, c_2, \cdots, c_n)^T$ and $\mathbf{d} = (d_1, d_2, \cdots, d_n)^T$ can be computed by

$$\begin{pmatrix} \mathbf{K} & \mathbf{P} \\ \mathbf{P}^T & \mathbf{0} \end{pmatrix} \begin{pmatrix} \mathbf{c} & \mathbf{d} \\ \mathbf{a} & \mathbf{b} \end{pmatrix} = \begin{pmatrix} \mathbf{x_o} & \mathbf{y_o} \\ \mathbf{0} & \mathbf{0} \end{pmatrix}$$

where $\mathbf{K}$, $\mathbf{P}$, $\mathbf{x_o}$, $\mathbf{y_o}$ and $k(\cdot, \cdot)$ are defined as in Eq.(9), but using the n matching pairs after the group shrinking.

In *group growing*, a Delaunay triangulation [23] is first constructed by taking all the predictions as the vertices, as shown in Fig. 6(a). Group growing is performed for each matching group (after the group shrinking) independently. Without loss of generality, let's consider a matching group $\left(\hat{\mathbf{s}}_q, \mathbf{o}_q\right)$, $q = 1, 2, \cdots, n$ shown by the dashed-line linked boxes (predictions) and circles (observations) in Fig. 6(a). The iterative growing of this group takes the following steps:

1) Label the predictions $\hat{\mathbf{s}}_q$, $q = 1, 2, \cdots, n$ as "processed" and all the other predictions as "unprocessed".
2) From all the "unprocessed" predictions that are adjacent to the matching group in the Delaunay triangulation graph, identify the nearest one and denote it as $\hat{\mathbf{s}}_a$ with the location $\hat{\mathbf{z}}_a$. If all predictions adjacent to the matching group have been "processed", exit and return the matching group as the group growing result.
3) Apply the TPS transform $\mathbf{f}$ computed in Eq.(10) to $\hat{\mathbf{z}}_a$ and search for the observation $\mathbf{o}_a$ that is nearest to $\mathbf{f}(\hat{\mathbf{z}}_a)$.
4) Check the consistency of the pair $(\hat{\mathbf{s}}_a, \mathbf{o}_a)$ against the matching group. If the consistency conditions are satisfied, the matching group is updated by including the pair $(\hat{\mathbf{s}}_a, \mathbf{o}_a)$. Relabel the predictions in the updated matching group as "processed" and all the other predictions as "unprocessed". Recalculate the TPS transform $\mathbf{f}$ using the updated matching group and go back to Step 2). If any consistency condition is not satisfied, simply label $\hat{\mathbf{s}}_a$ as "processed" and go back to Step 2).

Figure 6(b) shows a group-growing result, starting from the matching group given in Fig. 6(a).

In this paper, two consistency conditions are defined between a pair $(\hat{\mathbf{s}}_a, \mathbf{o}_a)$ and a matching group $\left(\hat{\mathbf{s}}_q, \mathbf{o}_q\right)$, $q = 1, 2, \cdots, n$. First, the distribution of the prediction-observation gap $(\mathbf{o}_q - \hat{\mathbf{z}}_q)$, $q = 1, 2, \cdots, n$ is computed in the matching group and this paper examines whether the gap $(\mathbf{o}_a - \hat{\mathbf{z}}_a)$ shows high likelihood in this distribution. More specifically, the gap is a 2D vector and two Gaussian distributions for the magnitude and slope angle are estimated, respectively. The first consistency condition is that the gap $(\mathbf{o}_a - \hat{\mathbf{z}}_a)$ falls in L_T times the standard deviations in both magnitude and slope angle distributions. Second, adding a new pair to a matching group may increase the TPS bending energy for the matching group. The small bending-energy increase, i.e., $\phi\left(\left(\hat{\mathbf{s}}_q, \mathbf{o}_q\right) | q = a, 1, 2, \cdots, n\right) - \phi\left(\left(\hat{\mathbf{s}}_q, \mathbf{o}_q\right) | q = 1, 2, \cdots, n\right) \leq \Delta_\phi$, is taken as the other consistency condition. If both consistency conditions are satisfied, the matching group is updated by including the new pair as stated in Step 4). Note that Gaussian distributions used in consistency conditions are also updated when the matching group is updated in the group growing.

After applying the group growing independently to all the matching groups, one prediction may be matched to different

Algorithm 1 Group-Wise TPS Association Algorithm

Init	$\{\hat{\mathbf{s}}_p^i\}_{p=1}^N$: N fiber predictions on slice i
	$\{\mathbf{o}_q^i\}_{q=1}^M$: M fiber observations slice i
1	Divide predictions to groups using K-means
2	**FOR** each group
3	TPS-RPM for initial association
4	Group shrinking to remove outlier matchings
5	Group growing to include more matching pairs
6	**END FOR**
7	Group merging for final association

observations in different matching groups and vice versa. A *group merging* is performed for the final association by two rounds of majority voting. In the first round, for each prediction $\hat{\mathbf{s}}_p$, the number of the votes an observation $\mathbf{o}_q$ receives is the number of matching groups that contain the pair $(\hat{\mathbf{s}}_p, \mathbf{o}_q)$ after group growing. The observation with the largest number of votes is matched to $\hat{\mathbf{s}}_p$. In the second round, for each observation $\mathbf{o}_q$, similar voting is performed for its corresponding prediction by only considering the matching pairs that are kept after the first round of voting. After these two rounds of voting, the resulting matching pairs are guaranteed to be one-on-one: No two observations are matched to a same prediction and vice versa. These final matching pairs are taken as the final association.

The whole group-wise TPS association algorithm is summarized in Algorithm 1.

IV. Experiments

The proposed method is tested on a set of material image sequences. Specifically, these images are collected in Air Force Research Laboratories (AFRL) using the RoboMet.3D automated serial sectioning instrument [24]. The tested material is S200, which is an amorphous SiNC matrix reinforced by continuous Nicalon fibers. 100 serial sections are produced with the dense inter-slice distance $1\mu m$. Because each serial section of the tested S200 sample is too large to be covered by a single shot of the microscopic image, it is divided into many tiles. In the experiments, the performance is tested on three independent tiles, which are named Data 1, Data 2 and Data 3, respectively. Each tile is an image sequence consisting of 100 slices and the resolution of each slice is 1292×968. An sample slice of a tile is shown in Fig. 2, which contains hundreds of crowded fibers.

A. Performance Evaluation Metrics

For objective and quantitative performance evaluation, fibers are manually annotated on these three test image sequences as the ground truth. As in many multi-target tracking tasks, it is very challenging to annotate all the fibers, which include the annotation of the fiber centers on each 2D slice and the linking of annotated 2D fiber centers between slices. In particular, for each tile not all the fibers are present in the whole sequence of 100 slices: some fibers may extend out of the perimeter of the tile when moving from one slice to another. In this paper, for each tile, the best efforts have been made to manually annotate as many 3D fibers as possible that extend through the whole image sequence. In total, 1,136 such fibers (393 in Data 1, 380 in Data 2 and 363 in Data 3) are annotated as the ground truth.[1]

In our experiments the fiber tracking performance is quantized by five widely used metrics [16], [25]: Multiple Object Tracking Accuracy (MOTA), Multiple Object Tracking Precision (MOTP), Identity Switches (IDSW), Mostly Tracked (MT) and Mostly Lost (ML), all of which measure the co-alignment between the tracked fibers and the annotated ground-truth fibers, but from different perspectives. In computing these metrics, a threshold of 20 pixels is used between the tracked fiber and the ground-truth fiber on each slice to count the hit/miss on the slice. MT is the number of ground-truth fibers that are hit in no less than 80% of slices while ML is the number of ground-truth boundary that are hit in no more than 20% of slices. For MOTA and MT, higher scores indicate better tracking, but for MOTP, IDSW and ML, lower scores indicate better tracking. Note that, for MOTP metric, there are two different definitions for evaluating the multiple target tracking. One of them applies to the tracking results in the form of the bounding boxes around the target. In this case [16], MOTP reflects the bounding-box overlap between the tracking results and the ground truth, in which higher MOTP denotes better tracking performance. In the other definition, tracking results are in the form of a sequence of target locations through the image sequence. In this case [25], MOTP reflects the distance between the tracking results and the ground truth, in which lower MOTP means better tracking performance. In this paper, the latter definition of MOTP is used.

Because we are not able to annotate all the fibers in each tile in building the ground truth, the number of tracked fibers is usually more than the number of ground-truth fibers. By directly comparing with the ground truth fibers, many tracked fibers may be mistakingly counted as false positives and the resulting evaluation metrics may undervalue the real performance of the fiber tracking. To address this problem, the tracked fibers that are far away from the ground truth fibers are pruned and the remaining tracked fibers are used for computing the above five metrics against the ground-truth fibers. Specifically, a tracked fiber $\{\mathbf{z}^i\}_{i=1}^I$, where $\mathbf{z}^i$ is the tracked-fiber centers at slice i, is pruned if

$$\frac{\sum_{i=1}^{I} \mathbb{I}(\| \mathbf{z}^i - \mathbf{z}_g^i \|_2 < t_d)}{I} < t_o \tag{11}$$

where $\mathbf{z}_g^i$ is the closest annotated fiber center (ground truth) to $\mathbf{z}^i$ at slice i, and $\mathbb{I}$ is an indicator function which equals to 1 if the condition is satisfied and 0 otherwise. The distance threshold t_d is set to 20 pixels and the overlapping threshold t_o is set to 50% for the experiments. Note that this is not a very strict condition – it does not overly prune the tracked fibers and reduce much the recall. Specifically, in all the experiments, after applying this step of pruning, the remaining number of tracked fibers is usually similar to the number of the annotated ground-truth fibers.

[1] Our publicized dataset and code:http://cvl.cse.sc.edu/project/cvpr2016.html.

To test the tracking performance under sparsely sampled image sequences, the original image sequence (slices $1, 2, \cdots, 100$ with the dense inter-slice distance $1\mu m$) is downsampled, by repeatedly skipping $C \geq 0$ slices before taking the next slice in the original sequence, until the end of original sequence is reached. For convenience, C is called the *sparsity* in this paper:

Definition 1: Sparsity C is the number of skipped slices between two consecutively selected slices in downsampling.

For the data tested in this paper, such a constructed image sequence with sparsity C actually has an inter-slice distance of $(C+1)\mu m$. For example, a sparsely sampled image sequence consisting of the original slices $1, 6, 11, \cdots, 96$ has a sparsity $C = 4$. Please note that the sparsity C defined in this paper describes the inter-slice distance along the cross-sectioning direction. It is not related to the resolution of each 2D slice and the density of the fiber distribution in the 2D or 3D spaces.

The larger the parameter C, the sparser the constructed image sequence. One issue is that, such constructed image sequences with large C are much shorter than the original image sequence and the tracking performance obtained on a single such short sequence may not be statistically reliable. To alleviate this issue, for a given sparsity C, a set of $C+1$ image sequences are constructed, starting from original slice $1, 2, \cdots, C+1$ respectively. These $C+1$ image sequences do not share any slice. Tracking is performed on each of them independently and then the average of their performances, e.g., MOTA and MOTP, is reported as the performance of tracking under the sparsity C. Note that when $C = 0$, tracking is directly performed and evaluated on a single image sequence: the original densely sampled image sequence without any downsampling. In our experiments, the sparsity C is continuously varied from 0 to 19 and the tracking performance is examined under different sparsity.

B. Comparison Methods

To justify the effectiveness of the proposed method, its performance is compared against three baseline Kalman filter methods and four other state-of-the-art multi-target tracking methods.

Kalman-NN is a baseline Kalman-filter tracking method where the association is computed by repeating nearest neighbor search in a greedy fashion. First, the pair of the prediction and observation with the minimum L2 distance is identified and associated. Then we exclude this identified pair and repeat the same nearest neighbor search on the remaining predictions and observations, until either predictions or observations are empty.

Kalman-Hung is a baseline Kalman-filter tracking method where the association is computed by Hungarian algorithm [26] for a minimum-total-distance bipartite matching and the pairwise distance between predictions and observations are their L2 distance. Because the number of predictions and observations are usually different, dummy nodes are introduced into Hungarian algorithm. The distance to a dummy node is set to be the maximum L2 distance of our tolerated matchings between predictions and observations on the considered slice (40 pixels used in our experiments).

Kalman-Global is another baseline Kalman-filter tracking method where the association is computed by directly applying the TPS-RPM [20] to match the predictions and observations in a global fashion.

The four other state-of-the-art multi-target tracking algorithms used for comparison are DPNMS [13], SMOT [14], CEM [16] and KTH [15]. *DPNMS* employs a min-cost flow model [11] with specific constraints and cost functions to handle target occlusions in tracking. *SMOT* considers the motion dynamics to handle similar-appearance targets. *CEM* uses a new continuous energy function that integrates the observation, appearance, motions and physical constraints for multi-target tracking. *KTH* considers the track splitting and merging in multi-target tracking based on a graph model. Different from the Kalman-filter tracking methods, these four comparison methods are all non-recursive methods.

The proposed tracking method, abbreviated as *Kalman-Groupwise* from now on, and the three baseline Kalman filter tracking methods, i.e., Kalman-NN, Kalman-Hung and Kalman-Global, all use the Kalman filter as defined in Section III-B and the observations are the ellipse centers detected on each slice as described in Section III-A. These four Kalman-filter tracking methods are implemented in Matlab. The initial state covariance P^0 is set to be a diagonal matrix with diagonal elements 10^3. The transition noise covariance Q is set to be a diagonal matrix with diagonal elements 10^{-3}. The observation noise covariance is set to be a diagonal matrix with diagonal elements 10^{-3}. In the proposed Kalman-Groupwise, predictions are always clustered to 10 groups and the percentage of removed fiber pairs in group shrinking is set to $\delta = 30\%$. The consistency thresholds in group growing are set to $L_T = 3$ and $\Delta_\phi = 0.01$.

Publicly available codes downloaded from their respective authors' websites are used for DPNMS,[2] SMOT,[3] CEM[4] and KTH.[5] For DPNMS, SMOT and CEM, the observations are the bounding boxes of the detected ellipses as described in Section III-A. For KTH, no source code is available and its binary executable software, which has its own integrated image segmentation and target detection components, is directly used. For these four comparison methods, their respective default parameters are used in the experiments.

C. Results

Figure 7 shows the MOTA and MOTP of the proposed Kalman-Groupwise method and the three baseline Kalman filter tracking methods on the three image sequences, under different sparsity C. It can be seen that, when the sparsity C is low, both the proposed method and the three baseline methods produce satisfactory fiber tracking, with very high MOTA and very low MOTP. With the increase of the sparsity, the performance of all these four methods drops. However, the proposed Kalman-Groupwise's performance, in terms of both MOTA and MOTP, drops much slower than the three baseline

[2] http://people.csail.mit.edu/hpirsiav/papers/tracking_cvpr11_release_v1.0.tar.gz

[3] https://bitbucket.org/cdicle/smot

[4] https://bitbucket.org/amilan/contracking

[5] http://www.codesolorzano.com/celltrackingchallenge/Cell_Tracking_Challenge/KTH-SE.html

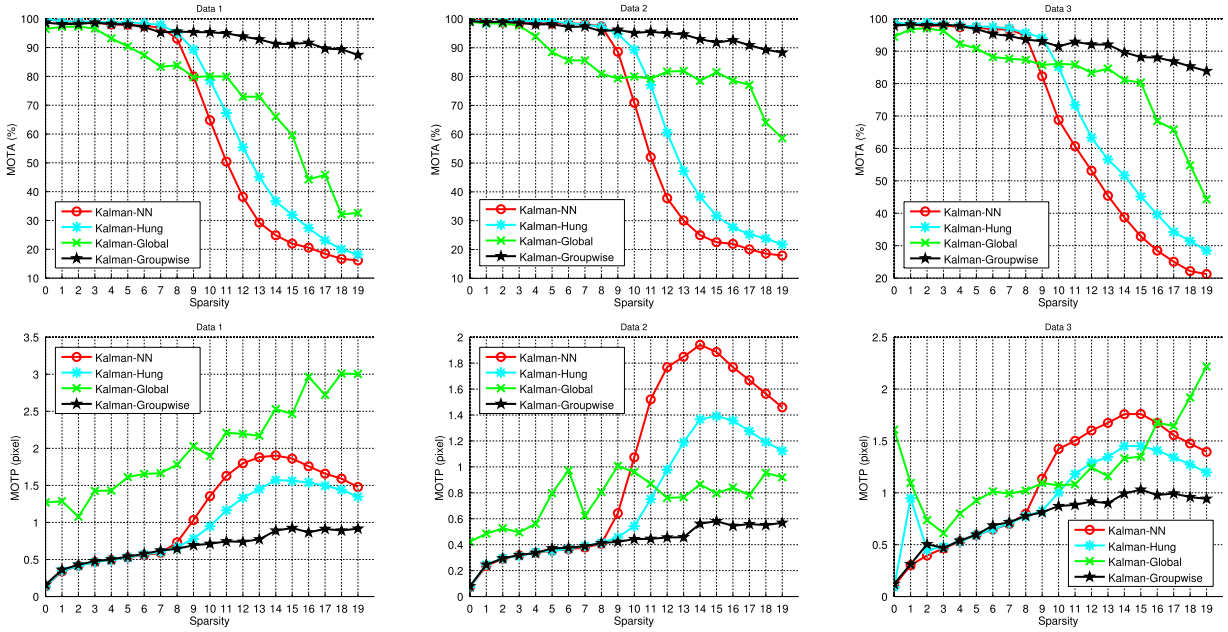


Fig. 7. MOTA and MOTP performance of the proposed Kalman-Groupwise method and the three baseline Kalman filter tracking methods: Kalman-NN, Kalman-Hung and Kalman-Global, under different sparsity C.

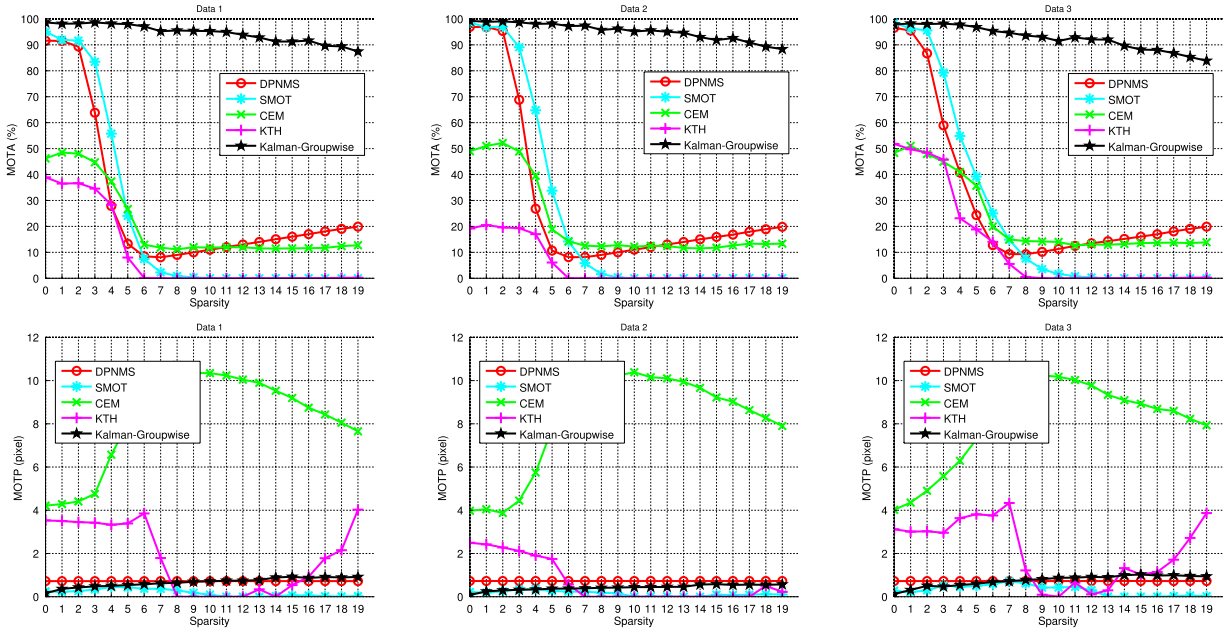


Fig. 8. MOTA and MOTP performance of the proposed Kalman-Groupwise method and the four non-recursive tracking methods: DPNMS, SMOT, CEM and KTH, under different sparsity C.

methods. Even if $C = 19$, i.e., increasing the inter-slice distance by 19 times, the proposed method can still achieve very high MOTA performance. For the poor performance of Kalman-Hung and Kalman-NN under high sparsity, the main reason is the large gap between a prediction and its corresponding observation. As a result, each prediction may not be corresponding to its nearest observation and vice versa. Table I shows the IDSW, MT and ML metrics of the proposed method and the three baseline Kalman-filter tracking methods under different sparsity. The performance shown in this table is the average over all three test image sequences. We can see that, all these four methods show increased IDSW and decreased MT with the increase of the sparsity. In general the proposed Kalman-Groupwise method achieves relatively lower IDSW and higher MT than the three baseline methods when the sparsity C increases. The ML metrics of the four methods are comparable to each other and they are quite low. These results show that the association step plays a key role in using Kalman filter for the proposed large-scale fiber tracking. The proposed group-wise TPS based algorithm can achieve much better association than the classical nearest neighbor and perfect matching algorithms.

Figure 8 shows the MOTA and MOTP of the proposed method and the four other state-of-the art non-recursive track-

TABLE I

IDSW, MT, AND ML PERFORMANCE OF THE PROPOSED KALMAN-GROUPWISE METHOD AND THE SEVEN COMPARISON METHODS UNDER DIFFERENT SPARSITY C. LAST ROW SHOWS THE FPS (FRAMES/SLICES PROCESSED PER SECOND) OF EACH METHOD. THE PERFORMANCE IS THE AVERAGE OVER ALL THREE IMAGE SEQUENCES.

Metrics		Kalman-NN	Kalman-Hung	Kalman-Global	DPNMS[13]	SMOT[14]	CEM[16]	KTH[15]	Kalman-Groupwise
IDSW	$C=0$	9.0	6.3	4.5	780.5	62.8	63.0	258.3	4.3
	$C=5$	6.3	3.3	41.0	5135.7	21.9	128.2	3.5	2.6
	$C=10$	596.6	209.9	118.6	2921.7	0.02	89.7	0	5.0
	$C=15$	1162.4	937.9	229.9	1871.0	0.7	118.2	0	21.3
	$C=19$	1100.5	999.1	453.0	1413.3	0	129.6	0.1	43.4
MT	$C=0$	376.3	377.0	366.0	370.5	363.0	113.8	84.0	376.3
	$C=5$	371.1	374.9	345.5	368.6	68.1	38.1	16.1	373.9
	$C=10$	309.2	337.6	313.0	360.1	0.7	2.3	0	364.6
	$C=15$	268.5	271.9	301.8	360.6	0	3.6	0	354.6
	$C=19$	307.5	297.5	280.2	362.0	0	5.2	0	347.5
ML	$C=0$	0.3	0.3	4.0	0.8	1.0	115.5	140.8	0.3
	$C=5$	1.6	1.8	8.5	0.8	214.4	166.0	324.6	2.7
	$C=10$	3.9	5.4	7.7	0.8	372.7	198.6	375.0	6.2
	$C=15$	8.8	11.2	11.4	1.6	374.3	231.0	375.0	12.4
	$C=19$	1.5	1.7	3.5	0.8	373.5	117.7	370.6	5.5
FPS		0.088	0.025	0.003	1.042	0.017	0.004	0.050	0.080

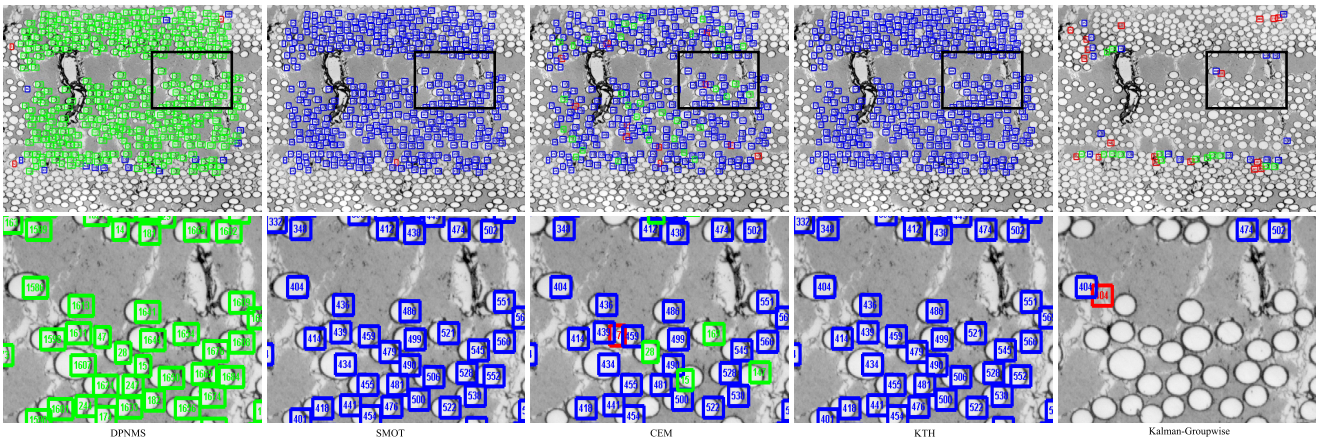


Fig. 9. An illustration of the tracking errors on one slice by using the proposed Kalman-Groupwise and four non-recursive methods, respectively, where the sparsity $C = 19$. False positives are in red, false negatives are in blue and ID switches are in green. Second row shows the zoomed images of the cropped regions (black box) in the first row.

ing methods: DPNMS, SMOT, CEM and KTH. All the four non-recursive comparison methods show low MOTA values when the sparsity increases. This is due to the large scale of the tracked fibers, their identical appearance and crowdedness, which break some of the assumptions made in these comparison methods. In terms of MOTP, the proposed Kalman-Groupwise, SMOT, and DPNMS achieve much lower MOTP values than CEM and KTH. Table I also shows the IDSW, MT and ML of SMOT, DPNMS, CEM and KTH under different sparsity. DPNMS generate very high IDSW errors, SMOT generates low MT errors and high ML errors, and CEM generates high IDSW and ML errors, while KTH generates low MT errors and high ML errors. In general, the proposed Kalman-Groupwise shows more competitive performance than these four comparison methods in the large-scale fiber tracking when the sparsity C is high.

Table I also shows the FPS numbers (frames/slices processed per second) of all the tracking methods. Higher FPS indicates lower computational cost. All the FPS numbers are obtained on a workstation with a 4-core 2.6GHz Intel CPU with 8GB memory. With an FPS of 0.080, the proposed method takes about $\frac{100}{0.08\times 60} = 21$ minutes to track all the fibers through a sequence of 100 slices. While this FPS number is not the best against other comparison methods as shown in Table I, the proposed method can substantially reduce the total time of data preparation. The most time-consuming step in data preparation is the mechanical cross sectioning – using RoboMet.3D [24], it takes about 15 minutes to grind for one slice. The proposed method can handle a sparsely sampled image sequence with larger inter-slice distance and therefore requires much fewer cross-sectioned slices. The time saved in cross sectioning is overwhelming to the time spent on fiber tracking.

Figure 9 shows the tracking errors of the proposed Kalman-Groupwise and the four non-recursive methods on one slice when the sparsity $C = 19$. It can be seen that DPNMS generates many ID switches (green boxes), SMOT and KTH generate many false negatives (blue boxes), and CEM generates both false positives (red boxes), false negatives and ID switches. On this slice, the proposed Kalman-Groupwise generates much less tracking errors of false positives, false negatives and ID switches than these four comparison methods.

To justify the steps of group shrinking and group growing in the proposed association algorithm, comparison experiments

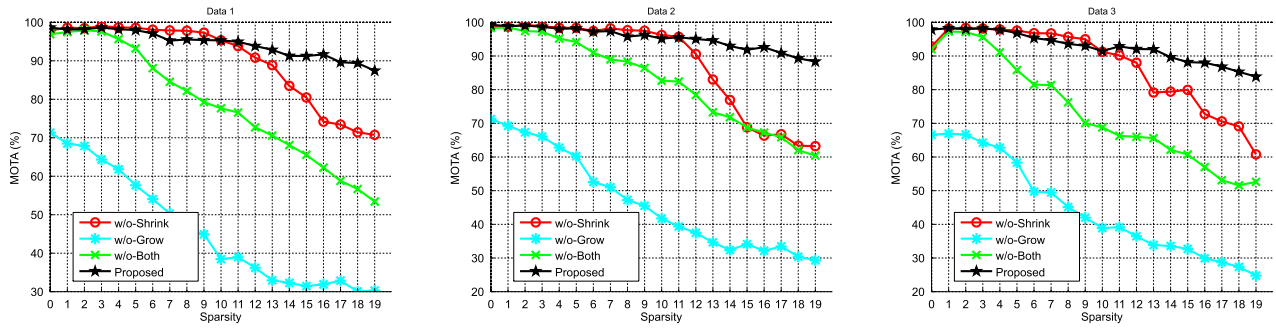


Fig. 10. MOTA performance of the proposed Kalman-Groupwise and its three variants by removing the step of group shrinking, the step of group growing and both of them, respectively.

are also conducted by using the proposed Kalman-Groupwise method (*Proposed*), but removing either or both of these two steps in the association. Results are shown in Fig. 10, where *w/o-Shrink, w/o-Grow,* and *w/o-Both* indicate three variants of the proposed method, by removing the step of shrinking, the step of growing and both of them, respectively. From these results, it can be seen that both the steps of group shrinking and group growing help increase the MOTA performance of the proposed method.

For the selection of group number in K-means clustering, different group numbers from 8 to 16 are tried on Data 2 with sparsity $C = 19$. The obtained mean MOTA is 85.6% with standard deviation 5.8%, indicating that the proposed method is not very sensitive to the choice of group number.

V. Conclusions

This paper proposed a Kalman-filter method for tracking large-scale fibers through microscopic image sequences. Our major contribution is the development of a new group-wise association algorithm that can match multiple predictions and observations even if the inter-slice distance is large and the fibers are crowded and share very similar appearance. The proposed algorithm employs the nonrigid thin-plate splines (TPS) to model the association of the fibers in a same fiber bundle. Without knowing the fiber bundle composition, all the fibers are first divided into a certain number of groups and a three-step algorithm was then developed for fiber association, consisting of group shrinking, group growing and group merging. Experiments were conducted on three real material image sequences. The tracking results show that the proposed method outperforms both the Kalman-filter methods using other classical association algorithms and four other state-of-the-art non-recursive multi-target tracking methods. The proposed method can be applied to speed up the imaging and image processing of the composite material images for fast fiber microstructure characterization.

Acknowledgment

The authors would like to thank Yuewei Lin, Xiaochuan Fan and Yang Mi for helpful discussions and data annotations. A preliminary version of this work has been published in a conference proceeding [27].

References

[1] M. P. Knox, "Continuous fiber reinforced thermoplastic composites in the automotive industry," in *Proc. Automotive Compos. Conf.*, 2001, pp. 1–5.

[2] S. Suresh, *Fundamentals of Metal-Matrix Composites*. Amsterdam, The Netherlands: Elsevier, 2013.

[3] C. Przybyla, T. Godar, J. Simmons, M. Jackson, L. Zawada, and J. Pearce, "Statistical characterization of SIC/SIC ceramic matrix composites at the filament scale with Bayesian segmentation Hough transform feature extraction, and pair correlation statistics," in *Proc. Int. SAMPE Tech. Conf.*, 2013, pp. 859–878.

[4] R. E. Kalman, "A new approach to linear filtering and prediction problems," *Trans. ASME, D, J. Basic Eng.*, vol. 82, pp. 35–45, 1960.

[5] J. Black, T. Ellis, and P. Rosin, "Multi view image surveillance and tracking," in *Proc. Workshop Motion Video Comput.*, 2002, pp. 169–174.

[6] D. B. Reid, "An algorithm for tracking multiple targets," *IEEE Trans. Autom. Control*, vol. 24, no. 6, pp. 843–854, Dec. 1979.

[7] M. D. Breitenstein, F. Reichlin, B. Leibe, E. Koller-Meier, and L. Van Gool, "Robust tracking-by-detection using a detector confidence particle filter," in *Proc. IEEE Int. Conf. Comput. Vis.*, Sep./Oct. 2009, pp. 1515–1522.

[8] K. Okuma, A. Taleghani, N. de Freitas, J. J. Little, and D. G. Lowe, "A boosted particle filter: Multitarget detection and tracking," in *Proc. Eur. Conf. Comput. Vis.*, 2004, pp. 28–39.

[9] J. Vermaak, A. Doucet, and P. Pérez, "Maintaining multimodality through mixture tracking," in *Proc. IEEE Int. Conf. Comput. Vis.*, Oct. 2003, pp. 1110–1116.

[10] H. Jiang, S. Fels, and J. J. Little, "A linear programming approach for multiple object tracking," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2007, pp. 1–8.

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

[12] J. Berclaz, F. Fleuret, and P. Fua, "Multiple object tracking using flow linear programming," in *Proc. IEEE Int. Workshop Perform. Eval. Tracking Surveill.*, Dec. 2009, pp. 1–8.

[13] H. Pirsiavash, D. Ramanan, and C. C. Fowlkes, "Globally-optimal greedy algorithms for tracking a variable number of objects," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2011, pp. 1201–1208.

[14] C. Dicle, O. I. Camps, and M. Sznaier, "The way they move: Tracking multiple targets with similar appearance," in *Proc. IEEE Int. Conf. Comput. Vis.*, Dec. 2013, pp. 2304–2311.

[15] K. E. G. Magnusson, J. Jaldén, P. M. Gilbert, and H. M. Blau, "Global linking of cell tracks using the Viterbi algorithm," *IEEE Trans. Med. Imag.*, vol. 34, no. 4, pp. 911–929, Apr. 2015.

[16] A. Milan, S. Roth, and K. Schindler, "Continuous energy minimization for multitarget tracking," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 36, no. 1, pp. 58–72, Jan. 2014.

[17] M. L. Comer and E. J. Delp, "The em/mpm algorithm for segmentation of textured images: Analysis and further experimental results," *IEEE Trans. Image Process.*, vol. 9, no. 10, pp. 1731–1744, Oct. 2000.

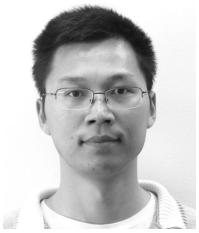
[18] J. Waggoner, Y. Zhou, J. Simmons, M. De Graef, and S. Wang, "3D materials image segmentation by 2D propagation: A graph-cut approach considering homomorphism," *IEEE Trans. Image Process.*, vol. 22, no. 12, pp. 5282–5293, Dec. 2013.

[19] Y. Xie and Q. Ji, "A new efficient ellipse detection method," in *Proc. Int. Conf. Pattern Recognit.*, 2002, pp. 957–960.

[20] H. Chui and A. Rangarajan, "A new point matching algorithm for non-rigid registration," *Comput. Vis. Image Understand.*, vol. 89, nos. 2–3, pp. 114–141, Feb. 2003.

[21] S. Belongie, J. Malik, and J. Puzicha, "Shape matching and object recognition using shape contexts," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 24, no. 4, pp. 509–522, Apr. 2002.

[22] F. L. Bookstein, "Principal warps: Thin-plate splines and the decomposition of deformations," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 11, no. 6, pp. 567–585, Jun. 1989.
[23] D.-T. Lee and B. J. Schachter, "Two algorithms for constructing a Delaunay triangulation," *Int. J. Comput. Inf. Sci.*, vol. 9, no. 3, pp. 219–242, 1980.
[24] *ROBO-MET.3D*, accessed on Jan. 8, 2015. [Online]. Available: http://www.ues.com/content/robomet3d
[25] K. Bernardin and R. Stiefelhagen, "Evaluating multiple object tracking performance: The CLEAR MOT metrics," *EURASIP J. Image Video Process.*, vol. 2008, Feb. 2008, Art. no. 246309.
[26] H. W. Kuhn, "The Hungarian method for the assignment problem," *Naval Res. Logistics Quart.*, vol. 2, nos. 1–2, pp. 83–97, Mar. 1955.
[27] H. Yu *et al.*, "Groupwise tracking of crowded similar-appearance targets from low-continuity image sequences," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, June 2016, pp. 952–960.



Youjie Zhou received the B.E. degree in software engineering from East China Normal University in 2010, and the Ph.D. degree in computer science and engineering from the University of South Carolina in 2015. He is currently a Software Engineer with Google. His research is focused on developing large-scale computer vision, image processing, and multimedia applications built upon modern machine learning and data mining techniques.

Hongkai Yu (S'16) received the B.S. and M.S. degrees from the Department of Automation, Chang'an University, Xi'an, China, in 2009 and 2012, respectively. He is currently pursuing the Ph.D. degree with the Department of Computer Science and Engineering, University of South Carolina. His current research interests include computer vision, machine learning, and intelligent transportation system.

Jeff Simmons received the B.S. degree in metallurgical engineering from the New Mexico Institute of Mining and Technology, Socorro, NM, USA, and the M.E. degree in metallurgical engineering and materials science and the Ph.D. degree in materials science and engineering from Carnegie Mellon University, Pittsburgh, PA, USA. After receiving the Ph.D. degree, he joined the Materials and Manufacturing Directorate, Wright Patterson SFB, as a Post-Doctoral Researcher and as a Research Contractor. In 1998, he joined the Air Force Research Laboratory as a Research Scientist. During his tenure, he has been involved in basic research as well as engineering applications in the aerospace industry. He is currently a Scientist with the Materials and Manufacturing Directorate, Air Force Research Laboratory. His research interests are in 3D materials science, particularly in the development of advanced algorithms for production and analysis of large image datasets. He has published and developed in both the materials science and signal processing fields on application of electronic imaging algorithms toward the analysis of digital microscopy. Other research interests have included mathematical representations of micro-structure and texture, physics-based modeling of micro-structure development, atomistic modeling of defect properties, and computational thermodynamics. Non-research experience has included leading teams to develop analytical tools for digital data, integration of computer resources for materials science simulations, computer security, and network design as well as manufacturing, particularly in machining applications. He has overseen execution of Air Force contracts on integrated computational materials science and engineering. He is a member of the TNS Committees of Advanced Characterization, Testing, and Simulations, Phase Transformations, and Integrated Computational Materials Science. He is a member of TS, ACM, and the IEEE Computer and Signal Processing Societies.

Craig P. Przybyla received the B.S. and M.S. degrees from the Department of Mechanical Engineering, Brigham Young University, in 2004 and 2005, respectively, and the Ph.D. degree from the Department of Materials Science and Engineering, Georgie Institute of Technology, in 2010. He is a Research Team Leader with the Materials and Manufacturing Directorate, Air Force Research Laboratory, Wright-Patterson Air Force Base, Ohio. Specifically, he is the Technical Lead of the Composites Performance Research Team in the Ceramics Branch whose mission is the performance characterization and prediction of advanced polymer and ceramic matrix composites for aerospace applications. He and his team support both composite material technology issues related to fielded systems and the transition of next generation composite technologies for future systems. Prior to his current position, he was a Senior Materials Engineer with the Composites Branch, where he was a Research Engineer and Subject Matter Expert for performance characterization and modeling of ceramic matrix composites for hot section components in gas turbine engines and/or thermal protection systems for hypersonic vehicles from 2010 to 2015.

Song Wang (SM'12) received the Ph.D. degree in electrical and computer engineering from the University of Illinois at Urbana–Champaign (UIUC), Urbana, IL, USA, in 2002. From 1998 to 2002, he was a Research Assistant with the Image Formation and Processing Group, Beckman Institute, UIUC. In 2002, he joined the Department of Computer Science and Engineering, University of South Carolina, Columbia, SC, USA, where he is currently a Professor. His current research interests include computer vision, medical image processing, and machine learning. He is currently serving as the Publicity/Web Portal Chair of the Technical Committee of Pattern Analysis and Machine Intelligence, the IEEE Computer Society, and an Associate Editor of *Pattern Recognition Letters*.