Robust Nanoparticles Detection From Noisy Background by Fusing Complementary Image Information

Yanjun Qian, Jianhua Z. Huang, Xiaodong Li, and Yu Ding, *Senior Member, IEEE*

Abstract—This paper studies the problem of detecting the presence of nanoparticles in noisy transmission electron microscopic (TEM) images and then fitting each nanoparticle with an elliptic shape model. In order to achieve robustness while handling low contrast and high noise in the TEM images, we propose an approach to fuse two kinds of complementary image information, namely, the pixel intensity and the gradient (the first derivative in intensity). Our approach entails two main steps: 1) the first step is to, after necessary pre-processing, employ both intensity-based information and gradient-based information to process the same TEM image and produce two independent sets of results and 2) the subsequent step is to formulate a binary integer programming (BIP) problem for conflict resolution among the two sets of results. Solving the BIP problem determines the final nanoparticle identification. We apply our method to a set of TEM images taken under different microscopic resolutions and noise levels. The empirical results show the merit of the proposed method. It can process a TEM image of 1024 × 1024 pixels in a few minutes, and the processed outcomes appear rather robust.

Index Terms—Binary integer programming, complementary information fusion, image segmentation, multi-expert system, nanoparticle analysis.

I. INTRODUCTION

A S MORE and more nanoparticle-embedded materials are moved from labs to commercial use, we witness an increasing need for automated nanoparticle detection and characterization based on the electron microscopic images of nanoparticles [1]–[3]. The images include those from both transmission electron microscope (TEM) and scanning electron microscope (SEM). Once the images are processed, material scientists would like to characterize the morphology of nanoparticles, or to quantify the dispersion of nanoparticles in the host material, as both traits are believed to have profound impact on the final material properties [4], [5]. To achieve these goals, the first job is to locate individual nanoparticles as accurately as possible, and then to characterize the shape and size of the nanoparticles. As such, automated detection and characterization of nanoparticles play important roles on nanomaterial exploration and production.

Park et al. [6] summarized the challenges associated with detection and characterization of nanoparticles from TEM images. The challenges lie in the facts that the nanoparticles are numerous and overlapped, and the variety of their shapes and sizes is also large. Park et al. [6] reviewed a number of image processing methods, including watershed transforms with different stopping criteria [7], [8], sliding band filter [9], graph cut [10], active contour [11], iterative voting [12], and a multiscale morphological method (a sophisticated variant of watershed) [13]. They argued that these methods can not be directly applied to the TEM images due to the technical challenges mentioned above. Park and his colleagues [6], [14] proposed image processing and shape analysis approaches, tailored to nanoparticle image processing. There are also some recent developments on detecting and measuring nanoparticles in TEM images. Yang and Ahuja [15] proposed a segmentation method to isolate the granular objects using a local density clustering and gradient barrier watershed. De Temmerman et al. [16] designed a semi-automatic approach to measure the size of the primary particles in the TEM images of powdered nanomaterials, also relying on the watershed transfer for segmentation. Muneesawang and Sirisathitkul [17] proposed a multi-level segmentation method for identifying nanoparticles. They applied a k-means method to segment the TEM image into several layers and then produced multiple binary images associated with different thresholds. After that, they separated particles by applying the watershed method to each binary image and merged the results to remove over-segmentations. Overall, these developments advance the state of the art in handling TEM images for material characterization and exploration. However, when we try to apply these methods to a set of TEM images at hand, the resulting performances are not satisfactory. In particular, the quality of detection and characterization is not robust under different resolutions and
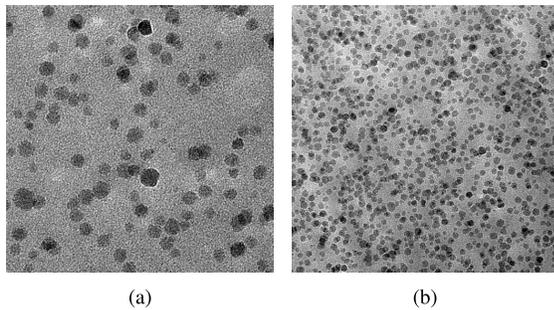
Fig. 1. Two examples of the TEM images of silica nanoparticles. (a) F3-2_7. (b) F10_8.

noise patterns. To understand the reason, we should discuss the difference between the TEM images we have and the images processed in those works.

Our TEM images are taken from an important kind of nanocomposite, which has bisphenol-F epoxy resin as the host material, blended with silicon dioxide nanoparticles, through a sol-gel process [18]. With epoxy resin as the polymer matrix and nanosilica as the nanofiller, it has enhanced mechanical properties such as modulus, hardness and fracture toughness while maintaining the optical properties (e.g., transparency), and is widely used in both academic research and industrial applications [19]. To attain a TEM image for such material, one typically takes a thin slice of sample, which has the thickness of 50 to 100 nanometers (nm) and is transparent to naked eyes. The slice is thin enough for electrons to pass through, producing an image. Two examples of the TEM images, under different instrumental resolutions, are shown in Figure 1, labeled as "F3-2_7" and "F10_8," respectively. In the images, the darker dots represent the nanoparticles, whereas the gray background represents the host material.

Compared with the nanoparticle images processed in [6] and [14]–[17], the TEM images at our disposal have much lower contrast and higher level of noises. This is due to the fact that our nanoparticles are silica particles, whose ability of shielding off or bouncing back electrons from passing through is weaker than metallic particles, and the silica particles are blended in a solid host material, whose density is not much less than the particles themselves. Meanwhile, the nonuniformity of the resin makes the background uneven (see Fig. 12(d)). By comparison, the nanoparticles used in most of the above-referenced works are metal ones, e.g., Au in [6] and [14], Ti in [16] and FePt in [17]. Those metal particles, considering its large mass, are particularly potent in bouncing back electrons, producing a sharp contrast between the particles and the background. The noisy nature of our TEM images makes the detection and characterization task more challenging. In this paper, our focus is to develop a new method for image segmentation targeting the nanoparticle detection problem in noisy and low contrast TEM images.

There are two kinds of information commonly used for image segmentation: the intensity information and the gradient information [20], [21]. The intensity-based approach is to classify the pixels with similar intensity to the same category (be it an object or the background). The gradient-based information is calculated as the first-order derivative of

the pixel intensity, signaling the magnitude of change along the way. A large gradient implies an edge (or a boundary, or a contour) separating an object from the background. To our best knowledge, many existing nanoparticle detection methods, for instance, [14], [16], [17], [22] among others, make primary use of one kind of image information (the use of intensity is more popular), causing them to only work well in certain circumstances. A natural remedy for that problem is an effective use of both kinds of image information.

The desire of combining the two kinds of image information has been raised. One strategy of combining information is to use different kinds of information sequentially, i.e. amend or enhance the segmentation results coming from the one kind of information by the other. For example, the method in [23] first over-segments an object based on the intensity and then subsequently removes false boundaries by using gradient information. Another example is the boundary refinement method [24], in which the initial boundary of an object comes from the intensity information, and the boundary is then adjusted by taking into account the gradient of pixel intensity. In the recent development, the sequential strategy is also tailored to handle specifically nanoparticle images. For instance, the approach in [6] is to first segment the foreground based on pixel intensity and find the location of centers via a modified watershed transform [7]. Then, a center is matched with the edge/boundary of the same particle, produced by Canny's edge detector [25]. At last, the approach in [6] combines the two image features (center and boundary) to locate each particle. This type of information-combining approaches work well when the boundaries detected based on gradient are similar to that of the intensity-based results, to make sure that combining the two kinds of information through a compromise could produce a better result. However, those approaches are not applicable to the noisy TEM images, since the segmentation outcomes produced by using each kind of image information alone can be drastically different (refer to Figure 8(a) and (b)), leaving little common ground for a compromise.

Another strategy of information combining is to design an energy functional, say the Mumford-Shah functional [26], integrating both kinds of information. Then the boundary of the foreground is evolving to maximize the energy functional until the local optima is found; doing so is supposed to produce the optimal separation [27]. To use the intensity information, Chan and Vese [11] assume that inside (or outside) the boundary, the variance of intensities of image should be small and Li et al. [28] assume that the intensities of pixels should change gradually, whereas to use the gradient information, Caselles et al. [29] assume that the gradient of the image along the boundary should be strong. Many recent works [30]–[32] consolidate these assumptions and design their versions, which, to certain extent, make use of both intensity and gradient information. However, applying this strategy alone cannot handle the segmentation problem of overlapped nanoparticles, because nearly all such methods, including [11], [32], can only segment the foreground from the background, leaving the overlapped objects intact within the foreground. Methods considering multiple objects detection,
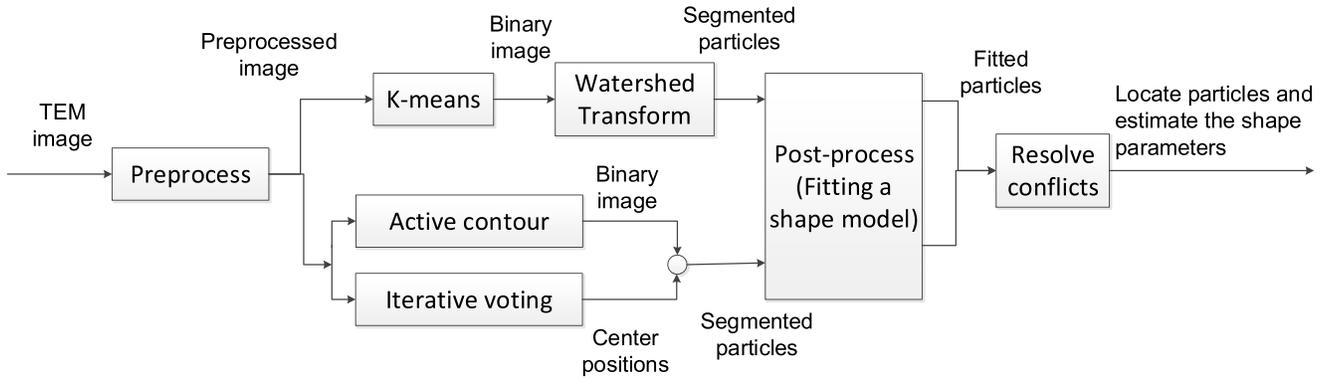
Fig. 2. The two pipelines of processing to make use of the complementary image information.

such as [28], require the objects at the foreground to have unconnected boundaries, namely that the multiple objects cannot overlap.

Recognizing the shortcomings (and strengths) of those strategies, we propose a new framework to fuse the two kinds of image information via a parallel approach. Our approach starts off with focusing the two kinds of image information separately on the same TEM image. In other words, a TEM image is handled by two pipelines of processing in parallel. One pipeline is using primarily the intensity information, segmenting the foreground by a *k*-means clustering [33] and then separating the particles according to the shape of the foreground found by a watershed transform [7], whereas the other pipeline is using primarily the gradient information, going through an active contour [11] procedure to find the foreground, followed by an iterative voting method [12] that finds the center of each particle. Intensity information is also used in the second pipeline but the main driving force therein is the gradient information, differentiating it from the first pipeline of processing.

The two pipelines of processing produce two sets of outcomes for the same image, and as expected, some of the particle detection outcomes agree with each other (which means two detections by different methods are almost the same), while many others do not. When the detection outcomes agree, it reinforces the belief that they both indicate a good detection, and when the detection outcomes differ, we then need to resolve the conflict and choose one of the outcomes. Based on a fitness criterion to be introduced later, we select the particle detection with the highest fitness score and discard those conflicting with it. A binary integer programming (BIP) is formulated and solved to obtain the optimal solution. In order to handle TEM images containing numerous particles, we also accelerate the optimizing procedure by using a subgraph decomposition technique. Our framework is illustrated in Figure 2.

We want to note a similarity between our fitness score approach and that in [34], which is in the context of tree detection. The approach in [34] is based on random point process and can be seen as a soft version of the optimization problem formulated in our paper, where overlapping is penalized but not forbidden. The random point processes are solved through Markov chain Monte Carlo, which is rather complex to optimize than the BIP formulation used in our approach.

The remaining parts of the paper are organized as follows. In Section II, we describe the basic thoughts behind the choice of the components in each processing pipeline. In Section III, we present the formulation and solution that resolves the conflicts between the two sets of processing outcomes. In Section IV, we apply our method to a set of TEM images, obtained under different instrumental resolutions and noise conditions, and assess the method's effectiveness and efficiency. Finally, we summarize our work in Section V.

II. BASIC COMPONENTS IN PROCESSING NANOPARTICLE IMAGES

The processing of nanoparticle images, illustrated in Figure 2, consists of three main steps: the preprocessing, the two pipelines of processing, and the postprocessing. This section intends to provide an overview of the basic components in the proposed framework.

The preprocessing is to enhance the image features from the noisy raw images, while the postprocessing is to fit a parametric shape model, once a nanoparticle is isolated. The two pipelines of image processing in between intend to locate the nanoparticles and isolate each of them as accurately as possible.

The two pipelines of processing are carried out on the same image in parallel. Specifically, one pipeline of processing uses primarily the intensity information, whereas the other uses primarily the gradient information. Each pipeline further involves two methods for separating and identifying the nanoparticles.

In this framework, many existing methods are used. In order to produce better results, however, certain methods, especially those used in the two pipelines of processing, are tailored towards the uniqueness of TEM images.

A. Preprocessing

The preprocessing intends to strengthen the contrast of the nanoimages and remove the unevenness in background. The background unevenness is a result of having non-uniform thickness in the slice of resin samples. Consequently, the resulting images are usually lighter on one corner/side and darker on the opposite corner/side; please see Figure 1 for
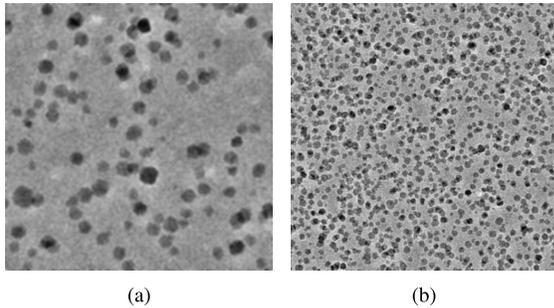
(a)   (b)

Fig. 3. The preprocessing results of the two image examples in Fig. 1. (a) F3-2_7. (b) F10_8.
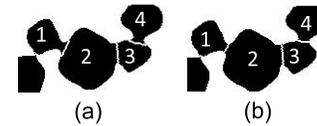

(a)   (b)

Fig. 4. The comparison of two watershed segmentations: (a) the result of the original UECS; (b) the result of the revised UECS with a timed erosion-dilation process.

an example. Two operations are used in the preprocessing: Butterworth high-pass filtering (4th order) [35] and Gaussian filtering [36].

Butterworth high-pass filtering removes the low frequency components of the image (related to the unevenness in background). A low frequency cutoff is used to identify background patterns of large size, supposedly far greater in size than a typical particle. We set this value as $1,024$ divided by four folds of the average particle diameter, where the factor of four is chosen empirically. Meanwhile, using a Gaussian filter intends to remove the high frequency components, weeding out the small objects that cannot possibly be a particle. The parameter in the Gaussian filter is set to be one-tenth of the nanoparticle's average diameter. By linking the filtering strength to the particle's average diameter, the strength of Gaussian filter's smoothing strength can be adaptively adjusted. We show the results of the preprocessing in Figure 3.

B. Intensity-Based Processing

In the first pipeline of processing, pixel intensity is used. This line of processing entails two steps: the first step is a k-means method [37] to separate the foreground from the background, producing the nanoparticle agglomerates, namely nanoparticle clusters. The second step is to use a watershed transform on the segmented foreground that further breaks the overlapped particles in the nanoparticle agglomerates into individual particles.

Each pixel in the first step is classified based on not only its intensity but also its coordinates. In [38], the image data is converted into a 5-dimensional vector $[x, y, l, u, v]$ for each pixel, where the x, y are the coordinates and l, u, v are the color values in LUV color space. For our gray-level TEM image, the input vector is set as $[w_s x, w_s y, R]$, where R is the intensity and w_s is a weighting coefficient to balance the effects between the coordinate value and the image intensity value. In this work, we set $w_s = 0.2$, as it is close to the ratio of the largest grayness value over the size of the TEM image, so we will have the similar ranges of the three coordinates. Then, we seek to find $k = 2$ clusters among the image pixels, corresponding to the foreground and background, respectively.

The second step is a watershed transform based on the shape of the foreground. A watershed transform goes through an erosion-dilation cycle, in which erosion produces the cores of neighboring objects (called markers) and dilation identifies the separating boundary lines between the objects. The specific variant of watershed transform we adopt is the Ultimate Erosion for Convex Sets (UECS) proposed by [6], which tailors its erosion stopping criterion towards convex objects, as the physical-chemical forces behind nanoparticle formation do drive nanoparticles to have convex shapes.

In the implementation of UECS, we found that the number of erosion steps can vary widely on different particles. One shortcoming of this variation in erosion steps is that the separating lines between particles tend to over-erode one of the particles. This phenomenon is illustrated in Figure 4(a); please note the over-erosion of boundary lines inside particle 2 and particle 3.

The remedy we devise to alleviate the over-erosion problem is to record the number of erosion steps, following a generic idea first introduced in [13]. Provided the number of erosion steps associated with each particle, the dilation process is then timed following the descending order of the number of erosion steps that had been performed on respective particles. For instance, suppose particle 1 was eroded 10 times to its final marker, while particle 2 was eroded 20 times. In dilation, we start with particle 2 and dilate its marker 10 times, and at which time, start the dilation of the marker of particle 1 in parallel, until the two dilated markers meet. It appears that this simple revision improves the accuracy of the boundary lines between particles appreciably; please see Figure 4(b).

C. Gradient-Based Processing

Gradient-based processing makes use of the gradient of an image to detect and separate the nanoparticles. As mentioned before, gradient-based processing also uses pixel intensity information; it is just that the gradient information plays a more deciding role here. This line of processing also entails two elements: an active contour method [32] that is based on the level set formulation and the iterative voting method [12].

The active contour method identifies the boundary (or edge) for the nanoparticles, without necessarily separating a particle agglomerate into individual particles. Then the iterative voting method locates the centers of individual particles from the preprocessed TEM images. Once the centers of individual particles as well as the boundaries of particle agglomerates are available, the connected particle contours can be separated and then assigned to individual particles by using an edge-to-marker association technique, similar to what was initially proposed in [6].

One difficulty of using the active contour method for the low-contrast, noisy TEM images is that the boundary of particles is blurred and the background is noisy, making the convergence of the recursive method sensitive to the choice
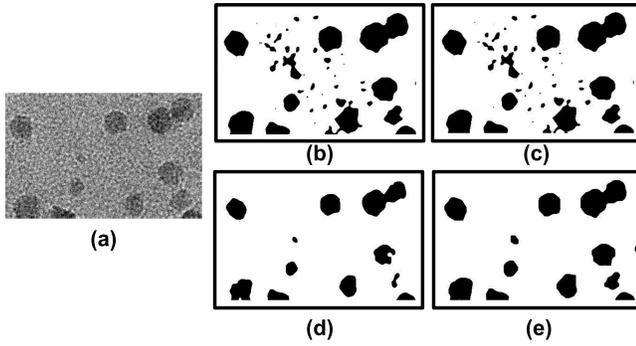
Fig. 5. The process of active contour with different initializations. (a) The original image; (b) The large mask M_1; (c) The convergent result B_1 from the large mask M_1; (d) The small mask M_2; (e) The convergent result B_2 from the small mask M_2.

of the initial contour (also known as a mask). We propose two remedies to ensure a robust convergence. Firstly we choose the active contour method, proposed by [32], that uses both intensity and gradient information, as it has a better convergence property than its counterparts that use only the intensity or the gradient information (e.g., [11], [29]).

Even with a capable method like [32], the choice of the initial contour still has a profound impact on the outcomes of contour detection for the silica nanoparticles. To find a proper initialization, we start the active contour method from either a large mask or a small mask of the foreground. Then the active contour algorithm can shrink the large mask or expand the small one to get the estimated contour. To see which mask leads to good detection outcome, we first apply Otsu's method [39] to get a binarization threshold R_t. And then we select an offset value R_s, so that we can choose masks of different sizes. Then the large mask, denoted by M_1, can be obtained by $M_1 = \{(x, y)|R(x, y) < R_t + R_s\}$, whereas the small mask, denoted by M_2, by $M_2 = \{(x, y)|R(x, y) < R_t - R_s\}$. Once M_1 and M_2 are used, the convergent outcomes are denoted by B_1 and B_2, respectively. We find that with Tian et al. [32]'s algorithm, B_2 (expanded from a small mask) is much better than B_1 (shrunk from a large mask); see an example in Figure 5. We believe that the noisy background of the TEM images makes the use of large masks ineffective (algorithm trapped in local optima). Therefore, we choose the small mask as the initial contour in the chosen active contour method.

Concerning the use of the iterative voting method, we also tailor the original method in [12] to our nano imaging problem. In the original approach, Parvin et al. [12] choose the pixels on the edge that are detected by Canny's edge detector [25] and use them to vote for locating the centers. The problem with this approach is that when some edges are hard to detect, such as in our noisy nanoparticle images, some nanoparticles will be missed. Figure 6(a) shows that using the original iterative voting method in a small region of about 20 particles produces three misses and two false detections; to produce Figure 6(a), we use the ImageJ plugin of the iterative voting method [12].

Our tailoring works as follows. Note that a large magnitude of gradient indicates that the corresponding pixel is more
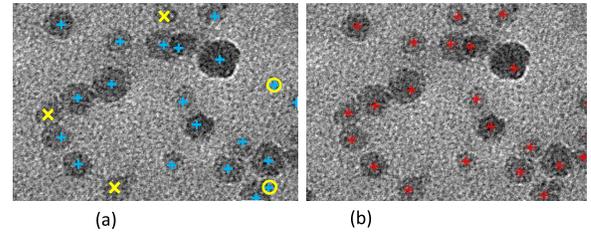

Fig. 6. The results of iterative voting: (a) ImageJ plugin. The three misses are indicated by yellow X's and two false detections are marked by yellow circles. (b) Our implementation outcome.
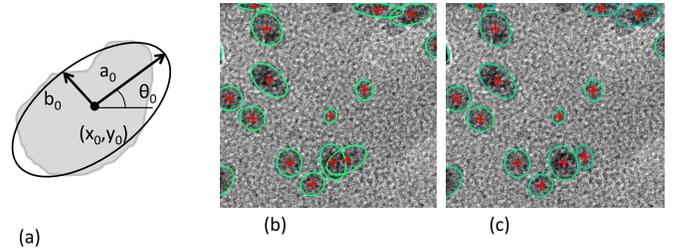

Fig. 7. Post-precessing: (a) Parametrization of an elliptical shape; (b) The fitting outcomes based on contour alone; (c) The fitting outcomes based on all the pixels in a detected particle region.

likely to be on the edge. We hence select all pixels of the preprocessed image whose gradient is larger than a threshold in magnitude, and deem them as our potential voter pixels. We then set the weight of each voter proportional to its magnitude. The threshold here is chosen as one-fifth of the maximal magnitude of gradient in the whole image. Because we skip the action of Canny's edge detection, we could not initialize the iterative voting procedure using the normal direction to the detected edges, as recommended in [12]. Instead, we let the voting direction initialized as opposite to the gradient direction at a voting pixel. Figure 6(b) shows the outcomes of our tailored approach, which is carried out on the same image and produces a result without misses and false detections.

D. Postprocessing

In post-processing, we fit each identified nanoparticle with a parametric shape model. Unlike in [6] where a particle is modeled by a B-spline, our treatment here is much simpler – we use an elliptic shape model that can be parameterized using five parameters $[x_0, y_0, a_0, b_0, \theta_0]$ (Figure 7(a)), where x_0 and y_0 are the coordinates of the center, a_0 and b_0 are the lengths of the long and short axes, and θ_0 is the orientation of the particle. The reason that we choose a simple shape model is that the silica nanoparticles produced by the sol-gel process are mostly of round or ellipse shapes; by contrast, the nanoparticles processed in [6] have shapes of wider varieties. In the meanwhile, given the noise level in the nanoimages processed in this work, it becomes less robust to use complicated shape models with too much flexibility, as a flexible shape model may be too eager to adapt itself to background noises surrounding a particle.

When it comes to the fitting of an elliptical shape, we choose to use the second-moment fitting method [40], which finds
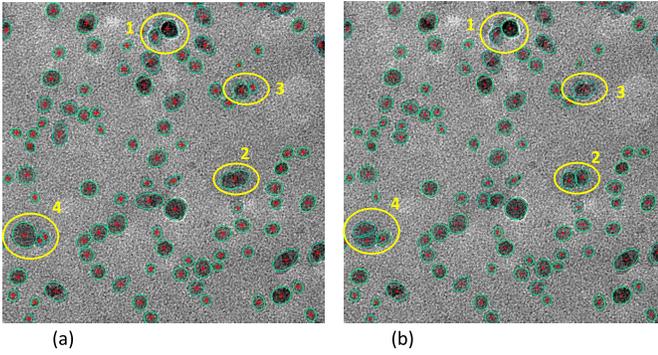
Fig. 8. The comparison of the results of the intensity-based and the gradient-based processing. (a) Intensity-based pipeline. (b) Gradient-based pipeline.

an ellipse that has the same mass center and same second moments as those of a detected particle region. This treatment uses all the pixels inside the contour of a detected particle, rather than rely on the detected contour of a particle. The drawback of using the detected contour alone is its sensitivity to shape noises, because many detected contours can end up with an irregular shape; see the example in Figure 7(a) (the gray region). This second-moment method produces much more robust shape fitting outcomes, as evident by the comparison between Figure 7(b) and (c).

E. Pros and Cons of the Two Pipelines of Processing

In Figure 8, we highlight four examples to illustrate the pros and cons of the two pipelines of processing. In example #1, the gradient-based processing produces a better boundary of the right-side particle than the intensity-based processing does. In example #2, the iterative voting in the gradient-based pipeline successfully segments two overlapped particles based on the intensity change inside the foreground region, whereas the intensity-based processing fails to do so. In example #3, the gradient-based processing fails to identify the right-side particle because of the blurred boundary, whereas the intensity-based process does detect. In example #4, the gradient-based processing over-segments the left-side particle, whereas the intensity-based processing over-segments the right-side one.

Generally speaking, our observations suggest that when the gradient is clear and accurate, the gradient-based processing works better (#1 and #2); otherwise the intensity-based processing will be more robust (#3). For some harder cases, such as #4, each pipeline of processing does half right, so only combining the two sets of the results can further improve the accuracy of the final detection. While the general observations make intuitive sense, it is not always so easy to tell which pipeline of processing will do better under a specific circumstance. This implies that the criterion that gradient information is clear and accurate sometimes can be difficult to assess and quantify manually. What is needed is an automatic conflict resolution procedure that can pick the better of the two detection outcomes.

III. FUSING THE COMPLEMENTARY INFORMATION

The next step is to make use of the detection results from the two pipelines of image processing and produce an enhanced
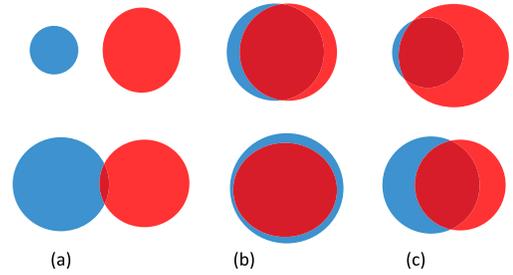

Fig. 9. Three possible relationships between $I(i)$ (blue) and $G(j)$ (red): (a) Two detection results are not related to the same particle. (b) Two results coincide with each other. (c) Two results are in conflict.

detection outcome. The problem is similar to multi-expert decision making [41], where the two pipelines of detection act as two experts and the sets of detected particles are their decisions. If both experts agree with each other on all decisions, then the problem is trivial, as one can choose either set of the outcomes. Otherwise we should devise a conflict resolving procedure to choose one of them or discard both.

Let us first introduce some notations. Let $I = \{I(i), i = 1, \cdots, N_I\}$ and $G = \{G(j), j = 1, \cdots, N_G\}$ denote the detected particles, respectively, by the intensity-based and gradient-based pipeline, where N_I and N_G are the corresponding numbers of particles detected.

The five shape parameters of $I(i)$, defined in Section II-D, are expressed as $[x_0(I(i)), y_0(I(i)), a_0(I(i)), b_0(I(i)), \theta_0(I(i))]$. The set of pixels within the fitted ellipses is labeled as $P_{I(i)}$, and its cardinality $|P_{I(i)}|$ represents the area of the corresponding region. The corresponding notations for $G(j)$ can be defined similarly.

We use the binary variables $b_{I(i)}$ and $b_{G(j)}$ to indicate the outcome of our resolution: if $I(i)$ (or $G(j)$) is chosen as the final detection outcome, then $b_{I(i)}$ (or $b_{G(j)}$) will be set as 1, otherwise it is set as 0. Aggregating all the decision variables associated with individual detections, the decision vector for the intensity-based approach is expressed as $\mathbf{b}_I = [b_{I(1)}, \cdots, b_{I(N_I)}]^T$, and that for the gradient-based approach is $\mathbf{b}_G = [b_{G(1)}, \cdots, b_{G(N_G)}]^T$. Our goal is to find an optimal solution of $\mathbf{b}_I$ and $\mathbf{b}_G$, which is to properly set elements of $\mathbf{b}_I$ and $\mathbf{b}_G$ to 1 or 0, according to an optimality criterion introduced below.

A. Basic Formulation for Conflict Resolution

One crucial step in making good use of the two types of detection outcomes is to understand the three possible relationships between $I(i)$ and $G(j)$. The relationship is illustrated in Figure 9. When two detection outcomes have only a slight overlap or no overlap at all, as shown in Column (a) of Figure 9, it is unlikely that they are related to the common particle in the image. When the two outcomes virtually coincide with each other, manifesting in a heavy overlap between the detection regions, they point to the same underlying particle and are then referred to as a consensus detection. When the two outcomes occupy the same region in the image, but the detected particles have serious disagreement, either in number (one approach detects one particle, while the other detects two, for instance) or in key shape parameters

(including the center location), these outcomes are referred to as the conflicting detections. The consensus detections and the conflicting detections are illustrated in Columns (b) and (c) of Figure 9, respectively. The unrelated and consensus detections are relatively straightforward to deal with. It is the conflicting detections that need further processing to decide which one to be the final detection outcome.

As such, there are two primary questions to be addressed:
1) How to determine which category of relation (unrelated, consensus, and conflicting) $I(i)$-versus-$G(j)$ belongs to?
2) Once this relation is determined as a conflicting detection, what criterion to use to make the final selection?

The answer to the first question apparently depends on the degree of overlap between two detections; the above description of the three relationships provides the intuition behind it. The specific formula will be presented later in Section III-B.

To address the second question, we assign each particle detection with a score, assessing its fitness to the original image. Intuitively speaking, the higher the score, the better a detection fits the original image. We denote the fitness score vector of a detection as $\mathbf{s}_I = [s_{I(1)}, \cdots, s_{I(N_I)}]^T$ for the intensity-based approach and $\mathbf{s}_G = [s_{G(1)}, \cdots, s_{G(N_G)}]^T$ for the gradient-based approach. The specific definition of the fitness score is provided in Section III-C.

When the two pipelines of processing reach a consensus, it enhances the credibility of the detection and makes such detection more reliable and trustworthy. It is safe to take the consensus outcomes and add them into the final detection results without further processing. We compute the shape parameters of the final particle by averaging the corresponding parameters of the two detections. Then we remove these particles from the sets of I and G, so that only the conflicting detections are left to be resolved. Denote the sets of the remaining particles as $\tilde{I} = \{\tilde{I}(1), \cdots, \tilde{I}(N_{\tilde{I}})\}$ and $\tilde{G} = \{\tilde{G}(1), \cdots, \tilde{G}(N_{\tilde{G}})\}$, where $N_{\tilde{I}}$ and $N_{\tilde{G}}$ are the numbers of particles in the two revised sets, respectively. In the subsequent conflict resolving procedure, we only need to solve for $\mathbf{b}_{\tilde{I}}$ and $\mathbf{b}_{\tilde{G}}$, which are a subset of $\mathbf{b}_I$ and $\mathbf{b}_G$, respectively, and have usually fewer than half of the original elements.

For the remaining conflicting detections, we use a conflict matrix $\mathbf{M} = (M_{ij})$ to connect them. $\mathbf{M}$ is an $N_{\tilde{I}} \times N_{\tilde{G}}$ binary matrix, with each row representing one particle in $\tilde{I}$ and each column representing one particle in $\tilde{G}$. If $\tilde{I}(i)$ and $\tilde{G}(j)$ are conflicting, $M_{ij} = 1$; if they are unrelated, $M_{ij} = 0$. Figure 10 shows a simple example of conflicting detections and the corresponding conflict matrix. In Figure 10, we observe that $\tilde{I}(1)$ is conflicting with $\tilde{G}(1)$, while $\tilde{I}(2)$ is conflicting with both $\tilde{G}(2)$ and $\tilde{G}(3)$; this is reflected in the 2×3 conflict matrix to the right.

With a fitness score chosen, we present the following constrained binary integer programming (BIP) problem for selecting the final detection out of a conflict:

$$\max_{\mathbf{b}_{\tilde{I}}, \mathbf{b}_{\tilde{G}}} \quad \mathbf{s}_I^T \mathbf{b}_{\tilde{I}} + \mathbf{s}_G^T \mathbf{b}_{\tilde{G}},$$

$$\text{subject to} \quad \mathbf{b}_{\tilde{I}}^T \mathbf{M} \mathbf{b}_{\tilde{G}} = 0. \tag{1}$$
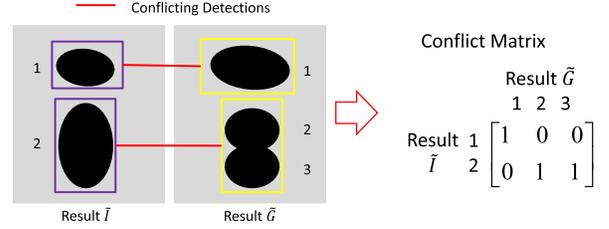

Fig. 10. An example of conflicting detections (left) and the corresponding conflict matrix (right).

The objective function is the summation of the fitness scores of all detections, and we aim to obtain the highest total fitness score for an image. The constraint function is to ensure that only one of the conflicting detections will be chosen. To see this, rewrite the constraint as

$$\sum_{i=1}^{N_{\tilde{I}}} \sum_{j=1}^{N_{\tilde{G}}} b_{\tilde{I}(i)} M_{ij} b_{\tilde{G}(j)} = 0, \tag{2}$$

meaning that if $\tilde{I}(i)$ and $\tilde{G}(j)$ are a pair of conflicting detections, namely $M_{ij} = 1$, then $b_{\tilde{I}(i)}$ and $b_{\tilde{G}(j)}$ cannot be 1 simultaneously. We solve this BIP problem by using the MATLAB solver 'bintprog'. To use it, we multiply a negative sign to the objective function to change the problem to a minimization problem.

We want to note that the authors in [34] also used model fitness in an application of tree detection. More specifically, they first use an unknown number of ellipses to model the trees on a plantation and then calculate the prior energy and the likelihood according to prior knowledge and the observed images. At last, they minimized the Bayesian energy using Markov chain Monte Carlo to find the ellipses that fit the tree crowns the best. There are a couple of differences between their method and our conflict resolution approach. Firstly, the approach in [34] is based on random point process. They wanted to minimize the overlapping of different ellipses, which penalizes the overlapping but does not forbid it, whereas in our approach, we have to choose one of the outcomes. Secondly, their Bayesian based solution procedure is more complicated than the BIP formulation we use.

B. Consensus and Conflicting Detections

From Figure 9, we can see that the degree of overlap between the two detection outcomes can be used to decide which category a pair of detections belongs to. When the Euclidean distance between the centers of the two detections is larger than $(a_0(I(i)) + a_0(G(j)))/2$, it means that there is no overlap between the two detected particles. The pair is then unrelated. When the distance is smaller than $(a_0(I(i)) + a_0(G(j)))/2$, we need to quantify the degree of overlap. The area of overlap is $|P_{I(i)} \cap P_{G(j)}|$. We calculate the maximum overlapping ratio $r_{\max}$ and minimum overlapping ratio $r_{\min}$ as follows:

$$r_{\max}(I(i), G(j)) = \max\{\frac{|P_{I(i)} \cap P_{G(j)}|}{|P_{I(i)}|}, \frac{|P_{I(i)} \cap P_{G(j)}|}{|P_{G(j)}|}\},$$

$$r_{\min}(I(i), G(j)) = \min\{\frac{|P_{I(i)} \cap P_{G(j)}|}{|P_{I(i)}|}, \frac{|P_{I(i)} \cap P_{G(j)}|}{|P_{G(j)}|}\}. \tag{3}$$
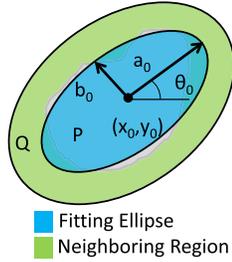
Fig. 11. The foreground region P (blue) and its neighboring region Q (green) for a detected particle.

We then set two thresholds, an upper ratio r_U and a lower ratio r_L, such that if $r_{\max}(I(i), G(j)) < r_L$, we deem the overlapping region small enough to declare $I(i)$ and $G(j)$ unrelated; if $r_{\min}(I(i), G(j)) < r_U$ and $r_{\max}(I(i), G(j)) > r_L$, we believe that the two detection outcomes are related but different, namely that they form a pair of conflicts; if $r_{\min}(I(i), G(j)) > r_U$, we consider this as a consensus detection.

C. Fitness Score of Detections

Essentially, calculating the fitness score for each particle is equivalent to evaluating the quality of the image segmentation, in which a regional part of TEM images is separated into the particle and its surrounding area. Zhang et al. [42] surveyed different evaluation methods for image segmentation quality when the ground truth is unknown. They pointed out a simple principle that is still widely used: the inter-region disparity should be large and intra-region variability should be small. For instance, Fisker et al. [43] maximize the difference in the average intensities between the foreground and its surrounding background for detecting a particle. To measure the inter-region disparity and the intra-region similarity, we need to define a neighboring region Q for particles in $\tilde{I}$ and $\tilde{G}$. Consider a particle $\tilde{I}(i)$ (the same can be done to $\tilde{G}(j)$). Its foreground information is in $P_{\tilde{I}(i)}$ and the surrounding background information is in $Q_{\tilde{I}(i)}$ (shown in Figure 11). In identifying $Q_{\tilde{I}(i)}$, we double the size of $P_{\tilde{I}(i)}$, namely $|Q_{\tilde{I}(i)} \cup P_{\tilde{I}(i)}| = 2|P_{\tilde{I}(i)}|$, so that $|Q_{\tilde{I}(i)}| = |P_{\tilde{I}(i)}|$.

Our measure of the inter-region disparity and the intra-region similarity is based on the sum of squares of pixel intensities. The sum of squares are proportional to the variance of the intensities within a region, so a large value indicates disparity while a small value indicates similarity. For a good segmentation, the sum of squares of the whole region should be much larger than that of separated background or foreground. For an arbitrary region A in the image, its sum of squares of the intensity, denoted by $SS(A)$, is calculated by:

$$SS(A) = \sum_{(x,y) \in A} [R(x, y) - \bar{R}(A)]^2, \qquad (4)$$

where $\bar{R}(A)$ is the average intensity of all pixels inside A. We then define the fitness score of $\tilde{I}(i)$ as:

$$s_{\tilde{I}(i)} = SS(P_{\tilde{I}(i)} \cup Q_{\tilde{I}(i)}) - [SS(P_{\tilde{I}(i)}) + SS(Q_{\tilde{I}(i)})] \\ - \lambda |P_{\tilde{I}(i)} \cup Q_{\tilde{I}(i)}|, \qquad (5)$$

where the first term $SS(P_{\tilde{I}(i)} \cup Q_{\tilde{I}(i)})$ measures the inter-region disparity, and the second term $[SS(P_{\tilde{I}(i)}) + SS(Q_{\tilde{I}(i)})]$ measures the intra-region similarity. The greater their difference, the stronger indication it is to think that $\tilde{I}(i)$ is part of the particle's foreground. The third term is a noise filter. Its inclusion forces the difference between the inter-region disparity and the intra-region similarity to be great enough so as to qualify $\tilde{I}(i)$ as a genuine particle, helping reduce false detections in a noisy image. If $\tilde{I}(i)$ is a single unrelated particles, which means it has no conflicting detection in another set of results, it will be selected if and only if $s_{\tilde{I}(i)}$ is larger then 0.

In equation (5), the first term is the total sum of squares of the whole region and the second term is the within-group sum of squares. According to the property of variance [44], their difference equals the between-group sum of squares, i.e.,

$$|P_{\tilde{I}(i)}|[\bar{R}(P_{\tilde{I}(i)}) - \bar{R}(P_{\tilde{I}(i)} \cup Q_{\tilde{I}(i)})]^2 \\ + |Q_{\tilde{I}(i)}|[\bar{R}(Q_{\tilde{I}(i)}) - \bar{R}(P_{\tilde{I}(i)} \cup Q_{\tilde{I}(i)})]^2, \qquad (6)$$

where $\bar{R}(P_{\tilde{I}(i)})$, $\bar{R}(Q_{\tilde{I}(i)})$ and $\bar{R}(P_{\tilde{I}(i)} \cup Q_{\tilde{I}(i)})$ are the average intensities of the foreground, its neighboring region, and the combined whole area, respectively. By the choice of neighboring region made above, namely $|Q_{\tilde{I}(i)}| = |P_{\tilde{I}(i)}|$ (they may not be exactly the same but the difference is negligible), it means:

$$\bar{R}(P_{\tilde{I}(i)} \cup Q_{\tilde{I}(i)}) = (\bar{R}(P_{\tilde{I}(i)}) + \bar{R}(Q_{\tilde{I}(i)}))/2. \qquad (7)$$

Plugging in equations (6) and (7) into equation (5), we have

$$s_{\tilde{I}(i)} = |P_{\tilde{I}(i)} \cup Q_{\tilde{I}(i)}| \left\{ \left(\frac{\bar{R}(P_{\tilde{I}(i)}) - \bar{R}(Q_{\tilde{I}(i)})}{2} \right)^2 - \lambda \right\}. \qquad (8)$$

It is now clear how the third term in equation (5) works – if the intensity difference between the foreground and background is smaller than the threshold $2\sqrt{\lambda}$, then, the fitness score $s_{\tilde{I}(i)}$ turns negative, and consequently, $\tilde{I}(i)$ will not be chosen as a particle.

D. Decomposition and Linearization

To solve the optimization problem (1) efficiently, we need to address two more problems: (a) There are hundreds to thousands of particles in $\tilde{I}$ and $\tilde{G}$ in a TEM image. Solving the optimization in its current form is time consuming. (b) The constraint in (1) is not linear, which prevents a straightforward application of some existing efficient methods. It is necessary to decompose the original problem into smaller-sized subproblems, and to linearize the constraint.

The way to decompose the original optimization problem is to decompose the conflict matrix $\mathbf{M}$. If $\mathbf{M}$ can be expressed in a block form with zero off-diagonal submatrices, then, each block submatrix can be used to form a separate BIP problem and be solved in parallel. A simple example is a two-block $\mathbf{M}$, such as

$$\mathbf{M} = \begin{bmatrix} \mathbf{M}_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{M}_2 \end{bmatrix}, \qquad (9)$$

then equation (1) can be decomposed into two BIP problems:

$$\max_{\mathbf{b}_{\tilde{I}_1}, \mathbf{b}_{\tilde{G}_1}} \mathbf{s}_{\tilde{I}_1}^T \mathbf{b}_{\tilde{I}_1} + \mathbf{s}_{\tilde{G}_1}^T \mathbf{b}_{\tilde{G}_1} \quad \max_{\mathbf{b}_{\tilde{I}_2}, \mathbf{b}_{\tilde{G}_2}} \mathbf{s}_{\tilde{I}_2}^T \mathbf{b}_{\tilde{I}_2} + \mathbf{s}_{\tilde{G}_2}^T \mathbf{b}_{\tilde{G}_2}$$

$$\text{subject to } \mathbf{b}_{\tilde{I}_1}^T \mathbf{M}_1 \mathbf{b}_{\tilde{G}_1} = 0 \quad \text{subject to } \mathbf{b}_{\tilde{I}_2}^T \mathbf{M}_2 \mathbf{b}_{\tilde{G}_2} = 0, \quad (10)$$

where $\mathbf{s}_{\tilde{I}} = [\mathbf{s}_{\tilde{I}_1}; \mathbf{s}_{\tilde{I}_2}]$ and $\mathbf{s}_{\tilde{G}} = [\mathbf{s}_{\tilde{G}_1}; \mathbf{s}_{\tilde{G}_2}]$. After solving those two subproblems, the minimizer of the original problem can be easily obtained by combining their individual solutions, namely $\mathbf{b}_{\tilde{I}} = [\mathbf{b}_{\tilde{I}_1}; \mathbf{b}_{\tilde{I}_2}]$ and $\mathbf{b}_{\tilde{G}} = [\mathbf{b}_{\tilde{G}_1}; \mathbf{b}_{\tilde{G}_2}]$.

The decomposition of the BIP can also be seen as a problem to find the connected independent subgraph. We regard the $N_{\tilde{I}} + N_{\tilde{G}}$ particles in $\tilde{I}$ and $\tilde{G}$ as nodes to build an undirected graph $\mathbf{G}$. Then, we connect two nodes if they form a fair of conflicting detection and obtain the corresponding adjacent matrix $\mathbf{W}$. If we can find an independent connected subgraph containing, for example, $\tilde{I}(1)$, $\tilde{I}(2)$ and $\tilde{G}(1)$, $\tilde{G}(2)$, $\tilde{G}(3)$, that means there is no conflicting relationship between them and any other particles. So we can form a subproblem only concerning those five particles, and the solution of that subproblem is the same as the corresponding part of the whole problem. To find all connected independent subgraphs in $\mathbf{G}$, we adopt the spectrum analysis method in [45].

The theory in [45] says that the number of independent connected subgraphs of $\mathbf{G}$ equals to the multiplicity of 0 eigenvalue of its normalized graph Laplacian matrix:

$$\mathbf{L} = \mathbf{I} - \mathbf{D}^{-\frac{1}{2}} \mathbf{W} \mathbf{D}^{-\frac{1}{2}}, \quad (11)$$

where $\mathbf{W}$ is the adjacent matrix of the graph $\mathbf{G}$, $\mathbf{I}$ is the identical matrix which has the same size of $\mathbf{W}$, and $\mathbf{D}$ is the diagonal matrix of the row (or column) sum of $\mathbf{W}$. Von Luxburg [45] provides a detailed procedure. Following their procedure, first check if the graph $\mathbf{G}$ is decomposable (i.e., check the multiplicity of 0 eigenvalue of $\mathbf{L}$), and if this multiplicity is $K > 1$, then $\mathbf{G}$ can be decomposed to a set of K independent connected subgraphs. Then we can break $\mathbf{M}$ into K block submatrices $\{\mathbf{M}_k\}_{i=1}^K$, and the fitness score vectors $\mathbf{s}_{\tilde{I}}$ and $\mathbf{s}_{\tilde{G}}$ into $\{\mathbf{s}_{\tilde{I}_k}\}_{i=1}^K$ and $\{\mathbf{s}_{\tilde{G}_k}\}_{i=1}^K$, respectively. As such, the original BIP can be decomposed to K smaller subproblems that can be solved in parallel. The kth subproblem is:

$$\max_{\mathbf{b}_{\tilde{I}_k}, \mathbf{b}_{\tilde{G}_k}} \mathbf{s}_{\tilde{I}_k}^T \mathbf{b}_{\tilde{I}_k} + \mathbf{s}_{\tilde{G}_k}^T \mathbf{b}_{\tilde{G}_k},$$

$$\text{subject to} \quad \mathbf{b}_{\tilde{I}_k}^T \mathbf{M}_k \mathbf{b}_{\tilde{G}_k} = 0. \quad (12)$$

Next, we show that the constraint in equation (1) can be linearized. Because $\mathbf{b}_{\tilde{I}}$, $\mathbf{b}_{\tilde{G}}$ and $\mathbf{M}$ are binary vectors/matrix, the original constraint can be replaced by the following inequality:

$$\mathbf{M}^T \mathbf{b}_{\tilde{I}} + N_{\tilde{I}} \mathbf{b}_{\tilde{G}} \le N_{\tilde{I}} \mathbf{1}_{N_{\tilde{G}}}, \quad (13)$$

where $\mathbf{1}_{N_{\tilde{G}}}$ represents an $N_{\tilde{G}} \times 1$ vector whose elements are all 1's.

We can show that the original constraint and equation (13) are equivalent. For the constraint in (1), it is obvious to see that the constraint is violated if and only if there exists any pair of i and j satisfying $M_{ij} = 1$, $b_{\tilde{I}(i)} = 1$ and $b_{\tilde{G}(j)} = 1$.

We want to show that equation (13) is violated under the same condition.

Equation (13) produces $N_{\tilde{G}}$ linear inequalities. Let us consider the jth inequality:

$$\sum_{i=1}^{N_{\tilde{I}}} M_{ij} b_{\tilde{I}(i)} + N_{\tilde{I}} b_{\tilde{G}(j)} \le N_{\tilde{I}}. \quad (14)$$

1) If $b_{\tilde{G}(j)} = 0$, because M_{ij} and $b_{\tilde{I}(i)}$ are both binary variables taking either 0 or 1, $\sum_{i=1}^{N_{\tilde{I}}} M_{ij} b_{\tilde{I}(i)} \le N_{\tilde{I}}$ is always true. This suggests that regardless the choice of $\mathbf{b}_{\tilde{I}}$, the constraint in (14) is satisfied.

2) If $b_{\tilde{G}(j)} = 1$, $N_{\tilde{I}} b_{\tilde{G}(j)}$ equals to $N_{\tilde{I}}$. If there exits any i satisfying $M_{ij} = 1$ and $b_{\tilde{I}(i)} = 1$, then $\sum_{i=1}^{N_{\tilde{I}}} M_{ij} b_{\tilde{I}(i)}$ is larger than 0, making the inequality untrue. In order for the inequality to hold, the first term must be 0, meaning when $b_{\tilde{G}(j)} = 1$, M_{ij} and $b_{\tilde{I}(i)}$ cannot be 1 at the same time.

The above argument extends to all j's.

As such, we can replace the original constraint with the inequality in (13), which is linear. As the objective function is also linear, we can use efficient linear binary programming methods (such as a branch-and-bound algorithm [46]) to solve the optimization problem.

IV. Experimental Results

A. Parameter Selection

One parameter used throughout the algorithm is the average diameter of the nanoparticle size, denoted by d_0. The d_0 can be considered as the average effect of a_0 and b_0 in the particle shape model, and it is used as the input to set a number of other settings in the algorithm. The value of d_0 in a TEM image is largely determined by a particle's actual size and the resolution level set in the TEM. Informed by our material science collaborators, we know about the average physical diameter of the nanoparticles to be blended in the host material. The physical size is used to estimate d_0 under a specific TEM resolution, which is a good enough initial estimate and can be refined once the TEM image is processed. The refined estimate of d_0 can be used to run the whole algorithm a second time so as to improve the quality of the processing.

In the main part of the algorithm, there are two other sets of parameters: (1) r_U and r_L that are used to categorize the detection outcomes into three groups, i.e., unrelated, consensus, and conflicting; (2) λ in the fitness score. We empirically choose $r_U = 0.8$ and $r_L = 0.2$. We test many TEM images and find that these choices produce rather robust categorizations consistent with human interpretation. We want to set the pixel intensity gap to be about one-tenth of the grayness levels from the brightest to the darkest in the TEM images, in order to differentiate a particle's foreground from its surrounding background. For noisy TEM images, this gap appears reasonable. Given that our TEM images have roughly 200 grayness levels, it suggests that the gap is going to be 20, and according to equation (8), this sets $\lambda = 100$.
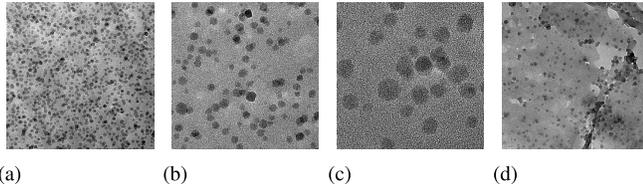
Fig. 12. Samples of TEM images. (a) Low resolution. (b) Medium resolution. (c) High resolution. (d) Uneven background.

TABLE I
ESTIMATES OF d_0 IN THE TEM IMAGES

	Low resolution	Medium resolution	High resolution	Uneven background
d_0	20	50	$100 - 120$	20

B. TEM Images Used in the Test

We test a total of 32 TEM images taken of the bisphenol-F epoxy resin samples that are blended with silica nanoparticles. These images can be grouped into four categories. The first three categories correspond to different resolution levels of TEM. All TEM images have $1,024 \times 1,024$ pixels. So the low resolution image is taken from a big view field of about $1,000 \times 1,000$ nm; the medium resolution taken from a view field of about 500×500 nm; the high resolution image taken from the smallest view field of 250×250 nm. The last category, and also the fourth, of the images is the one having an uneven background of particular patterns. This background pattern is a result of inconsistency in the resin properties, so that the nanoparticles do not disperse well as they are blended in. This set of images with uneven background is taken under the low resolution. Figure 12 shows one typical image for each of the kinds, in which Figure 12(a) and (b) are the same images as those shown in Figure 1.

As we explained earlier, the average particle size d_0 in an image is affected by the resolution of TEM. The ranges of these d_0's in the aforementioned four categories of images are presented in Table I.

C. Comparing the Integrated Approach With Individual Pipeline of Processing

Using the two images in Figure 1, we want to show where the integrated approach improves upon the individual pipelines of processing. Figure 13 presents the detection outcomes. The two images illustrate the results of the integrated processing. The images are color coded: green means a consensus detection, blue means that an intensity-based detection prevails, and yellow means that a gradient-based detection prevails. In the low resolution image ("F10_8"), there are 721 consensus detections, out of $1,100$ particles finally detected. Among the 379 conflicting detections, 162, or 43%, final outcomes come from the intensity-based processing, whereas 217, or 57%, come from the gradient-based processing. The respective numbers for the medium resolution image ("F3-2_7") are: 103 total particles, 85 consensus detections, 18 conflicting detections, and among those particles, 9, or 50%, are from the
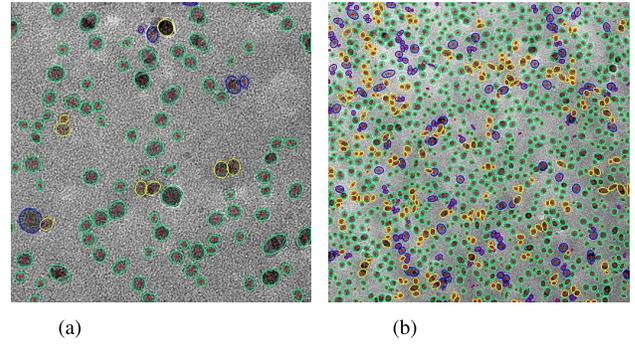

Fig. 13. Comparison of individual pipelines of processing. The left image corresponds to Figure 1 (medium-resolution image), whereas the right image corresponds to Figure 1 (low-resolution image). Green particles are those from the consensus detections; blue particles are an intensity-based detection; yellow particles are an gradient-based detection. (a) Medium resolution. (b) Low resolution.
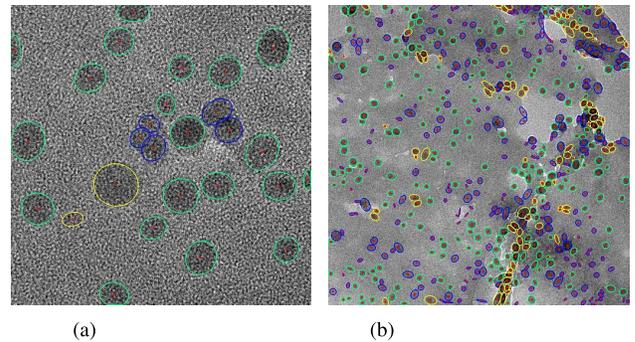

Fig. 14. Comparison of individual pipelines of processing for other two categories. The left image corresponds to Figure "F3-2_11" (high-resolution image), whereas the right image corresponds to Figure "F8-2_16" (image with uneven background). Color coding is the same as in Figure 13. (a) High resolution. (b) Uneven background.

intensity-based processing, whereas the other 9, or 50%, from the gradient-based processing. Figure 14 presents the outcomes of the integrated processing for other two categories of TEM images; the same color code applies. We observe again that the integrated processing improves upon the individual pipeline of processing and we believe that this is a key advantage of the integrated approach, as it makes use of the image information fully and compensate for the limitations of the approaches emphasizing too much on one type of image information.

D. Test Outcomes of Four Kinds of TEM Images

To quantify the performance of our method, we run the algorithm on all 32 TEM images and report the number of particles they are able to identify. For the medium and high resolution images, we are able to manually label the particles and treat the manual outcome as our ground truth. These detection results are included in Table II. In Table II, for the individual pipeline of processing, we report the numbers of the total particle detections as well as the numbers of the consensus detections and conflicting detections selected by the integrated method. The percentages of the conflicting detections selected from each pipeline are also shown in the table. For further comparison, we define the dissimilarity
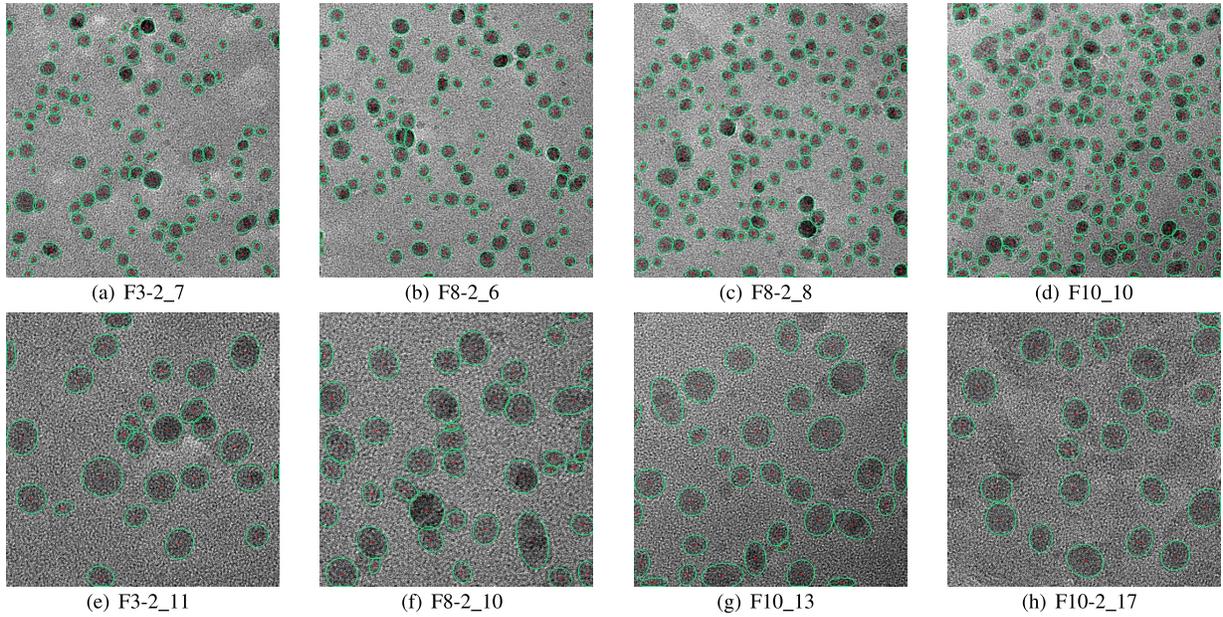
Fig. 15. The processed outcomes of medium-resolution (top row) images and high-resolution (bottom row) images. (a) F3-2_7. (b) F8-2_6. (c) F8-2_8. (d) F10_10. (e) F3-2_11. (f) F8-2_10. (g) F10_13. (h) F10-2_17.

TABLE II
COMPARISON OF PARTICLE DETECTIONS FOR MEDIUM
AND HIGH RESOLUTION IMAGES

TEM image	Ground truth	Intensity Based	Gradient Based	Consensus Detections	Selected From Intensity		Selected From Gradient		Integrated approach
					Number	Percentage	Number	Percentage	
Medium resolution images									
F3-2_6	103	99	97	85	9	56.3%	7	43.7%	101
F3-2_7	104	100	99	85	9	50%	9	50%	103
F3-2_8	100	99	98	73	10	35.7%	18	64.3%	101
F8-2_6	113	108	111	98	5	35.7%	9	64.3%	112
F8-2_7	134	126	131	119	1	7.7%	12	92.3%	132
F8-2_8	148	143	142	119	14	51.8%	13	48.2%	146
F10_10	214	201	195	114	64	66%	33	34%	211
F10_12	179	175	162	141	30	88.2%	4	11.8%	175
High resolution images									
F3-2_9	24	24	25	14	0	0%	10	100%	24
F3-2_10	26	26	24	11	4	28.6%	10	71.4%	25
F3-2_11	26	33	25	20	5	71.4%	2	28.6%	27
F8-2_10	42	44	37	17	12	50%	12	50%	41
F8-2_11	44	41	35	20	16	69.6%	7	30.4%	43
F10_13	37	41	36	23	4	33.3%	8	66.7%	35
F10_15	47	50	34	17	24	82.8%	5	17.2%	46
F10-2_17	25	31	25	19	4	66.7%	2	33.3%	25

Fig. 16. The boxplot of the dissimilarity metric for (a) medium resolution images and (b) high resolution images.

The results presented in Table II and Figure 16 demonstrate the effectiveness of the integrated approach. Both of intensity-based and gradient-based processing contribute to the intergraded results and combining their strengths allows the proposed method to achieve a high degree of accuracy consistently across the samples. We also conduct an analysis of variance (ANOVA) [47] on the dissimilarity of three groups (integrated approaches, intensity-based only and gradient-based only) for the medium and high resolution images. For the medium resolution images, the p-value of an one-way ANOVA test is 0.0124 between the integrated approach and the intensity-based approach and 0.0013 between the integrated approach and the gradient-based approach. For the high resolution images, the p-value is 0.0001 between the integrated approach and the intensity-based approach and 0.0025 between the integrated approach and the gradient-based approach.
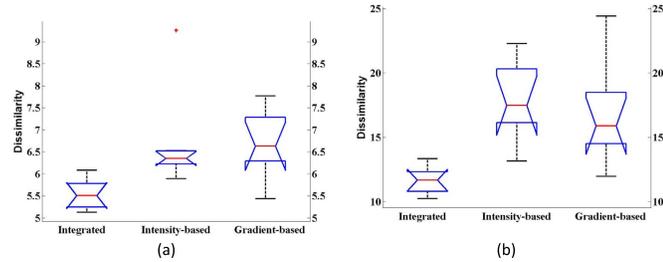
For the low resolution images including the ones with uneven background, it is difficult to manually count and identify all the particles, as they usually have over hundreds or even thousands of particles. What we do here is to present the processed outcomes of individual images in Figure 17, so that people can visually sense how the method performs. We still only show half of the results due to the images large size. We present a table, similar to Table II, but it does not have the ground truth column. For the intensity-based and gradient-based approaches, we again report the numbers of particles it detects and the numbers of the conflicted outcomes selected by the integrated method. Combining both Table III and Figure 17, we believe that the proposed method presents an advantage in achieving robust detections when the image quality varies.

Table II and Table III also suggest that the two pipelines of processing make similar contributions for the low, medium

between the detected outcomes and the ground truth as the average distance between the nearest centers of different point sets, and show the boxplots of the comparison results in Figure 16. The smaller the dissimilarity, the better a detection outcome. Because of the availability of the results in Table II and Figure 16, we only present half of the processed images in Figure 15, as inclusion of all images makes the paper's file size too large.
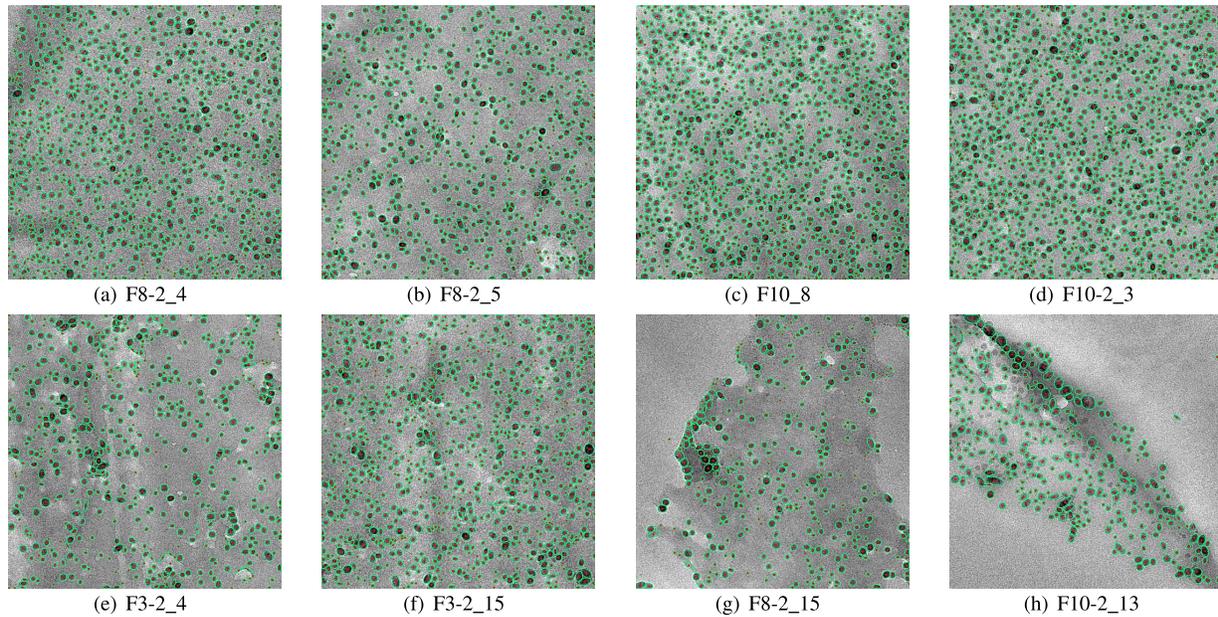
Fig. 17. The processed outcomes of low-resolution images (top row) and the images with uneven background (bottom row). (a) F8-2_4. (b) F8-2_5. (c) F10_8. (d) F10-2_3. (e) F3-2_4. (f) F3-2_15. (g) F8-2_15. (h) F10-2_13.

TABLE III
COMPARISON OF PARTICLE DETECTIONS FOR LOW RESOLUTION IMAGES

TEM image	Intensity Based	Gradient Based	Consensus Detections	Selected From Intensity Number	Selected From Intensity Percentage	Selected From Gradient Number	Selected From Gradient Percentage	Integrated approach
			Low resolution images					
F3-2_16	826	695	403	257	61.6%	160	38.4%	820
F8_8	1197	997	595	425	67.9%	201	32.1%	1221
F8-2_4	871	822	575	189	62.4%	141	37.6%	878
F8-2_5	633	678	510	65	40.4%	96	59.6%	671
F10_7	885	924	667	109	43.3%	143	56.7%	919
F10_8	1041	1077	721	162	42.7%	217	57.3%	1100
F10_9	1115	1153	730	211	46.6%	242	53.4%	1183
F10-2_3	1053	1096	763	153	43.5%	199	56.3%	1115
			Uneven background images					
F3-2_4	502	487	294	133	60.7%	86	39.3%	513
F3-2_5	465	463	228	150	55.3%	121	44.7%	499
F3-2_15	815	712	466	222	64.5%	122	35.5%	810
F8_13	291	200	95	124	73.8%	44	26.2%	263
F8-2_15	327	309	159	133	65.3%	60	34.7%	332
F8-2_16	556	398	199	247	73.3%	90	26.7%	536
F10-2_12	480	187	102	303	95.3%	15	4.7%	420
F10-2_13	290	259	165	80	60.6%	52	39.4%	297

Fig. 18. Computational time of the algorithm. The horizontal axis is d_0, and the vertical axis is the processing time in minutes.

and high-resolution TEM images. But for those images with uneven background, more conflicting outcomes are selected from the intensity-based processing than from the gradient-based processing. We believe that the unevenness in background intensity causes confusion in using the gradient information, making the intensity-based processing more accurate and the gradient-based processing less so.

E. Computation Time

People perceive that the time spent to process a TEM image is proportional to the number of particles in an image. This turns out untrue. The processing time in fact depends heavily on the resolution level of an image; see Figure 18. The horizontal axis is the value of d_0 related to an image's resolution level. As the resolution gets higher, d_0 gets bigger, even for particles of the same physical size. Our algorithm spent longer time to process the high resolution images than the low-resolution ones. But the overall time is manageable. For the 32 images, the longest processing time is about 10 minutes. Recall that our method is intended to be an offline
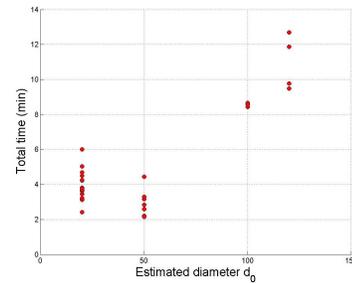
processing tool, so our material science collaborators deem 10 minutes very much acceptable.

The most time consuming part of our algorithm is the two iterative processing components: the active contour and the iterative voting. When processing the high resolution images, the heavier noise and lower contrast make it harder for the active contour method to find the optimal solution of their energy functional. So it takes a longer time to converge. When the iterative voting is applied to the high resolution images, the large diameter of particles d_0 leads to a large voting region for each step in its iteration, also causing a longer time for the method to execute.

F. Parameter Sensitivity

In this section, we discuss the effect of the input parameter d_0 on the detection results. We test a given set of TEM images using different d_0's and generate the box-plot of the dissimilarity and total processing time in Figure 19. We choose the medium resolution TEM images because (a) these images contain a good number of nanoparticles and (b) the number of particles is manageable so that we can manually verify the ground-truth. The recommended value of d_0 is 50, which
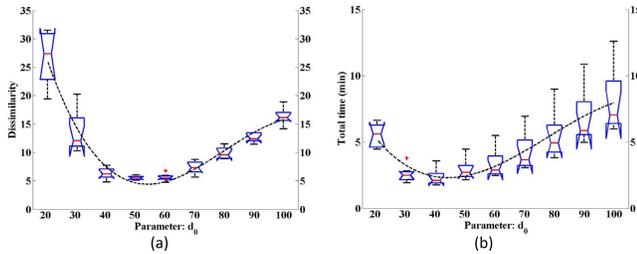
Fig. 19. The box-plot and response curve of (a) dissimilarity and (b) total processing time, with respect to d_0, for the medium resolution TEM images.

is the middle value of the parameter's range. Following the evaluation methodology suggested in [48] and [49], we also fit a 3-degree polynomial of d_0 for its mean values of the dissimilarity and total time; this 3-degree polynomial is shown as the black dashed line in the respective plots.

Figure 19 shows that d_0 does play an importance role in affecting detection quality as well as detection time. If d_0 is chosen too small or too big, both detection quality and processing time will be adversely affected. We also observe that underestimation of d_0 harms the detection quality more than overestimation, while overestimation prolongs the processing time more. Nevertheless, both detection quality and processing time remain reasonably stable when d_0 is chosen between 40 and 60, namely within 20% deviation of the nominal particle size. This range of allowance makes it practical to use a rough estimate of the particle diameter in the proposed method to produce robust detection results.

V. Summary

In this paper, we proposed a new method to detect the nanoparticles in noisy (TEM) images. The main contribution of the paper is that we present a framework leading to robust processing capability. This framework entails two pipelines of processing in parallel, making use of complementary image information, followed by a binary integer optimization procedure to resolve detection conflicts and select better outcomes. Our method can solve the particle detection problem for TEM images with low contrast and heavy noise, making the new method particularly useful in the application of non-metallic nano material analysis.

We want to point out a few possible extensions. When a new pipeline of processing is discovered to complement the existing two processing pipelines, our BIP formulation does allow an extension to include those. What needs to be done is to amend the constraint conditions to incorporate more than two detection outcomes and devise a conflict matrix **M**, making sure that still only one outcome is chosen eventually. Application front, a possible extension of the current work is to explore the method's applicability to bio-image processing like cell detection or object detection from satellite images especially when the quality of the images is low.

References

[1] H. Zheng, R. K. Smith, Y.-W. Jun, C. Kisielowski, U. Dahmen, and A. P. Alivisatos, "Observation of single colloidal platinum nanocrystal growth trajectories," *Science*, vol. 324, no. 5932, pp. 1309–1312, 2009.

[2] J. Watt, S. Cheong, and R. D. Tilley, "How to control the shape of metal nanostructures in organic solution phase synthesis for plasmonics and catalysis," *Nano Today*, vol. 8, no. 2, pp. 198–215, 2013.

[3] T. J. Woehl, C. Park, J. E. Evans, I. Arslan, W. D. Ristenpart, and N. D. Browning, "Direct observation of aggregative nanoparticle growth: Kinetic modeling of the size distribution and growth rate," *Nano Lett.*, vol. 14, no. 1, pp. 373–378, 2013.

[4] C. L. Nehl, H. Liao, and J. H. Hafner, "Optical properties of star-shaped gold nanoparticles," *Nano Lett.*, vol. 6, no. 4, pp. 683–688, 2006.

[5] Y. Pan *et al.*, "Size-dependent cytotoxicity of gold nanoparticles," *Small*, vol. 3, no. 11, pp. 1941–1949, 2007.

[6] C. Park, J. Z. Huang, J. X. Ji, and Y. Ding, "Segmentation, inference and classification of partially overlapping nanoparticles," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 35, no. 3, pp. 669–681, Mar. 2013.

[7] L. Shafarenko, M. Petrou, and J. Kittler, "Automatic watershed segmentation of randomly textured color images," *IEEE Trans. Image Process.*, vol. 6, no. 11, pp. 1530–1544, Nov. 1997.

[8] Y.-C. Lin, Y.-P. Tsai, Y.-P. Hung, and Z.-C. Shih, "Comparison between immersion-based and toboggan-based watershed image segmentation," *IEEE Trans. Image Process.*, vol. 15, no. 3, pp. 632–640, Mar. 2006.

[9] P. Quelhas, M. Marcuzzo, A. M. Mendonça, and A. Campilho, "Cell nuclei and cytoplasm joint segmentation using the sliding band filter," *IEEE Trans. Med. Imag.*, vol. 29, no. 8, pp. 1463–1473, Aug. 2010.

[10] D. Freedman and T. Zhang, "Interactive graph cut based segmentation with shape priors," in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, vol. 1, pp. 755–762, Jun. 2005.

[11] T. F. Chan and L. A. Vese, "Active contours without edges," *IEEE Trans. Image Process.*, vol. 10, no. 2, pp. 266–277, Feb. 2001.

[12] B. Parvin, Q. Yang, J. Han, H. Chang, B. Rydberg, and M. H. Barcellos-Hoff, "Iterative voting for inference of structural saliency and characterization of subcellular events," *IEEE Trans. Med. Imag.*, vol. 16, no. 3, pp. 615–623, Mar. 2007.

[13] O. Schmitt and M. Hasse, "Morphological multiscale decomposition of connected regions with emphasis on cell clusters," *Comput. Vis. Image Understand.*, vol. 113, no. 2, pp. 188–201, 2009.

[14] C. Park *et al.*, "A multistage, semi-automated procedure for analyzing the morphology of nanoparticles," *IIE Trans.*, vol. 44, no. 7, pp. 507–522, 2012.

[15] H. Yang and N. Ahuja, "Automatic segmentation of granular objects in images: Combining local density clustering and gradient-barrier watershed," *Pattern Recognit.*, vol. 47, no. 6, pp. 2266–2279, 2014.

[16] P.-J. De Temmerman, E. Verleysen, J. Lammertyn, and J. Mast, "Semi-automatic size measurement of primary particles in aggregated nanomaterials by transmission electron microscopy," *Powder Technol.*, vol. 261, pp. 191–200, Jul. 2014.

[17] P. Muneesawang and C. Sirisathitkul, "Size measurement of nanoparticle assembly using multilevel segmented tem images," *J. Nanomaterials*, vol. 16, no. 1, p. 58, 2015.

[18] L. L. Hench and J. K. West, "The sol-gel process," *Chem. Rev.*, vol. 90, no. 1, pp. 33–72, 1990.

[19] H. Zhang *et al.*, "Wear-resistant and transparent acrylate-based coating with highly filled nanosilica particles," *Tribol. Int.*, vol. 43, nos. 1–2, pp. 83–91, 2010.

[20] J. Freixenet, X. Muñoz, D. Raba, J. Martí, and X. Cufí, "Yet another survey on image segmentation: Region and boundary information integration," in *Proc. Eur. Conf. Comput. Vis.*, 2002, pp. 408–422.

[21] X. Muñoz, J. Freixenet, X. Cufí, and J. Martí, "Strategies for image segmentation combining region and boundary information," *Pattern Recognit. Lett.*, vol. 24, nos. 1–3, pp. 375–392, 2003.

[22] L.-C. Chen, "Investigation on morphology measurement and evaluation of TiO_2 nanoparticles synthesized by sanss," *J. Alloys Compounds*, vol. 483, nos. 1–2, pp. 366–370, 2009.

[23] T. Pavlidis and Y.-T. Liow, "Integrating region growing and edge detection," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 12, no. 3, pp. 225–233, Mar. 1990.

[24] M. Sato, S. Lakare, M. Wan, A. Kaufman, and M. Nakajima, "A gradient magnitude based region growing algorithm for accurate segmentation," in *Proc. Int. Conf. Image Process.*, vol. 3. 2000, pp. 448–451.

[25] J. Canny, "A computational approach to edge detection," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. PAMI-8, no. 6, pp. 679–698, Nov. 1986.

[26] D. Mumford and J. Shah, "Optimal approximations by piecewise smooth functions and associated variational problems," *Commun. Pure Appl. Math.*, vol. 42, no. 5, pp. 577–685, 1989.

[27] M. Kass, A. Witkin, and D. Terzopoulos, "Snakes: Active contour models," *Int. J. Comput. Vis.*, vol. 1, no. 4, pp. 321–331, 1988.
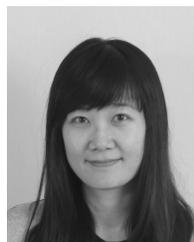
[28] C. Li, C.-Y. Kao, J. C. Gore, and Z. Ding, "Minimization of region-scalable fitting energy for image segmentation," *IEEE Trans. Image Process.*, vol. 17, no. 10, pp. 1940–1949, Oct. 2008.

[29] V. Caselles, R. Kimmel, and G. Sapiro, "Geodesic active contours," *Int. J. Comput. Vis.*, vol. 22, no. 1, pp. 61–79, 1997.

[30] M. S. Allili and D. Ziou, "Globally adaptive region information for automatic color–texture image segmentation," *Pattern Recognit. Lett.*, vol. 28, no. 15, pp. 1946–1956, 2007.

[31] K. Zhang, L. Zhang, H. Song, and W. Zhou, "Active contours with selective local or global segmentation: A new formulation and level set method," *Image Vis. Comput.*, vol. 28, no. 4, pp. 668–676, 2010.

[32] Y. Tian, F. Duan, M. Zhou, and Z. Wu, "Active contour model combining region and edge information," *Mach. Vis. Appl.*, vol. 24, no. 1, pp. 47–61, 2013.

[33] J. A. Hartigan and M. A. Wong, "Algorithm as 136: A K-means clustering algorithm," *J. Roy. Statist. Soc. C, Appl. Statist.*, vol. 28, no. 1, pp. 100–108, 1979.

[34] G. Perrin, X. Descombes, and J. Zerubia, "A marked point process model for tree crown extraction in plantations," in *Proc. IEEE Int. Conf. Image Process.*, vol. 1. Sep. 2005, pp. 661–664.

[35] A. Cichocki and S. Amari, *Adaptive Blind Signal and Image Processing: Learning Algorithms and Applications*. New York, NY, USA: Wiley, 2002.

[36] R. Jain, R. Kasturi, and B. G. Schunck, *Machine Vision*. New York, NY, USA: McGraw-Hill, 1995.

[37] A. Likas, N. Vlassis, and J. J. Verbeek, "The global k-means clustering algorithm," *Pattern Recognit.*, vol. 36, no. 2, pp. 451–461, Feb. 2003.

[38] G. Hamerly and C. Elkan, "Alternatives to the k-means algorithm that find better clusterings," in *Proc. Int. Conf. Inf. Knowl. Manage.*, pp. 600–607, 2002.

[39] N. Otsu, "A threshold selection method from gray-level histograms," *Automatica*, vol. 11, nos. 285–296, pp. 23–27, 1975.

[40] W. R. Taylor, J. M. Thornton, and W. G. Turnell, "An ellipsoidal approximation of protein shape," *J. Molecular Graph.*, vol. 1, no. 2, pp. 30–38, 1983.

[41] F. Herrera and L. Martinez, "A model based on linguistic 2-tuples for dealing with multigranular hierarchical linguistic contexts in multi-expert decision-making," *IEEE Trans. Syst., Man, Cybern. B, Cybern.*, vol. 31, no. 2, pp. 227–234, Apr. 2001.

[42] H. Zhang, J. E. Fritts, and S. A. Goldman, "Image segmentation evaluation: A survey of unsupervised methods," *Comput. Vis. Image Understand.*, vol. 110, no. 2, pp. 260–280, 2008.

[43] R. Fisker, J. M. Carstensen, M. F. Hansen, F. Bødker, and S. Mørup, "Estimation of nanoparticle size distributions by image analysis," *J. Nanoparticle Res.*, vol. 2, no. 3, pp. 267–277, 2000.

[44] K. V. Mardia, J. T. Kent, and J. M. Bibby, *Multivariate Analysis*. New York, CA, USA: Academic, 1980.

[45] U. von Luxburg, "A tutorial on spectral clustering," *Statist. Comput.*, vol. 17, no. 4, pp. 395–416, 2007.

[46] P. M. Narendra and K. Fukunaga, "A branch and bound algorithm for feature subset selection," *IEEE Trans. Comput.*, vol. 100, no. 9, pp. 917–922, Sep. 1977.

[47] H. Scheffé, *The Analysis of Variance*. New York, NY, USA: Wiley, 1999.

[48] F. A. Sadjadi, "Experimental design methodology: The scientific tool for performance evaluation," *Proc. SPIE*, vol. 1310, pp. 100–107, Sep. 1990.

[49] F. A. Sadjadi, "Infrared target detection with probability density functions of wavelet transform subbands," *Appl. Opt.*, vol. 43, no. 2, pp. 315–323, 2004.

Yanjun Qian received the B.S. and M.S. degrees in automation from Tsinghua University, China, in 2009 and 2012, respectively. He is currently pursuing the Ph.D. degree in industrial and systems engineering with Texas A&M University. His research interests are in system informatics and quality and reliability engineering, image and video processing, and machine learning. He is a Student Member of INFORMS.

Jianhua Z. Huang received the B.S. and M.S. degrees in probability and statistics from the Beijing University of China in 1989 and 1992, respectively, and the Ph.D. degree in statistics from the University of California at Berkeley in 1997. He is currently a Professor of Statistics with Texas A&M University. His research interests are in statistical machine learning and applied statistics. He is a fellow of ASA and IMS, and an elected member of the ISI.

Xiaodong Li received the B.S. degree in applied mathematics and the M.S. degree in statistics from Peking University, China, in 2003 and 2006, respectively, and the Ph.D. degree in statistics from Columbia University in 2010. She was an Assistant Professor with the Academy of Mathematics and Systems Sciences, Chinese Academy of Sciences. Her research interests include spatial point pattern and image processing, with special application in nanosciences.

Yu Ding (M'01–SM'11) received the B.S. degree in precision engineering from the University of Science and Technology of China in 1993, the M.S. degree in precision instruments from Tsinghua University, China, in 1996, the M.S. degree in mechanical engineering from Pennsylvania State University in 1998, and the Ph.D. degree in mechanical engineering from the University of Michigan in 2001. He is currently the Mike and Sugar Barnes Professor of Industrial and Systems Engineering and a Professor of Electrical and Computer Engineering with Texas A&M University. His research interests are in system informatics and quality and reliability engineering. He is a fellow of IIE, a fellow of ASME, and a member of INFORMS.