Robust Endoscope Motion Estimation Via an Animated Particle Filter for Electromagnetically Navigated Endoscopy

Xiongbiao Luo*, *Member, IEEE*, and Kensaku Mori, *Member, IEEE*

Abstract—**Electromagnetically navigated endoscopy, which is increasingly applied in endoscopic interventions, utilizes an electromagnetic sensor attached at the endoscope tip to measure the endoscope movements and to navigate the endoscope in the region of interest in the body. Due to patient motion and magnetic field distortion, sensor electromagnetic tracking (EMT) measurement inaccuracy and dynamic jitter errors remain challenging for electromagnetic navigation. This paper proposes a new tracking framework of an animated particle filter that integrates adaptive particle swarm optimization into a generic particle filter to significantly boost electromagnetic trackers. We validate our method on a dynamic phantom and compare it to state-of-the-art EMT methods. Our experimental results demonstrate the effectiveness and robustness of our method, which provides position and orientation accuracy of 2.48 mm, 4.69° that significantly outperformed previous methods at least with tracking error of 4.19 mm, 7.75°. The tracking smoothness was improved from 4.09 mm, 3.37° to 1.84 mm, 2.52°. Our method successfully tackled the particle impoverishment better than standard particle filters.**

Index Terms—**Adaptive particle swarm optimization (APSO), animated particle filter (APF), bronchoscopy navigation, electromagnetically navigated endoscopy (ENE), endoscope 3-D motion tracking.**

I. BACKGROUND

MINIMALLY invasive diagnosis and surgery (MIDS) techniques, which can reduce patient trauma and recovery times, are becoming more and more prevalent in medical communities. As one MIDS technique, surgical endoscopy, where an endoscope is inserted into the body through a natural orifice (e.g., nose or mouth) for organ observation, is a common medical procedure that is performed for cancer diagnosis and treatment: laparoscopy for abdominal disease detection, neurosurgery applied to brain tumor surgery or removal, and colonoscopy adopted to colon or rectum cancer examinations.

However, conventional endoscopic techniques are somewhat limited to medical applications; physicians have difficulty either performing transbronchial lung biopsy to successfully sample suspicious tissues or resecting precancerous tumors inside complex airway trees, since only 2-D bronchoscopic video images are used without any other guidance information (endoscope position and orientation) during bronchoscopic conventions. To address the lack of guidance information in the conventional endoscopy and obtain successful diagnosis and cancer treatment, surgeons require useful multimodal information fusion or visualization guidance tools to assist them to perform safe and secure operations during the endoscopic examinations.

As a result of recent technological advances in augmented reality (AR), real-time endoscopic imaging, and sensor techniques, endoscopy navigation systems can be constructed to compensate conventional endoscopy and extend its functions or abilities for medical applications [4], [5], [9], [10], [15], [26], [27]. Such endoscopy systems fuse pre- and intraoperative imaging information, e.g., 3-D computed tomography (CT) or magnetic resonance (MR) data and endoscopic video sequences, providing AR guidance environments with visual information by comparatively displaying live video images and CT- or MR-driven virtual renderings by indicating the relative positions between endoscopes or surgical tools and target tumors. To realize endoscopy navigation, one vital problem must be resolved: accurate and stable synchronization between prebuilt 3-D anatomical models and real time 2-D endoscopic video images. Unfortunately, it remains challenging to precisely and stably coordinate pre- and intraoperative imaging spaces when using endoscopy navigation in operating rooms, because it is somewhat difficult to accurately and robustly track endoscope movements for the current available methods, e.g., electromagnetic tracking (EMT) techniques.

The remainder of this paper is organized as follows. We discuss EMT techniques in Section II. Our animated particle filter (APF) framework is described in Section III. Sections IV and V show evaluation strategies and some details of our experimental setups. Section VI compares evaluation results. We discuss our method in Section VII before concluding this paper in Section VIII.

*X. Luo is with the Information and Communications Headquarters, Nagoya University, Nagoya 464-8603, Japan (e-mail: xiongbiao.luo@gmail.com).

K. Mori is with the Information and Communications Headquarters, Nagoya University, Nagoya 464-8603, Japan (e-mail: kensaku@is.nagoya-u.ac.jp).
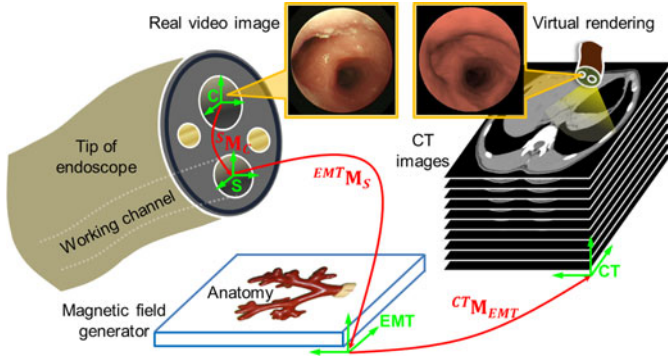
Fig. 1. ENE involves coordinate systems $C, S,$ EMT, and CT, respectively, corresponding to endoscopic camera, EMT sensor, EMT system, and CT space. Relationships among these coordinate systems can be represented by ${}^{\mathrm{CT}}\mathbf{M}_C = {}^{\mathrm{CT}}\mathbf{M}_{\mathrm{EMT}}{}^{\mathrm{EMT}}\mathbf{M}_S{}^{S}\mathbf{M}_C$.

II. ELECTROMAGNETICALLY NAVIGATED ENDOSCOPY

Electromagnetically navigated endoscopy (ENE) is now increasingly introduced to cancer diagnosis and treatment [2], [8]. During such endoscopic navigations, an EMT sensing coil (sensor) with six degrees of freedom (6 DoF) measurements is usually attached to the endoscope working channel and localized by a magnetic field generating system, such as a commercially available superDimension (superDimension Inc., Minneapolis, MN, USA) navigation system [24]. Sensor outputs are utilized to synchronize EMT and preoperative image (here refer to CT) coordinate systems. Hence, endoscope location in the CT space can be determined nominally during endoscopic procedures (see Fig. 1).

Although EMT-based methods provide global synchronization between patients and their CT image coordinate systems since they can follow the immediate movements of an endoscope inside the body [14], [31], such techniques still suffer from two major drawbacks: 1) inaccurate location due to such patient movements as respiratory motion; since EMT sensor outputs provide position and orientation measurements relative to a fixed, world coordinate system, they do not correspond exactly to the current camera location and 2) unstable measurements with stochastic jitter due to the magnetic field distortion that is caused by ferrous metals or conductive material in the working volume. Because of these two drawbacks, ENE is somewhat limited to the diagnostic yield; e.g., its indicative accuracy only ranges between 59% and 74% [25].

Various previous works have addressed these two drawbacks. To correct jitter measurements caused by stationary objects, many calibration techniques have been investigated [13] that usually collect distorted and undistorted measurements to generate a field distortion function in terms of lookup tables or polynomials. However, such functions cannot be utilized for different setups in operating rooms before endoscopic procedures and dynamic changes of the magnetic field distortion because of the relocation of the field generator or the movement of instruments during interventions. Mori *et al.* [19] proposed a hybrid tracking method by integrating EMT and image registration techniques that are usually criticized since they fail to easily track the global trajectory of the endoscope due to image

artifacts (e.g., bubbles). Their hybrid method was enhanced by Kalman filtering and a respiratory motion compensation method using a surrogate sensor [28]. Gergel *et al.* [7] tackled airway deformation using particle filtering and bronchial structures, assuming that an endoscope is always moving along the airway centerlines. However, such an assumption is broken easily in clinical applications. On the other hand, EMT output jitters are difficult to correct, unless they are combined with optical tracking [6].

Even though many papers have tackled the drawbacks of ENE, a more robust and accurate EMT-based endoscope 3-D motion estimation approach is still needed to improve ENE diagnostic accuracy during interventions. This paper partially extends our previous work [16], [18]. We modified a generic particle filter to our APF, which not only utilizes adaptive particle swarm optimization (APSO) to tackle the particle impoverishment problem but also a new likelihood model to precisely characterize endoscopic video images to enhance particle robustness. We proposed a fully automatic visualization assessment to evaluate our tracking results from different methods to replace our previous manual inspection method. Furthermore, thorough evaluation and comparison were performed on our dynamic phantom. Our proposed method clarifies the following highlights.

1) Technically, our APF framework provides a promising means to effectively fuse raw sensor measurements and endoscopic video images to achieve more robust and accurate endoscope tracking than all previous methods. Our APF-based multimodal tracking can successfully deal with the drawbacks of inaccurate location and jitter measurements in EMT procedures. Our experimental results show that high tracking performance can be obtained with our APF strategy, which also provides an effective scheme to fuse other external tracking sources in navigated endoscopy.

2) Our method also successfully tackles the particle diversity loss problem in standard particle filters. We introduced APSO to address the diversity loss problem, to alleviate particle impoverishment, and to obtain particle diversity in APSO iterations to enhance the particle exploration ability in APF. On the other hand, for evolutionary methods, we presented APSO that enables standard particle swarm optimization (PSO) to automatically control evolutionary parameters, which are estimated on the basis of spatial constraints and endoscopic video image intensity information.

3) From an application point of view, to the best of our knowledge, our APSO-involved framework is a novel PSO application for endoscope motion tracking in computer-assisted intervention communities. Additionally, our APF approach can potentially be applied to many endoscopic guidances, e.g., colonoscopes.

III. APPROACHES: ANIMATED PARTICLE FILTER

This paper seeks to accurately and robustly estimate endoscope motion in accordance with the EMT sensors (with 6 DoF position and orientation information) and an endoscopic video

camera (with 2-D endoscopic image information) during endoscopy. The challenge for endoscope motion estimation lies in the measurement or the data distortion; i.e., sensor output uncertainty (inaccuracy and jitter errors as discussed in Section II) and image ambiguity or artifact (e.g, motion blurring), which commonly occur in interventions, hinder the successful registration between pre- and intra-operative information spaces. Due to measurement distortion, endoscope tracking is a dynamic, multimodal, and nonlinear procedure.

Particle filter, a sequential Monte Carlo (SMC) technique, has the capacity to avoid local minima convergence and applicability to uncertainty or non-Gaussian data and provides an effective way to recover measurement distortion in a dynamic system; one major problem of particle impoverishment often results in a loss of particle diversity. In this paper, we propose an APSO strategy to tackle the particle impoverishment problem and develop an APF algorithm to deal with measurement distortion or uncertainty in EMT sensor outputs and video images.

A. Motion Parameterization

Endoscopic camera motion parameters ${}^{\mathrm{CT}}\mathbf{M}_C$ include position ${}^{\mathrm{CT}}\mathbf{t}_C$ and rotation ${}^{\mathrm{CT}}\mathbf{R}_C$ from the camera coordinate system to the CT coordinate system. Position vector ${}^{\mathrm{CT}}\mathbf{t}_C$ is presented by ${}^{\mathrm{CT}}\mathbf{t}_C = \left({}^{\mathrm{CT}}t_C^x, {}^{\mathrm{CT}}t_C^y, {}^{\mathrm{CT}}t_C^z\right)$, where components ${}^{\mathrm{CT}}t_C^x$, ${}^{\mathrm{CT}}t_C^y$, and ${}^{\mathrm{CT}}t_C^z$ are the coordinate values of the camera position in the x-, y-, and z-axes of the CT space. For the rotation part, we use the quaternion but not the rotation matrix ${}^{\mathrm{CT}}\mathbf{R}_C$ to describe the rotation part in our implementation since the quaternion can very powerfully characterize the rotation part in terms of such advantages as compactness and the avoidance of discontinuous jumps compared to other representations (e.g., Euler angles).

A quaternion representation of rotation can be conveniently considered a normalized vector with four components:

$$\mathbf{q} = [\; q_0 \quad q_x \quad q_y \quad q_z \;], \qquad q_0{}^2 + q_x{}^2 + q_y{}^2 + q_z{}^2 = 1. \tag{1}$$

In our APF framework, camera motion state or particle $\mathbf{p}_i$ can be parameterized by a 7-D vector

$$ {}^{\mathrm{CT}}\mathbf{M}_C \;\longmapsto\; \mathbf{p}_i = \left({}^{\mathrm{CT}}\mathbf{t}_C^i \quad {}^{\mathrm{CT}}\mathbf{q}_C^i \right)^T \tag{2}$$

where i means the camera motion state at time i or denotes the ith 6 DoF EMT senor measurement in endoscope tracking. Hence, particle $\mathbf{p}_i$ is a 7-D vector to parameterize the 6 DoF camera pose with the camera position and rotation.

Additionally, for each particle, we define its weight w_i^j and cumulative weight c_i^j ($j = 1, 2, \ldots, M, j$ is the particle index and M is the particle number). One particle can be represented by $(\mathbf{p}_i^j, w_i^j, c_i^j)$, which also denotes a potential solution of current bronchoscope camera motion parameters ${}^{\mathrm{CT}}\mathbf{M}_C^i$.

B. Particle Randomness

Before particle animation, a particle randomness step is performed to diversify initial set $\mathcal{P}_0 = \{(\mathbf{p}_0^j, w_0^j, c_0^j)\}$ for $\mathcal{P}_i = \{(\mathbf{p}_i^j, w_i^j, c_i^j)\}$ at time i. Initialized particle set $\mathcal{P}_0$ is basically randomized by a Gaussian transition model. Randomized
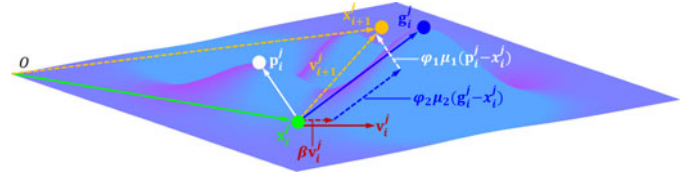

Fig. 2. Particle $\mathbf{x}_i^j$ in our proposed APF framework is animated to $\mathbf{x}_{i+1}^j$ by *inertial*, *social*, and *cognitive* elements, $\beta\mathbf{v}_i^j$, $\varphi_2\mu_2(\mathbf{g}_i^j - \mathbf{x}_i^j)$, and $\varphi_1\mu_1(\mathbf{p}_i^j - \mathbf{x}_i^j)$ that correspond to *red*, *blue*, and *white* dashes, respectively.

particle set $\mathcal{X}_i = \{(\mathbf{x}_i^j, w_i^j, c_i^j)\}$ is generated by

$$\mathbf{x}_i^j \;\sim\; \mathcal{N}(\mathbf{p}_0^j, \Sigma) \tag{3}$$

where Σ is a fixed noise term. We then update the particles in $\mathcal{P}_i$ from $\mathcal{X}_i$ in terms of the particle weights or the fitnesses (8). Note that such randomness is only performed once: at frame $i = 1$.

After obtaining swarms $\mathcal{X}_i$ and $\mathcal{P}_i$, our method moves $\mathcal{X}_i$ to $\mathcal{X}_{i+1}$ and transmits $\mathcal{P}_i$ deterministically and stochastically to approximate the posterior density of motion state $\mathbf{p}_{i+1}$ at frame $(i + 1)$ in accordance with three steps: 1) particle animation, 2) state propagation, and 3) likelihood model. They are described in the following three sections.

C. Particle Animation: APSO

PSO, which was originally presented in [12], is a multiple-population parallel search algorithm that makes conceptual particles fly through nonlinear and high-dimensional solution spaces. Its major attractiveness is its ability to tackle nonlinear, nondifferentiable, and multimodal optimization problems by dynamically interacting all particles in a similar analogy with the *cognitive* and *social* properties of populations [23].

Each particle $\mathbf{x}_i^j$ with its velocity $\mathbf{v}_i^j$ that is moved to new state $\mathbf{x}_{i+1}^j$ and velocity $\mathbf{v}_{i+1}^j$ can be presented by a matrix equation

$$\begin{pmatrix} \mathbf{x}_{i+1}^j \\ \mathbf{v}_{i+1}^j \end{pmatrix} = \begin{pmatrix} 1-\alpha & \beta \\ -\alpha & \beta \end{pmatrix} \cdot \begin{pmatrix} \mathbf{x}_i^j \\ \mathbf{v}_i^j \end{pmatrix} + \begin{pmatrix} \alpha \\ \alpha \end{pmatrix} \cdot \Upsilon \tag{4}$$

$$\alpha = \frac{\varphi_1\mu_1 + \varphi_2\mu_2}{2}, \quad \Upsilon = \frac{\varphi_1\mu_1\mathbf{p}_i^j + \varphi_2\mu_2\mathbf{g}_i^j}{\varphi_1\mu_1 + \varphi_2\mu_2} \tag{5}$$

where inertial weight β indicates how much $\mathbf{v}_i^j$ will be preserved in $\mathbf{v}_{i+1}^j$ and $\mathbf{p}_i^j$ and $\mathbf{g}_i^j$ are the best states found by particle j so far and the best state found by the whole swarm so far. Uniformly distributed random numbers μ_1 and μ_2, respectively, control the effect of individual best $\mathbf{p}_i^j$ and global all best $\mathbf{g}_i^j$ on new velocity $\mathbf{v}_{i+1}^j$, and acceleration coefficients φ_1 and φ_2, respectively, show the self- and swarm-confidences [3]. Fig. 2 further illustrates how particle $\mathbf{x}_i^j$ is animated.

Since evolutionary factors β, φ_1, and φ_2 determine the PSO performance, φ_1, φ_2 can adaptively calculated by [32]

$$\varphi_1 = \frac{2\Psi(\mathbf{p}_i^j)}{\Psi(\mathbf{p}_i^j) + \Psi(\mathbf{g}_i^j)}, \quad \varphi_2 = \frac{2\Psi(\mathbf{g}_i^j)}{\Psi(\mathbf{p}_i^j) + \Psi(\mathbf{g}_i^j)} \tag{6}$$

where $\Psi(\cdot)$ is a fitness value that is computed by (19).

Inertia weight β can be calculated on spatial constraints and image information. We compute distance d_i^j from one particle to all other particles and find $d_{\min}$ and $d_{\max}$ from $\{d_i^j\}_{j=1}^M$. We then normalize distance d_g^j between one particle and the current global best particle with $d_{\min}$ and $d_{\max}$, obtain $\gamma_i^j = (d_g^j - d_{\min})/(d_{\max} - d_{\min})$, and assign γ_i^j to each particle. Finally, since β was suggested within interval $[0.4\ 0.9]$ for weighting the global and local searching abilities [23], it can be adaptively computed by

$$\beta(\Psi(\mathbf{x}_i^j), \gamma_i^j) = \frac{2}{2 + 3\exp(-1.28(\Psi(\mathbf{x}_i^j) + \gamma_i^j))} \quad (7)$$

which shows a novel strategy to automatically determine β by integrating spatial continuity constraint γ_i^j and endoscopic image information $\Psi(\mathbf{x}_i^j)$ that will be computed on the basis of video image intensity into our modified PSO algorithm.

Eventually, we update $\mathbf{g}_{i+1}^j$ and each particle in $\mathcal{P}_i$ to $\mathcal{P}_{i+1}$ in accordance with $\mathcal{X}_{i+1}$ by the following equations:

$$\mathbf{p}_{i+1}^j = \begin{cases} \mathbf{x}_{i+1}^j, & \text{if } \Psi(\mathbf{x}_{i+1}^j) > \Psi(\mathbf{p}_i^j) \\ \mathbf{p}_i^j, & \text{otherwise} \end{cases} \quad (8)$$

$$\mathbf{g}_{i+1}^j = \arg\max_{\mathbf{p}_{i+1}^j} \Psi(\mathbf{p}_{i+1}^j). \quad (9)$$

D. State Propagation

After getting animated set $\mathcal{P}_{i+1} = \{(\mathbf{p}_{i+1}^j, w_{i+1}^j, c_{i+1}^j)\}$, we resample $\mathcal{P}_{i+1}$ and obtain set $\hat{\mathcal{P}}_{i+1} = \{(\hat{\mathbf{p}}_{i+1}^j, \hat{w}_{i+1}^j, \hat{c}_{i+1}^j)\}$ that has deterministically drifted and stochastically diffused to new set $\tilde{\mathcal{P}}_{i+1} = \{(\tilde{\mathbf{p}}_{i+1}^j, \tilde{w}_{i+1}^j, \tilde{c}_{i+1}^j)\}$ [11]

$$\tilde{\mathbf{p}}_{i+1}^j = \mathcal{F}(\mathbf{A}\hat{\mathbf{p}}_{i+1}^j, \mathbf{B}(n_i^j)) \quad (10)$$

where $\mathcal{F}$ is a transition function and stochastic diffusion part $\mathbf{B}(n_i^j)$ is a noise term with independent stochastic variable n_i^j, which was assumed to yield a normal distribution in our implementation. Deterministic drift part $\mathbf{A}$ is calculated on the basis of consecutive EMT sensor outputs $\hat{\mathbf{M}}_i$ and $\hat{\mathbf{M}}_{i+1}$

$$\mathbf{A} = \hat{\mathbf{M}}_{i+1}(\hat{\mathbf{M}}_i)^{-1}. \quad (11)$$

Since n_i^j is assumed to follow a normal distribution, state transition density $Pr(\tilde{\mathbf{p}}_{i+1}^j | \hat{\mathbf{p}}_{i+1}^j)$ can be computed by [1]

$$Pr(\tilde{\mathbf{p}}_{i+1}^j | \hat{\mathbf{p}}_{i+1}^j) \propto \exp\left(-\mathbf{B}^{-1}(\mathcal{F}^{-1} - \mathbf{A}\hat{\mathbf{p}}_{i+1}^j)^2/2\right). \quad (12)$$

E. Likelihood Model

After particle propagation and obtaining $\{\tilde{\mathbf{p}}_{i+1}^j\}_{j=1}^M$, we must determine observation probability $Pr(\mathbf{y}_{i+1} | \tilde{\mathbf{p}}_{i+1}^j)$ ($\mathbf{y}_{i+1}$ is an observation variable), particle weight $\tilde{w}_{i+1}^j$ (or fitness value $\Psi(\tilde{\mathbf{p}}_{i+1}^j)$), and cumulative weight $\tilde{c}_{i+1}^j$. In this paper, $Pr(\mathbf{y}_{i+1} | \tilde{\mathbf{p}}_{i+1}^j)$ is defined as an image intensity likelihood function on the basis of a modified structural similarity measure (MoSSM) from the structural similarity [30]. As mentioned in previous work [30], it is more consistent to locally compute similarity for separate patches in an image than directly and globally calculating the similarity for one whole image, since

human observers are usually more sensitive to local statistical features. In an endoscopic video, e.g., bronchoscopic images, specific shapes like bifurcations, and folds are easily noticed by observers. Hence, extracting these shapes and computing similarity only in the shape regions is very useful.

First, we separate endoscopic video image $\mathbf{I}_R^{i+1}$ with size $W \times H$ into $K \times U$ patches $\{D_{k,u}^{i+1}(a, b)\}$ presented by

$$D_{k,u}^{i+1}(a, b) = \begin{pmatrix} (k-2)\dfrac{W}{K} \le a \le (k+1)\dfrac{W}{K} \\ (u-2)\dfrac{H}{U} \le b \le (u+1)\dfrac{H}{U} \end{pmatrix} \quad (13)$$

where k and u range with $2 \le k \le K-1$ and $2 \le u \le U-1$.

Next, we select patches from $\{D_{k,u}^{i+1}\}$ for the local similarity computation in terms of its two intensity-based features: standard deviation $\delta_{D_{k,u}^{i+1}}$ and brightness $\lambda_{D_{k,u}^{i+1}}$:

$$\delta_{D_{k,u}^{i+1}} = \sqrt{\frac{1}{\left|D_{k,u}^{i+1}\right|} \sum_{(a,b)\in D_{k,u}^{i+1}} \left(\mathbf{I}_R^{i+1}(a,b) - \hat{\mathbf{I}}_R^{i+1}\right)^2} \quad (14)$$

$$\lambda_{D_{k,u}^{i+1}} = \frac{1}{\left|D_{k,u}^{i+1}\right|} \sum_{(a,b)\in D_{k,u}^{i+1}} \Omega_{k,u}\left(\mathbf{I}_R^{i+1}(a,b)\right) \quad (15)$$

where $\hat{\mathbf{I}}_R^{i+1}$ is the image intensity mean. Function $\Omega_{k,u}(\cdot)$, which depends on the pixel color information of saturation $H(a, b)$ and lightness $L(a, b)$ in the hue–saturation–lightness (HSL) color model, is used to evaluate whether pixel (a, b) belongs to the highlights and lighter areas

$$\Omega_{k,u}(\mathbf{I}_R^{i+1}(a,b)) = \begin{cases} 1, & H(a,b) \le \pi_H \text{ and } L(a,b) \ge \delta_L \\ 0, & \text{otherwise} \end{cases} \quad (16)$$

where π_H and δ_L are two predetermined thresholds. We remove the white patches without structural information by $\lambda_{D_{k,u}^{i+1}} > \lambda_*$ (a fixed constant), descendingly sort the remaining patches in terms of $\delta_{D_{k,u}^{i+1}}$, and choose $\eta \cdot K \cdot U$ patches $\{S_{k,u}^{i+1}\}$ for the similarity computation.

Now similarity η_{i+1}^j between $\mathbf{I}_R^{i+1}$ and virtual rendering $\mathbf{I}_V(\tilde{\mathbf{p}}_{i+1}^j)$ in the chosen regions $\{S_{k,u}^{i+1}\}$ can be computed by

$$\eta_{i+1}^j = \text{MoSSM}(\ \mathbf{I}_R^{i+1},\ \mathbf{I}_V(\tilde{\mathbf{p}}_{i+1}^j)\ )$$

$$= \frac{1}{\left|S_{k,u}^{i+1}\right|} \sum_{D_{k,u}^{i+1} \in S_{k,u}^{i+1}} \frac{1}{\left|D_{k,u}^{i+1}\right|} \sum_{D_{k,u}^{i+1}} Q \quad (17)$$

where $|S_{k,u}^{i+1}|$ and $|D_{k,u}^{i+1}|$ are the numbers of the selected patches and pixels in the detected region $D_{k,u}^{i+1}$, respectively. We define Q as the patch index in terms of a previously robust image measure [30], which significantly outperformed a mean squared error-based function, and compute Q by

$$Q = \frac{(2\xi_{i+1,r}\xi_{i+1,v} + C_1)(2\sigma_{i+1,r,v} + C_2)}{(\xi_{i+1,v}^2 + \xi_{i+1,r}^2 + C_1)(\sigma_{i+1,r}^2 + \sigma_{i+1,v}^2 + C_2)} \quad (18)$$

where $\sigma_{i+1,r,v}$ is the correlation between $\mathbf{I}_R^{i+1}$ and $\mathbf{I}_V(\tilde{\mathbf{p}}_{i+1}^j)$ at patch $D_{k,u}^{i+1}$; $\xi_{i+1,r}$ and $\xi_{i+1,v}$ are the mean intensity values; $\sigma_{i+1,r}$ and $\sigma_{i+1,v}$ are the intensity variances.
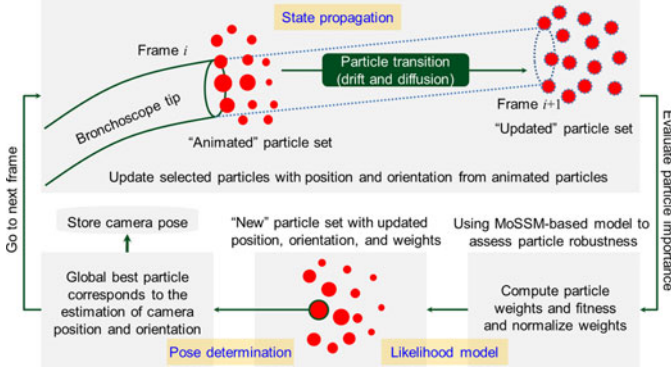
Fig. 3. Processing of state propagation, particle evaluation by our MoSSM-based likelihood model, and pose determination in our APF algorithm.

Fig. 4. Our dynamic phantom with a rubber bronchi model fixed at an acrylic box, a servo motor, and nylon threads.

We now compute likelihood $Pr(\mathbf{y}_{i+1}|\tilde{\mathbf{p}}_{i+1}^j)$, weight $\tilde{w}_{i+1}^j$, and fitness Ψ_{i+1}^j and update cumulative weight $\tilde{c}_{i+1}^j$

$$Pr(\mathbf{y}_{i+1}|\tilde{\mathbf{p}}_{i+1}^j) = \tilde{w}_{i+1}^j = \Psi_{i+1}^j = \exp\left(\eta_{i+1}^j\right) \quad (19)$$

$$\tilde{c}_{i+1}^j = \tilde{c}_i^j + \tilde{w}_{i+1}^j, \tilde{c}_{i+1}^j = \tilde{c}_{i+1}^j \left(\sum_{j=1}^{M} \tilde{c}_{i+1}^j\right)^{-1}. \quad (20)$$

F. Pose Determination

Our goal is to estimate the endoscope 6 DoF position and the orientation parameters. We integrate the EM sensor measurements and the endoscopic images into our APF algorithm. Based on (13)–(19), we compute weight $\tilde{w}_{i+1}^j$ or fitness Ψ_{i+1}^j for each particle in $\{\tilde{\mathbf{p}}_{i+1}^j\}_{j=1}^M$. The best solution to the current endoscope pose corresponds to the position and orientation of the particle with maximal weight or fitness. Hence, the APF output is global best solution $\mathbf{g}_{i+1}^*$ that is found from $\{\tilde{\mathbf{p}}_{i+1}^j\}_{j=1}^M$ for current pose $^{\mathrm{CT}}\mathbf{M}_C^{i+1}$ at frame $(i+1)$

$$\mathbf{g}_{i+1}^* = \arg\max_{\tilde{w}_{i+1}^j}\{\tilde{\mathbf{p}}_{i+1}^j\}, \ \mathbf{g}_{i+1}^* \longmapsto \ ^{\mathrm{CT}}\mathbf{M}_C^{i+1}. \quad (21)$$

Fig. 3 summarizes the three steps of our APF method.

IV. EVALUATION CRITERION

A. Accuracy

We investigate the tracking accuracy of different methods by calculating position error ζ and orientation error ϕ

$$\begin{cases} \zeta = \|\mathbf{t} - \mathbf{t}_G\|, \\ \phi = \arccos((\mathrm{trace}(\mathbf{R}\mathbf{R}_G{}^T) - 1)/2), \end{cases} \quad (22)$$

where ζ denotes the Euclidean distance between ground truth position $\mathbf{t}_G$ and estimated position $\mathbf{t}$ and ϕ indicates the rotation difference between ground truth matrix $\mathbf{R}_G$ and estimated rotation matrix $\mathbf{R}$ by various tracking methods.

B. Smoothness

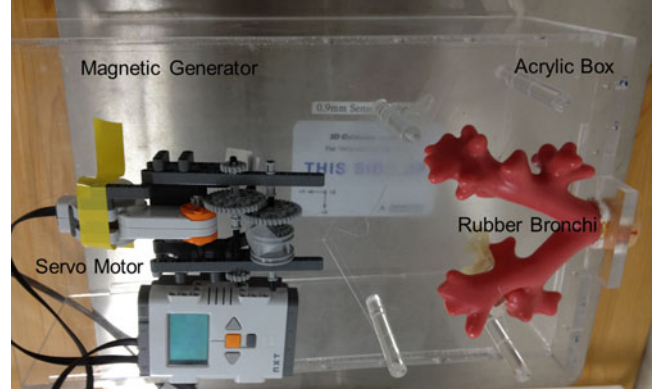Smooth tracking implies little jitter or jump involved in the endoscope movement estimates and might provide two main advantages: 1) reduction of physician discomfort and distraction during operations and 2) enhancement of such posterior procedures as respiratory motion modeling and clinical comparison by temporally and spatially aligning different video sequences from several examinations on the same patient.

Smoothness, which is defined as the average Euclidean distance and the standard deviation of the estimated positions among continuous frames as well as orientations, is calculated by

$$\tau = \frac{\sum_{i=1}^{N-1}\|\mathbf{t}_{i+1} - \mathbf{t}_i\|}{N-1} \quad (23)$$

$$\psi = \frac{\sum_{i=1}^{N-1}\arccos((\mathrm{trace}(\mathbf{R}_{i+1}\mathbf{R}_i{}^T) - 1)/2)}{N-1} \quad (24)$$

where N is the frame number. Large τ and ψ usually show that greater jitter or jump occurred in the estimated motion results.

C. Visualization Quality

It is very important to properly interpret tracking results, which are usually stored as numbers, to enable surgeons to easily and directly visualize tracking results during online interventions in operating rooms. Fortunately, camera position and orientation estimates can be visually described by the virtual images generated by such visualization techniques as volume rendering [20]. Hence, it is natural to evaluate the visualization quality by comparing the virtual images to the endoscopic video images to investigate whether a method is more accurate or robust than another.

Our visual quality assessment is performed by computing two values of similarity and sensitivity. According to a universal image quality index [29], similarity χ between one virtual image and an endoscopic video image can be calculated by

$$\chi = \frac{1}{2}\left(1 + \frac{4\vartheta_{r,v}\kappa_r\kappa_v}{(\vartheta_r^2 + \vartheta_v^2)(\kappa_r^2 + \kappa_v^2)}\right) \quad (25)$$

where $\vartheta_{r,v}$ is the correlation between the virtual and video images, ϑ_r and ϑ_v are the covariance, and κ_r and κ_v are the average values.
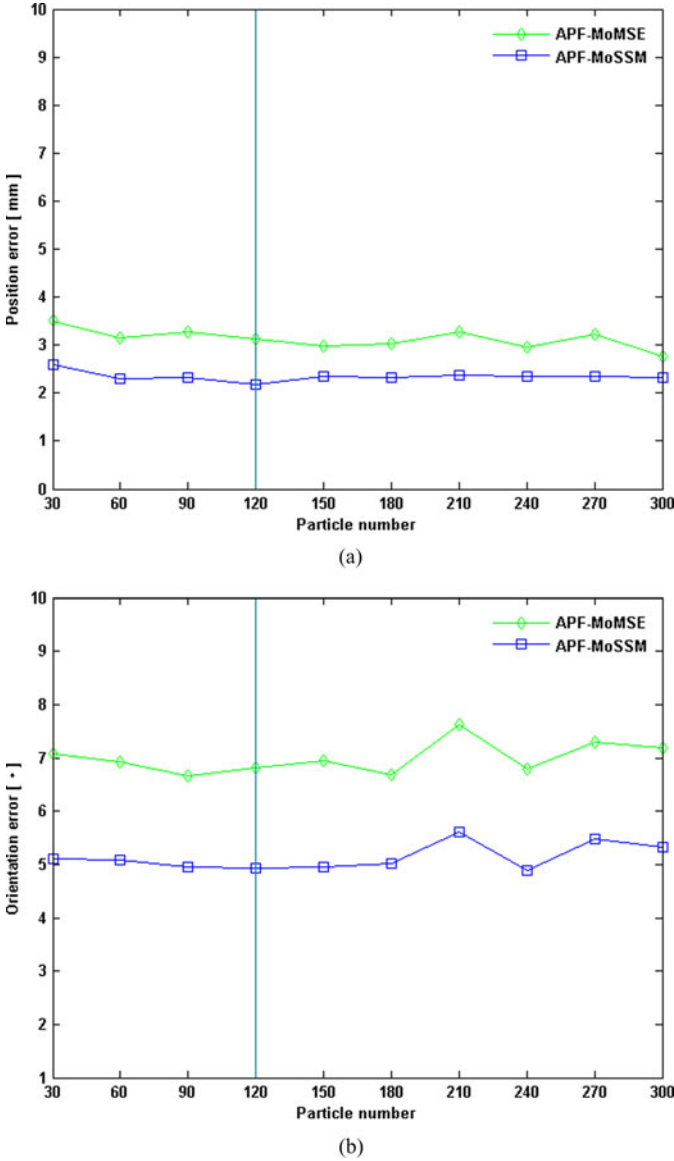
Fig. 5.   Illustration of tracking performance of our APF-based method using different particle numbers. (a) Position error. (b) Orientation error.
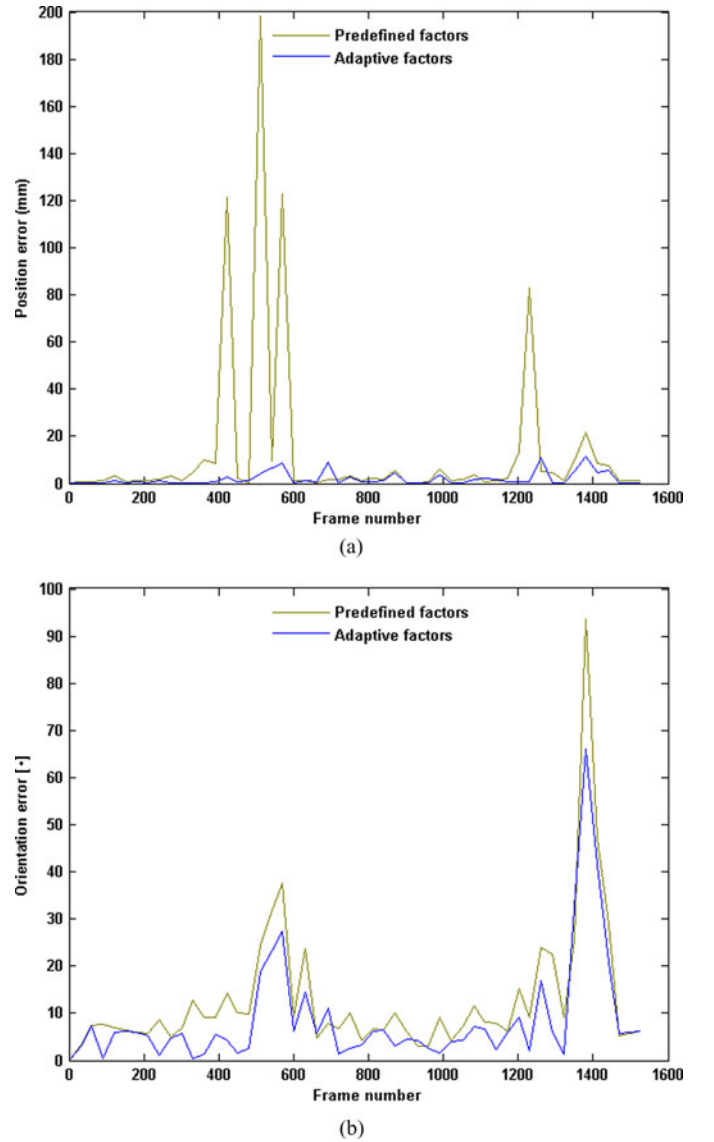

Fig. 6.   Comparison of tracking error from predetermined and adaptive evolutionary factors in PSO-based particle animation. (a) Position error ρ_e. (b) Orientation error $\hat{\chi}_e$.

Based on χ, we empirically predetermine threshold ϵ and decide whether good or bad quality is achieved at frame i

$$C_i = \begin{cases} 1, & \text{if } \chi_i \geq \epsilon \\ 0, & \text{otherwise.} \end{cases} \qquad (26)$$

We add C_i from all frames and compute sensitivity ratio $\rho_e = N^{-1} \sum_{i=1}^{N} C_i$ in each experiment (e is the experiment indication). We also calculate average similarity ratio $\chi_e = N^{-1} \sum_{i=1}^{N} \chi_i$ for each experiment. Additionally, for each method, we compute the average sensitivity and similarity values, $\tilde{\rho} = E^{-1} \sum_{e=1}^{E} \rho_e$ and $\tilde{\chi} = E^{-1} \sum_{e=1}^{E} \chi_e$, which are also used to evaluate different tracking methods.

V. EXPERIMENTAL SETUPS

We evaluated our method on a dynamic phantom that can simulate respiratory motion with a breath rate range of 0–10 per minute and a motion domain of 0–24 mm (see Fig. 4). The phantom CT acquisition parameters were 512×512 pixels, 1021 slices, a 0.68-mm reconstruction pitch, and 0.5-mm thick slices. One drawback of our dynamic phantom must be noted; its breath rates are somewhat unrealistic, compared to human breath rates of 14–30. Our previous work showed more details about constructing our dynamic phantom [17]. We performed 25 experiments with ten ground truth datasets that were generated based on a previous method [16].

A bronchoscope (BF-P260F, Olympus, Tokyo) was used to collect 362×370 pixel, endoscopic video images at 30 frames/s. A 3-D Guidance medSAFE tracker with a 9-coil flat type transmitter (Ascension Technology Corporation, Milton, VT, USA) was employed as our EMT system. Additionally, we performed camera and hand-eye calibrations and CT-to-physical space registration to align the EMT and CT coordinate systems, as previously discussed [17].
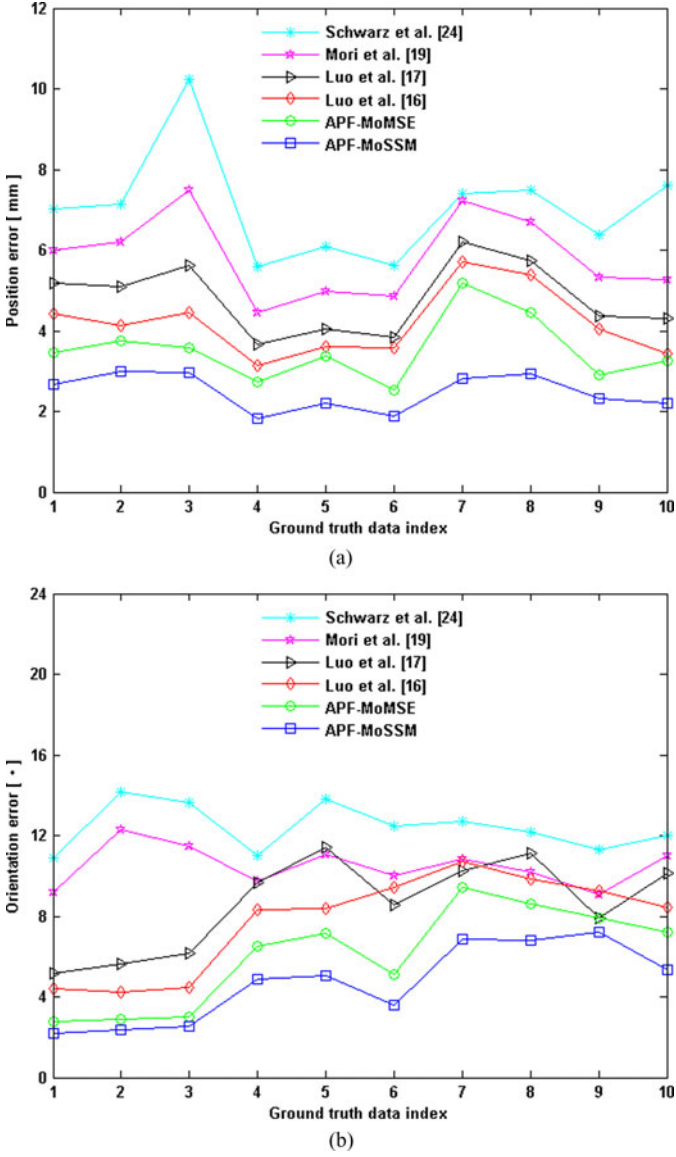
Fig. 7. Tracking accuracy of six methods plotted against ten ground truth datasets. (a) Position error. (b) Orientation error.
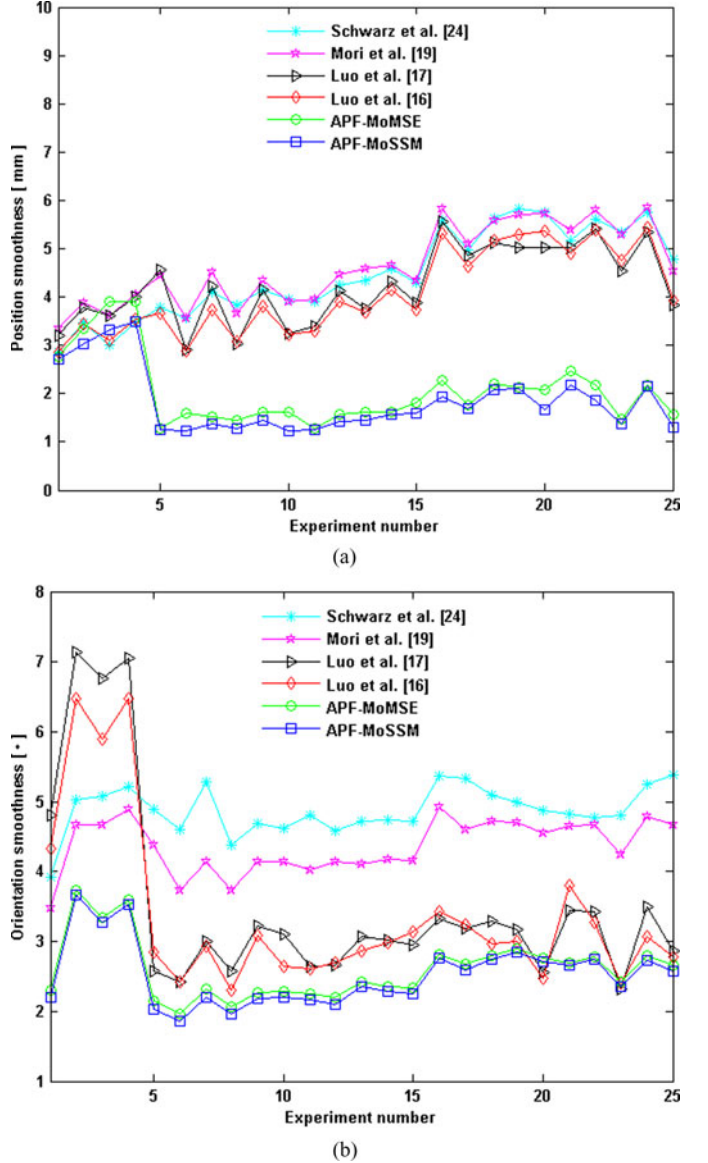

Fig. 8. Plotted smoothness against tracking results obtained from six methods on 25 datasets. Small smoothness of τ and ψ means that one method can track the endoscope more smoothly than another. (a) Position smoothness τ. (b) Orientation smoothness ψ.

For a lateral comparison of our proposed approach, we investigated several currently available methods: 1) Schwarz *et al.* [24], only using an EMT-based method, 2) Mori *et al.* [19], a hybrid tracking scheme, 3) Luo *et al.* [17], an improved hybrid method by image registration and selective EMT sensor outputs, 4) Luo *et al.* [16], using SMC sampling to enhance the EMT-based method, 5) APF-MoMSE, our APF method with a modified mean square error (MoMSE)-based likelihood model, and 6) APF-MoSSM, our APF method with MoSSM-based likelihood model, as presented in Section III.

Additionally, during the likelihood model (see Section III-E), we set the following parameters: $K = U = 30, \eta = 0.3, \pi_H = 0.6, \delta_L = 0.7,$ and $\lambda_* = 0.9$ [4].

VI. COMPARATIVE RESULTS

In our validation, we experimentally determined particle number M. Fig. 5 shows the tracking error under different particle numbers. We choose $M = 120$. Fig. 6 displays the tracking error of PSO and APSO using prefixed and adaptive evolutionary factors. We predefined acceleration coefficients $\varphi_1 = \varphi_2 = 2.0$ and inertia weight $\beta = 1.0$ for PSO. Fig. 7 plots the tracking errors of the six methods against ten ground truth datasets. The smoothness, sensitivity, and similarity of the six methods evaluated on 25 datasets are shown in Figs. 8 and 9. Table I summarizes the tracking results by computing accuracy, smoothness, and visualization index for each method validated on ten ground truth datasets. The average tracking error (position and orientation) was reduced from at least 4.19 mm, 7.75° to 2.48 mm, 4.69°. The smoothness improved from 4.09 mm, 3.37° to 1.84 mm, 2.52°. The visualization quality of the sensitivity and similarity also increased from 74.6%, 0.60 to 85.9%, 0.83. The tracking performance of our APF-based method was
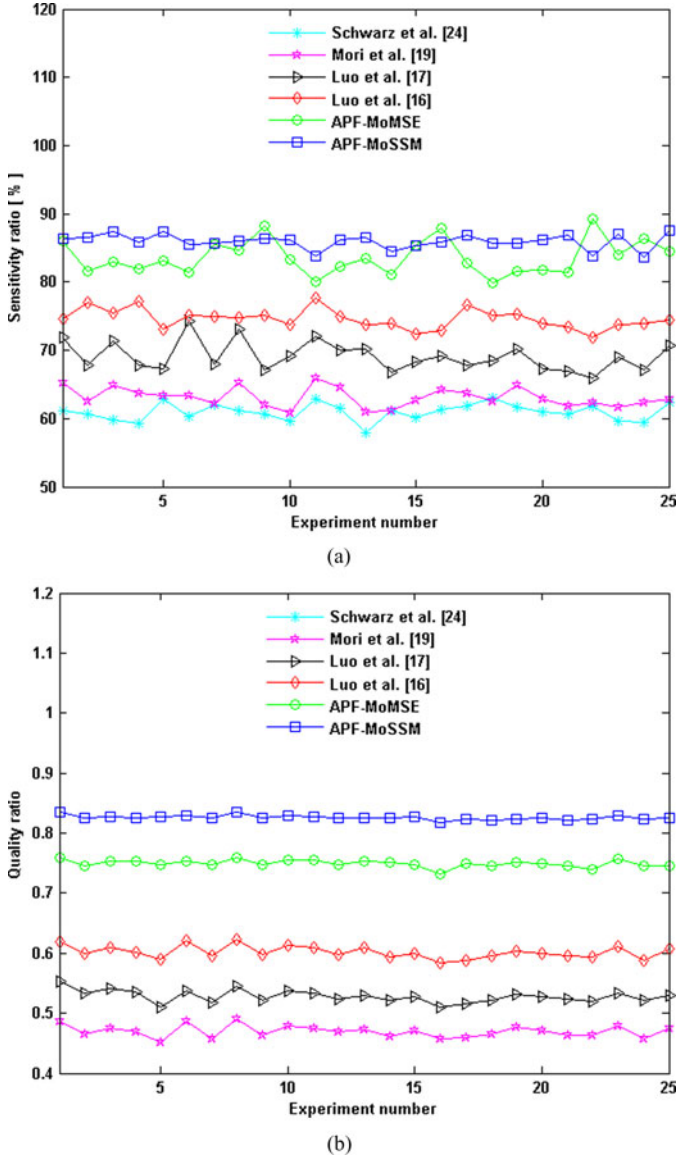
(a)

(b)

Fig. 9.   Plotted sensitivity and similarity against tracking results obtained from six methods on 25 datasets. Large sensitivity and similarity ratios indicate good visualization quality and robust tracking performance. (a) Sensitivity ratio ρ_e. (b) Similarity ratio χ_e.
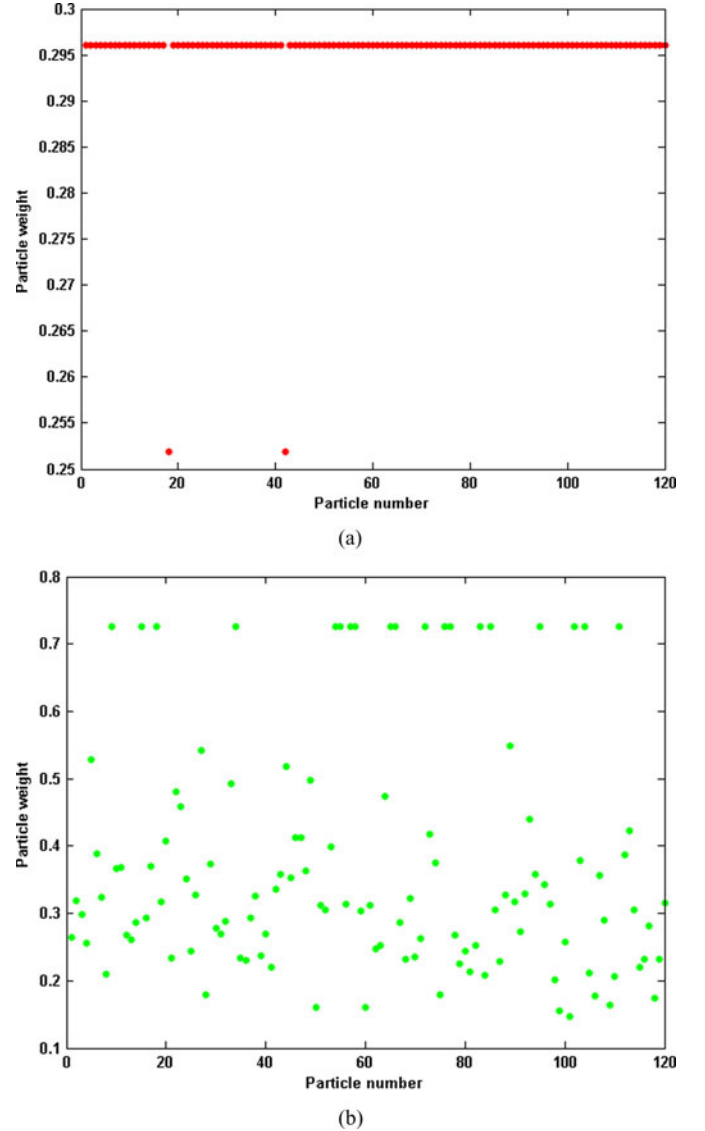

(a)

(b)

Fig. 10.   Comparison of particle diversity of standard particle filter [16] and our animated one. The more different weights that exist, the more potential modes in solution space and less particle impoverishment. (a) Standard particle filter [16]. (b) APF.

significantly better than the other methods. Fig. 11(a) and (b) compares the estimated camera moving path of the methods of Luo *et al.* [16] and APF-MoSSM. The estimated path of APF-MoSSM is much closer to the ground truth than Luo *et al.* [16], particularly the circled region where the simulated respiratory motion was quite big. Fig. 11(c) shows the visualization quality of the virtual images generated from the position and orientation parameters estimated by the six methods. Our method outperforms the other approaches since its virtual images resemble the selected video frames much better than the others. Fig. 12 illustrates the runtime of the different methods. Note that the method of Schwarz *et al.* [24] is real time processing.

VII. DISCUSSION

This paper aims to efficiently use endoscopic images and EMT sensor outputs to accurately and robustly track endoscope movements under image artifacts and EMT measurement inaccuracy and distortion. In general, our APF framework successfully achieved precise and stable tracking to tackle the limitations in ENE. We contribute our achievement to the following aspects.

First, we believe that the particle diversity of our APF was significantly enhanced, as demonstrated in Fig. 10, which shows the diversity of the 120 particle weights of the standard particle filter [16] and our animated one. The standard is trapped in particle impoverishment since its particles almost had the same weights [see Fig. 10(a)]. The particle weights or fitnesses of our animated one are obviously distributed more diversely than the standard one; i.e., the more different weights that exist, the more potential modes in the solution space and less particle impoverishment. Next, our method benefits from the adaptive evolutionary factors in the APSO-based particle animation step,

TABLE I
ACCURACY, SMOOTHNESS, AND VISUALIZATION QUALITY OF SIX METHODS EVALUATED ON TEN GROUND TRUTH DATASETS

Investigated	Accuracy		Smoothness		Visualization quality	
approaches	Position ζ (mm)	Orientation ϕ (°)	Position τ (mm)	Orientation ψ (°)	Sensitivity $\tilde{\rho}$	Similarity $\tilde{\chi}$
Schwarz et al. [24]	7.06±8.34	12.4±13.1	4.48±0.91	4.88±0.34	60.9%	0.34
Mori et al. [19]	5.85±6.04	10.5±10.0	4.65±0.80	4.37±0.39	63.1%	0.47
Luo et al. [17]	4.80±3.87	8.60±9.31	4.24±0.78	3.52±1.40	69.1%	0.53
Luo et al. [16]	4.19±4.83	7.75±8.43	4.09±0.88	3.37±1.19	74.6%	0.60
APF-MoMSE	3.52±2.80	6.05±5.48	2.05±0.74	2.60±0.45	83.6%	0.75
APF-MoSSM	2.48±1.45	4.69±3.93	1.84±0.66	2.52±0.46	85.9%	0.83

They were calculated in terms of (22)–(26). Our APF-MoMSE and -MoSSM approaches outperformed other methods.
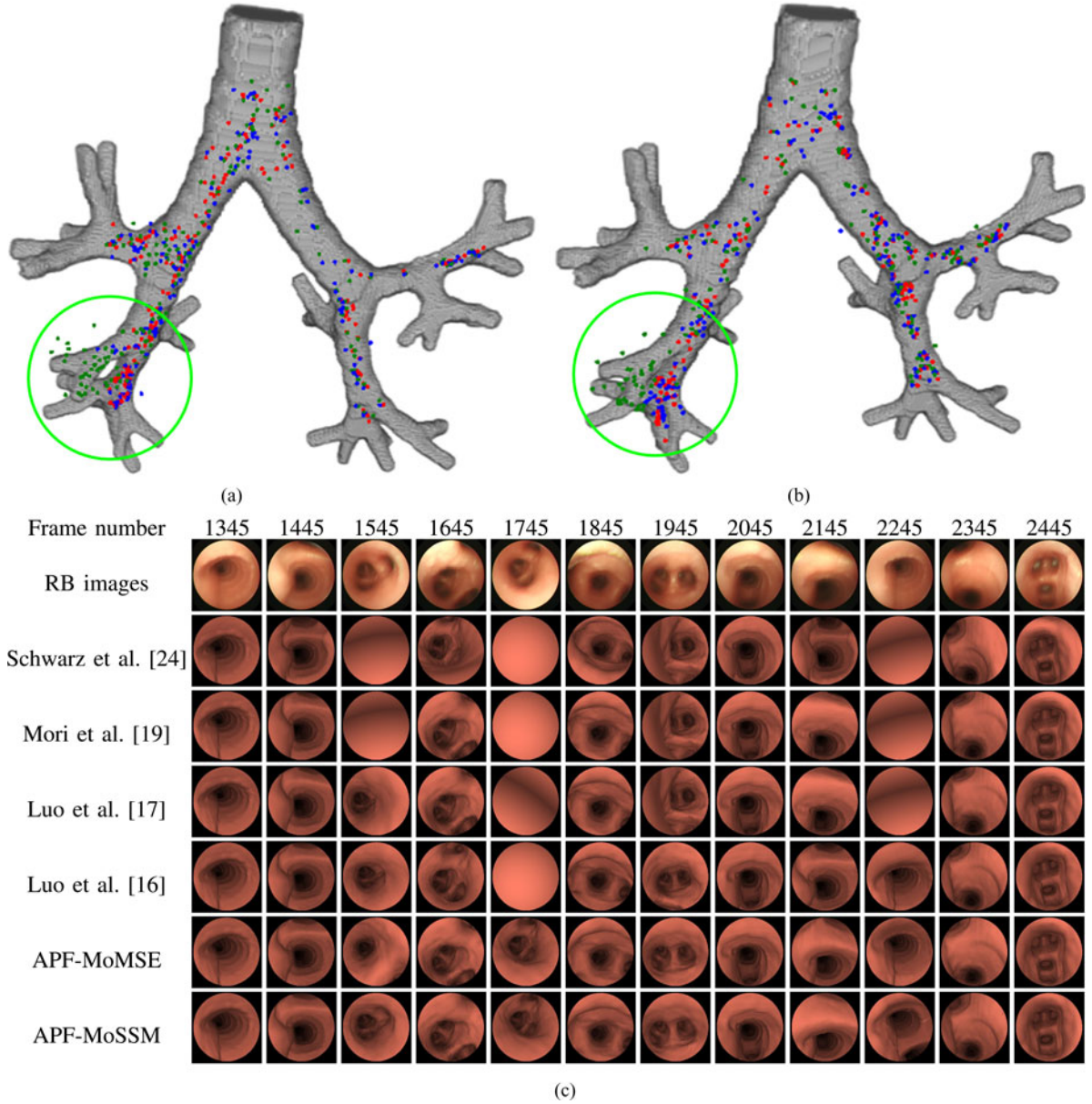

Fig. 11. Examples of visual comparison of tracking results. (a) and (b) Ground truth (*red dots*) and tracked camera paths from methods of Luo *et al.* [16] (*green dots*) and APF-MoSSM (*blue dots*). (c) Top row shows frame numbers selected uniformly every 100 frames, and second row shows corresponding real images. Other rows display virtual images generated from tracking results using methods discussed above. Our proposed APF-based frameworks display much better performance. (a) One example path of Data 03. (b) One example path of Data 05.

as proved in Fig. 6. Adaptive evolutionary parameters are very useful to renew particles with more diversity. They were automatically computed on the basis of spatial distance constraints and image intensity information to more flexibly control the global and local search abilities and determining a reasonable velocity limitation to update the particles. Finally, we are absolutely convinced that the improvement of the tracking performance can also be attributed to our MoSSM-based likelihood model. Beyond the particle diversity that is vital to any stochastic filtering methods, a likelihood model also plays an important role during state estimation, since each particle is weighed by it. The particles should be weighed as accurately as possible since the particle weight or fitness indicates the particle's searching ability. Since it is common to use an image appearance model as the likelihood model, we follow such a generic model and expect it to be insensitive or robust to image artifacts. The MoSSM model satisfies our requirement in endoscope tracking, as previously discussed [30]. MoSSM greatly outperforms MoMSE since the latter commonly assumes that corresponding pixels have a similar color value called color consistency [30], which is easily violated when computing the similarity between video and virtual images as the particle weight. However, it remains somewhat challenging to evaluate whether a particle is with an important weight or a potential optimal solution for motion states. A more robust likelihood model might be developed to further tackle image artifacts and improve our tracking performance.

We need to clarify the potential limitations of our method. Two main limitations come from data acquisition: 1) the CT data utilized here were collected without breathing motion information and 2) the image artifacts or uninformative endoscopic video sequences. The former complicates compensating for breathing motion during endoscopy. The latter remains challenging to get the correct computation of the particle weight. Another open issue is the computational time of our method, which needs about 0.81 s to process one frame. The particle evaluation was time-consuming, including the similarity computation and virtual image generation by volume rendering. Graphics processing unit or multithreading techniques can be used to accelerate the processing. Additionally, we currently did not evaluate our method on clinical datasets. Undoubtedly, we must validate it on patient data. However, obtaining them is really difficult since few clinical datasets are available worldwide, even though a commercially available EMT-based bronchoscopy system called superDimension is increasingly being introduced in operating rooms. Unfortunately, we failed to obtain any clinical datasets. We will obtain some for future clinical validation.

Finally, our proposed APF method for endoscope tracking and navigation is significantly related to visual simultaneous localization and mapping (SLAM) [21], [22]. Visual SLAM algorithms enable robots or mobile objects to generate a map within a dynamic scenario or update a map within a prebuilt scenario using a monocular camera. They also simultaneously allow robots to determine their 6 DoF motion parameters or moving trajectories. They usually use image feature points as well Kalman or particle filtering for that purpose. In our case of endoscope navigation or its 3-D motion estimation, we used
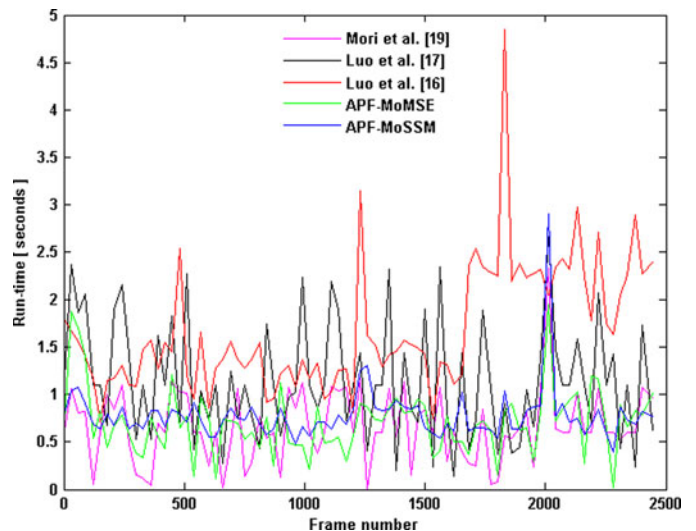

Fig. 12.   Average run times of Mori *et al.* [19], Luo *et al.* [17], Luo *et al.* [16], APF-MoMSE, and APF-MoSSM were 0.61, 1.17, 1.72, 0.73, and 0.81 s/frame, respectively.

our APF method to accurately and robustly locate the endoscope tip inside the preoperative images (e.g., CT or MR) that are segmented for a map. Therefore, at the point of the camera 3-D location or the trajectory recovery within a map, our method closely resembles visual SLAM algorithms. However, our APF method did not perform any image feature point detection and matching procedures. We only extracted the structural regions in the video images by simple image processing (see Section III-E). We used these regions to compute the similarity between the video and virtual images for evaluating particle weights. Since APF looks very similar to visual SLAM, we believe that our work also contributes to the SLAM community.

VIII. CONCLUSION

This paper proposed an accurate and stable endoscope 3-D motion tracking approach on the basis of an APF that uses APSO to deal with the particle impoverishment of standard particle filter algorithms. Our experimental results demonstrate that the particle impoverishment problem was significantly diminished by APSO. Furthermore, our method successfully tackled the limitations of endoscopic image artifacts and inaccurate and distorted EMT sensor measurements caused by respiratory motion and EMT system jitter errors. Its current tracking accuracy and smoothness were 2.48 mm, 4.69° and 1.84 mm, 2.52°. Both were significantly improved compared to previous methods. Future work includes improvement of our method's performance, reduction of its computational time, and patient data validation.

REFERENCES

[1] M. S. Arulampalam, S. Maskell, and N. Gordon, "A tutorial on particle filters for nonlinear/non-Gaussian Bayesian tracking," *IEEE Trans. Signal Process.*, vol. 50, no. 2, pp. 174–188, Feb. 2002.

[2] L. Y. Ching, K. Moller, and J. Suthakorn, "Non-radiological colonoscope tracking image guided colonoscopy using commercially available electromagnetic tracking system," in *Proc. IEEE Conf. Robot. Autom. Mechatron.*, 2010, pp. 62–67.

[3] S. Das, A. Abraham, and A. Konar, "Particle swarm optimization and differential evolution algorithms: Technical analysis, applications and hybridization perspectives," *Stud. Comput. Intell.*, vol. 116, pp. 1–38, 2008.

[4] D. Deguchi, K. Mori, M. Feuerstein, T. Kitasaka, C. R. Maurer, Jr, Y. Suenaga, H. Takabatake, M. Mori, and H. Natori, "Selective image similarity measure for bronchoscope tracking based on image registration," *Med. Image Anal.*, vol. 13, no. 4, pp. 621–633, 2009.

[5] F. Deligianni, A. J. Chung, and G. Z. Yang, "Nonrigid 2-D/3-D registration for patient specific bronchoscopy simulation with statistical shape modeling: Phantom validation," *IEEE Trans. Med. Imag.*, vol. 25, no. 11, pp. 1462–1471, Nov. 2006.

[6] M. Feuerstein, T. Reichl, J. Vogel, J. Traub, and N. Navab, "Magneto-optical tracking of flexible laparoscopic ultrasound: Model-based online detection and correction of magnetic tracking errors," *IEEE Trans. Med. Imag.*, vol. 28, no. 6, pp. 951–967, Jun. 2009.

[7] I. Gergel, T. R. dos Santos, R. Tetzlaff, L. Maier-Hein, H.-P. Meinzer, and I. Wegner, "Particle filtering for respiratory motion compensation during navigated bronchoscopy," in *Proc. SPIE Med. Imag.*, 2010, vol. 7625, p. 76250W.

[8] C. Hayhurst, P. Byrne, P. R. Eldridge, and C. L. Mallucci, "Application of electromagnetic technology to neuronavigation: A revolution in image-guided neurosurgery," *J. Neurosurg.*, vol. 111, no. 6, pp. 1179–1184, 2009.

[9] J. P. Helferty, A. J. Sherbondy, A. P. Kiraly, and W. E. Higgins, "Computer-based system for the virtual-endoscopic guidance of bronchoscopy," *Comput. Vis. Image Understand.*, vol. 108, no. 1–2, pp. 171–187, 2007.

[10] L. Hong, A. Kaufman, Y. C. Wei, A. Viswambharan, M. Wax, and Z. Liang, "3D virtual colonoscopy," in *Proc. Biomed. Vis.*, 1995, pp. 26–32.

[11] M. Isard and A. Blake, "Condensation—Conditional density propagation for visual tracking," *Int. J. Comput. Vis.*, vol. 29, no. 1, pp. 5–28, 1998.

[12] J. Kennedy and R. C. Eberhart, "Particle swarm optimization," in *Proc. IEEE Int. Conf. Neural Netw.*, 1995, pp. 1942–1948.

[13] V. V. Kindratenko, "A survey of electromagnetic position tracker calibration techniques," *Virt. Reality: Res., Dev., Appl.*, vol. 5, no. 3, pp. 169–182, 2000.

[14] E. B. Levy, J. Tang, D. Lindisch, N. Glossop, F. Banovac, and K. Cleary, "Implementation of an electromagnetic tracking system for accurate intrahepatic puncture needle guidance: Accuracy results in an in vitro model," *Acad. Radiol.*, vol. 14, pp. 344–354, 2007.

[15] X. Luo, M. Feuerstein, D. Deguchi, T. Kitasaka, H. Takabatake, and K. Mori, "Development and comparison of new hybrid motion tracking for bronchoscopic navigation," *Med. Image Anal.*, vol. 16, no. 3, pp. 577–596, 2012.

[16] X. Luo, M. Feuerstein, T. Kitasaka, and K. Mori, "Robust bronchoscope motion tracking using sequential Monte Carlo methods in navigated bronchoscopy: Dynamic phantom and patient validation," *Int. J. Comput. Assist. Radiol. Surg.*, vol. 7, no. 3, pp. 371–387, 2012.

[17] X. Luo, M. Feuerstein, T. Sugiura, T. Kitasaka, K. Imaizumi, Y. Hasegawa, and K. Mori, "Towards hybrid bronchoscope tracking under respiratory motion: Evaluation on a dynamic motion phantom," in *Proc. SPIE*, 2010, vol. 7625, p. 76251B.

[18] X. Luo, T. Kitasaka, and K. Mori, "Endoscope 3-d motion tracking using an aggressive particle filtering for boosting electromagnetic guidance endoscopy," in *Proc. 21st Int. Conf. Pattern Recog.*, 2012, pp. 117–120.

[19] K. Mori, D. Deguchi, K. Akiyama, T. Kitasaka, C. R. Maurer, Jr, Y. Suenaga, H. Takabatake, M. Mori, and H. Natori, "Hybrid bronchoscope tracking using a magnetic tracking sensor and image registration," in *Proc. Mid. Image Comput. Comput.-Assisted Intervention Conf.*, 2005, vol. 3750, pp. 543–550.

[20] K. Mori, Y. Suenaga, and J. Toriwaki, "Fast software-based volume rendering using multimedia instructions on PC platforms and its application to virtual endoscopy," in *Proc. SPIE*, 2003, vol. 5031, pp. 111–122.

[21] G. K. D. Murray, "Parallel tracking and mapping for small AR workspaces," in *Proc. IEEE 6th IEEE ACM Int. Symp. Mixed Augmented Reality*, 2007, pp. 225–234.

[22] R. A. Newcombe, S. J. Lovegrove, and A. J. Davison, "DTAM: Dense tracking and mapping in real-time," in *Proc. IEEE Int. Conf. Comput. Vis*, 2001, pp. 2320–2327.

[23] D. Parrott and X. Li, "Locating and tracking multiple dynamic optima by a particle swarm model using speciation," *IEEE Trans. Evol. Comput.*, vol. 10, no. 4, pp. 440–458, Aug. 2006.

[24] Y. Schwarz, J. Greif, H. D. Becker, A. Ernst, and A. Mehta, "Real-time electromagnetic navigation bronchoscopy to peripheral lung lesions using overlaid CT images: The first human study," *Chest*, vol. 129, no. 4, pp. 988–994, 2006.

[25] L. M. Seijo, J. P. de Torres, M. D. Lozano, G. Bastarrika, A. B. Alcaide, M. M. Lacunza, and J. J. Zulueta, "Diagnostic yield of electromagnetic navigation bronchoscopy is highly dependent on the presence of a bronchus sign on CT imaging: Results from a prospective study," *Chest*, vol. 138, no. 6, pp. 1316–1321, 2010.

[26] L. Serra, T. Poston, W. L. Nowinski, B. C. Chua, H. Ng, and P. K. Pillay, "The brain bench planner and trainer for minimal access surgery," in *Proc. ACM Symp. Virt. Reality Softw. Technol.*, 1996, pp. 191–192.

[27] M. Solaiyappan, T. Poston, P. A. Heng, E. R. McVeigh, M. A. Guttman, and E. A. Zerhouni, "Interactive visualization for rapid noninvasive cardiac assessment," *IEEE Comput.*, vol. 29, no. 1, pp. 55–62, Jan. 1996.

[28] T. D. Soper, D. R. Haynor, R. W. Glenny, and E. J. Seibel, "In vivo validation of a hybrid tracking system for navigation of an ultrathin bronchoscope within peripheral airways," *IEEE Trans. Biomed. Eng.*, vol. 57, no. 3, pp. 736–745, Mar. 2010.

[29] Z. Wang and A. C. Bovik, "A universal image quality index," *IEEE Signal Process. Lett.*, vol. 9, no. 3, pp. 81–84, Mar. 2002.

[30] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: From error visibility to structural similarity," *IEEE Trans. Image Process.*, vol. 13, no. 4, pp. 600–612, Apr. 2004.

[31] Z. Yaniv, P. Cheng, E. Wilson, T. Popa, D. Lindisch, E. Campos-Nanez, H. Abeledo, V. Watson, K. Cleary, and F. Banovac, "Needle-based interventions with the image-guided surgery toolkit (IGSTK): From phantoms to clinical trials," *IEEE Trans. Biomed. Eng.*, vol. 57, no. 4, pp. 922–933, Apr. 2010.

[32] L. Zhang, T. Mei, Y. Liu, and D.Tao, H. Q. Zhou, "Visual search reranking via adaptive particle swarm optimization," *Pattern Recog.*, vol. 44, no. 8, pp. 1811–1820, 2011.
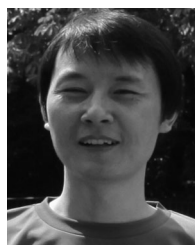
Xiongbiao Luo (M'12) received the B.E. degree in mechanical engineering and automation, the M.E. degree in materials processing engineering, and the Ph.D. degree in information science, in 2001, 2005, and 2011, respectively.

He is currently a Research Assistant Professor at the Information and Communications Headquarters, Nagoya University, Nagoya, Japan. His current research interests include video image computing, camera 3-D motion tracking, surgical instrument tracking and navigation, intraoperative patient-to-model registration, stochastic filtering, machine learning, and evolutionary computation for medical applications, endoscopy location and navigation systems, image-guided diagnosis and surgery, or information processing in computer-assisted interventions.

Kensaku Mori (M'04) received the B.S degree in electronics engineering, the M.S and Ph.D. degrees in information engineering from Nagoya University, Japan, in 1992, 1994 and 1996, respectively. He was a Research Fellow of the Japanese Society for the Promotion of Science (JSPS) from 1994 to 1997, Research Associate at the Department of Computational Science and Engineering, Nagoya University from 1997 to 2000, an Assistant Professor in 2000. He was Associate Professor at the Research Center for Advance Waste and Emission Management of Nagoya University from 2001 to 2003. He was also Visiting Associate Professor at the Department of Neurosurgery, Stanford University, USA, from 2001 to 2002. He was Associate Professor of the Graduate School of Information Science, Nagoya University from 2003 to 2009. He has been Professor of Information and Communications Headquarters, Nagoya University since 2009. His current research interests include three dimensional image processing, computer graphics, virtual reality and their applications to medical image.

Dr. Mori has received many awards from several institutes includes the Young Researcher Award, Ministry of Education, Culture, Sports, the best paper awards from the Japanese Society of Medical Imaging Technology. He is appointed as the General Chair of the 16th International Conference of Medical Image Computing and Computer Assisted Intervention (MICCAI 2013). He is a member of the Institute of Electronics, Information and Communication Engineering of Japan, Japanese Society of Biomedical Engineering, and Japanese Society of Medical Imaging Technology, MICCAI, and SPIE.