Robust electromagnetically guided endoscopic procedure using enhanced particle swarm optimization for multimodal information fusion

Xiongbiao Luo,[a)] Ying Wan[b)], and Xiangjian He[c)]

[a]*Robarts Research Institute, Western University, London, Ontario, N6A 5K8, Canada*

[b,c]*School of Computing and Communications, University of Technology, Sydney, New South Wales, 2007, Australia*

(Dated: 1 October 2014)

b) has equal contribution with corresponding author

Purpose: Electromagnetically guided endoscopic procedure, which aims at accurately and robustly localizing the endoscope, involves multimodal sensory information during interventions. However, it still remains challenging in how to integrate these information for precise and stable endoscopic guidance. To tackle such a challenge, this paper proposes a new framework on the basis of an enhanced particle swarm optimization method to effectively fuse these information for accurate and continuous endoscope localization.

Methods: We use the particle swarm optimization method, one of stochastic evolutionary computation algorithms, to effectively fuse the multimodal information including pre-operative information (i.e., computed tomography images) as a frame of reference, endoscopic camera videos, and positional sensor measurements (i.e., electromagnetic sensor outputs). Since the evolutionary computation method usually limits its possible premature convergence and evolutionary factors, we introduce the current (endoscopic camera and electromagnetic sensor's) observation to boost the particle swarm optimization and also adaptively update evolutionary parameters in accordance with spatial constraints and the current observation, resulting in advantageous performance in the enhanced algorithm.

Results: The experimental results demonstrate that our proposed method provides a more accurate and robust endoscopic guidance framework than state-of-the-art methods. The average guidance accuracy or error of our framework was about 3.0 mm and 5.6° while the previous methods show at least 3.9 mm and 7.0°. The average position and orientation smoothness of our method was 1.0 mm and 1.6°, which is significantly better than the other methods at least with (2.0 mm and 2.6°). Additionally, the average visual quality of the endoscopic guidance was improved to 0.29.

Conclusions: A robust electromagnetically guided endoscopy framework was proposed on the basis of an enhanced particle swarm optimization method with using the current observation information and adaptive evolutionary factors. Our proposed framework greatly reduced the guidance errors from (4.3, 7.8) to (3.0 mm, 5.6°), compared to state-of-the-art methods.

PACS numbers: ——

Electromagnetically guided endoscopic procedure

Keywords: Endoscopic navigation, particle swarm optimization, electromagnetic tracking, multimodal information fusion, computer-assisted interventions

I. INTRODUCTION

Endoscopic interventions are widely performed for diagnosis and treatment of different cancers, e.g., bronchus and lung cancers via bronchoscopic surgery[1–4], bowel cancer by colonoscopic surgery[5], brain tumors using neurosurgery[6,7]. Guided endoscopic procedure is endoscopy system is generally recognized as the next generation of interventional endoscopy. Such a guidance usually involves different modalities of sensory information: (1) pre-operative information, e.g., 3-D computed tomography (CT) images, (2) two-dimensional (2-D) endoscopic video sequences, and (3) positional sensor measurements, e.g., a full six degrees of freedom (6DoF) sensor output from electromagnetic (EM) tracker systems (*Northern Digital Inc.*[8], Waterloo, Canada). However, solely using one of three kinds of information is very difficult to accurately navigate or localize the endoscope to cancerous regions where surgical interventions must be performed. At this point, the endoscopic guidance is considered as a procedure of multimodal information fusion.

Unfortunately, it still remains challenging to effectively and efficiently combine three modalities of sensory information for the endoscopic guidance since these information are *incomplete* that means their limitations: (1) pre-operative information (e.g., CT slices) before endoscopic surgery cannot record any information about patient movements (e.g., respiration) which exactly happen in the operating room, (2) image artifacts (e.g., inter-reflection or motion blurring) unavoidably occur in endoscopic videos, resulting in a large number of low-quality images, and (3) the positional EM sensor's accuracy is heavily deteriorated by patient movements and the magnetic field distortion that is caused by ferrous metals or conductive material within or close to the working volume of EM trackers. Due to the incompleteness of these information, current commercially available tracking systems (e.g., superDimension), which are increasingly used in clinical applications[9–11], were reported their diagnostic accuracy between 59% and 74%, without sensitivity to tumor size[12,13].

To tackle these *incomplete* information for accurate endoscopic guidance, it is common to formulate such multimodal information fusion as an optimization process, which is usually solved by deterministic[14,15] or stochastic[16–18] approaches. Deterministic methods, typically 2-D/3-D image registration algorithms[4], usually define an optimization function to minimize the pixel difference between endoscopic video images and virtual renderings generated from pre-operative information. On the other hand, stochastic approaches were demonstrated to

be more precise and stable than deterministic methods since they take dynamic uncertainties into account and adapt themselves to *incomplete* information for the endoscope localization.

Even though several publications have been discussed about stochastic methods[17–19], a more robust and accurate stochastic approach is still vastly expected to effectively fuse these *incomplete* information for guided endoscopic procedures. Recently, a numerous population-based stochastic evolutionary algorithm, particle swarm optimization (PSO), which was originally proposed by Kennedy and Eberhart[20], has been increasingly applied as a successful optimization technique to address complex problems[21–23]. The algorithm simulates natural and biological behaviors such as birds flocking and fish schooling to find optimal solutions in nonlinear and high-dimensional spaces. Moreover, one of most attractive aspects of PSO is able to deal with nonlinear, non-differentiable, and multimodal optimization problems by dynamically interacting all particles in a similar analogy with the *cognitive* and *social* properties of populations[24]. However, the standard PSO method limits to its evolutionary parameters, particularly it does not take the current observation information into consideration when propagating the particle swarm to the new state. These limitations potentially result in a premature convergence.

This paper proposes an accurate and robust electromagnetically endoscopic guidance framework that uses an enhanced particle swarm optimization (EPSO) algorithm to address these limitations to improve the PSO's performance. It is worthwhile to highlight several technical contributions of this work. First, EPSO integrates the current observation information into the particle swarm. The current observation results in the more powerful searching ability for particles to approximate the optimal solution. We also introduce a new strategy to adaptively control evolutionary parameters on the basis of the particle fitness value and spatial constraints, which are very useful to tackle the diversity loss problem, alleviate particle impoverishment, and obtain various particle diversity in the EPSO's iterations. Next, with application to the *incomplete* multimodal information fusion, an EPSO-based endoscopic guidance method was developed for surgical procedures. We successfully formulated the problem of the endoscopic guidance as an EPSO-based stochastic optimization process. The EPSO algorithm exactly provides an effective scheme to fuse multimodal information of the CT images, endoscopic videos, and EM sensor measurements for achieving a robust and accurate guidance procedure. Additionally, although our EPSO algorithm currently was applied to the medical image community, it should be definitely applicable to other dif-

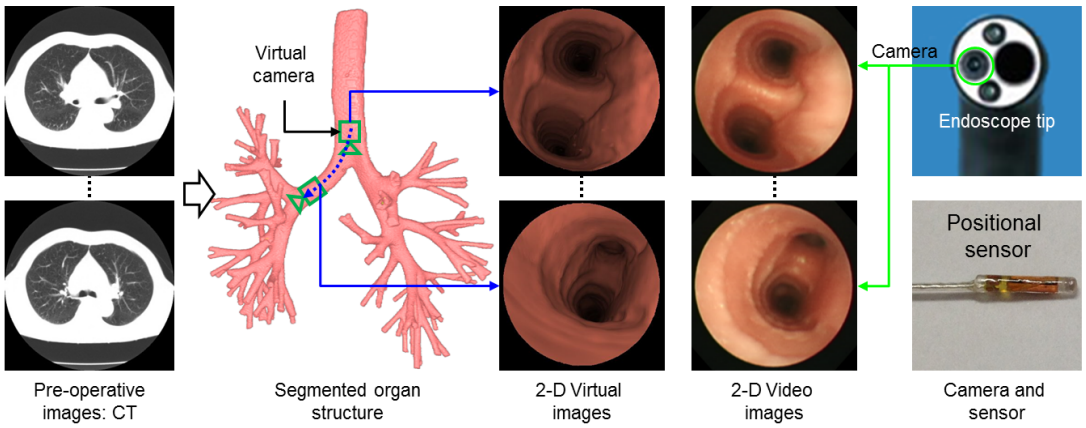


FIG. 1 Multimodal information including the CT images used to generate 2-D virtual rendering images, endoscopic video images, and positional sensor measurements involved in an electromagnetically guided endoscopic procedure.

ferent optimization problems in the computer vision field, e.g., image segmentation, visual tracking, or object recognition.

II. METHODS AND MATERIALS

Electromagnetically guided endoscopy seeks to accurately and continuously determine the endoscope location or its 6DoF motion parameters including position and orientation in a reference frame (e.g., the CT image coordinate system). The core of our endoscopic guidance framework is to use the EPSO algorithm to integrate multimodal information to precisely navigate the endoscope. Before discussing EPSO, we slightly introduce different sensory information involved in the endoscopic guidance.

II.A. Multimodal information

II.A.1. Pre-operative information

Before performing endoscopic examinations, physicians usually use 3-D imaging devices such as CT scanners to acquire pre-operative images. We segment these pre-operative images to obtain organ structures and pre-build 3-D anatomy model before endoscopic procedures[4]. We use the CT images to generate 2-D virtual images by volume rendering techniques[25]

Electromagnetically guided endoscopic procedure

(Fig. 1).

II.A.2. Endoscopic video sequences

The endoscopic camera is directly inserted to observe organ interior structures. Endoscopic video images display sufficient organ interior surface information (Fig. 1). However, the camera only provides 2-D live video images without depth information to target regions. Hence, we require to integrate 2-D endoscopic images with 3-D CT images to guide endoscopic procedures.

II.A.3. Positional sensor measurements

Ultra-miniature sensors, i.e, EM sensors (Fig. 1), which are externally attached to the endoscope's distal tip surface or its working channel, are increasingly used to measure the endoscopic camera movement information. However, the sensor measurements might be inaccurate due to tissue deformation, e.g., respiration.

II.B. Enhanced particle swarm optimization

The PSO algorithm is a population-based stochastic optimization technique. It seeks to propagate the population with a number of particles to approximate the optimal solution. Suppose that we generate a population of particles $\mathcal{P}_j = \{(\mathbf{p}_{i,j}, f(\mathbf{p}_{i,j}), \gamma_{i,j})\}_{i=1}^{M}$ at j-th iteration, where $f(\mathbf{p}_{i,j})$ is the particle fitness value, $\gamma_{i,j}$ is a weight on the basis of spatial distribution constraint, M is the particle number, and $j = 1, 2, \ldots N$, N is the iteration number. Our proposed EPSO algorithm is to update these particles with several steps: (1) randomization, (2) propagation, (3) evolutionary parameters analysis, and (4) population update, as discussed as follows.

II.B.1. Population randomization

We first initialize particle $\mathbf{p}_{0,0}$ and $f(\mathbf{p}_{0,0}) = \gamma_{0,0} = 1/M$ before iterations. We perform a randomization only once to increase the diversity of particles and avoid the particle impoverishment. Particle $\mathbf{p}_{0,0}$ is randomized by using the Gaussian propagation model, and we

Electromagnetically guided endoscopic procedure

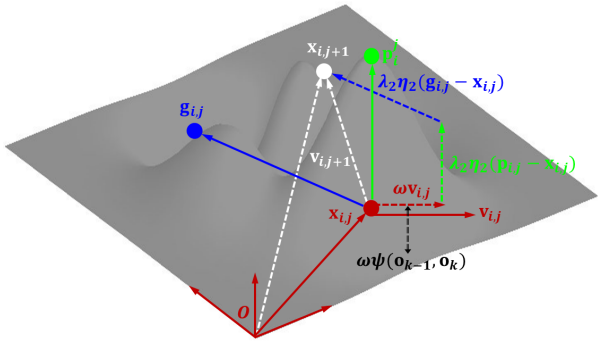


FIG. 2 Population propagation to obtain new particle states.

obtain particle state $\mathbf{x}_{0,0}$:

$$\mathbf{x}_{0,0} = \mathcal{G}(\mathbf{p}_{0,0}, \mu\Lambda), \tag{1}$$

where $\mathcal{G}$ is the transform function that is used to add changeable part $\mu\Lambda$ to particle $\mathbf{p}_{0,0}$ and obtain new particle $\mathbf{x}_{0,0}$, μ is a Gaussian distribution random number: $\mu \sim \mathcal{N}(0,1)$, and Λ is a pre-determined constant vector. Note that the randomization for particle diversification does not perform a *resampling process*, as particle filter methods do[18], since the local best particles provide compact samples for propagation[26].

II.B.2. Population propagation

After the randomization, we obtain particles $\{\mathbf{x}_{i,j}\}_{i=1}^{M}$ at j-th iteration and propagate $\mathbf{x}_{i,j}$ to $\mathbf{x}_{i,j+1}$ at iteration j+1 on the basis of velocity $\mathbf{v}_{i,j}$ and inertia weight ω that is used for deciding how much $\mathbf{v}_{i,j}$ to be preserved in $\mathbf{v}_{i,j+1}$ (Fig. 2):

$$\mathbf{v}_{i,j+1} = \omega\mathbf{v}_{i,j} + \underbrace{\lambda_1\eta_1(\mathbf{p}_{i,j} - \mathbf{x}_{i,j})}_{\Delta_p} + \underbrace{\lambda_2\eta_2(\mathbf{g}_{i,j} - \mathbf{x}_{i,j})}_{\Delta_g}, \tag{2}$$

$$\mathbf{x}_{i,j+1} = \mathbf{x}_{i,j} + \mathbf{v}_{i,j+1}, \tag{3}$$

where λ_1 and λ_2 are acceleration constants and η_1 and η_2 are randomly generated from the uniform distribution with interval [0.0 1.0]. Particles $\mathbf{p}_{i,j}$ (for the local individual best) and $\mathbf{g}_{i,j}$ (for the global all best) are the best state found by particle i so far and the best state found by the whole swarm so far.

The PSO's performance depends on this particle progapation (Eq. 3), where velocity $\mathbf{v}_{i,j}$ plays an important role since it aims at keeping all the particles approaching to the optimal

solution. It is difficult to accurately determine $\mathbf{v}_{i,j}$. In the standard PSO algorithm, it is empirically determined within a reasonable range, which usually loses the relative continuity between two consecutive times $k-1$ and k from iterations j to $j+1$. PSO might get trapped in a local minima convergence. To address this problem, our idea is to integrate current observation at time k to compute velocity $\mathbf{v}_{i,j}$:

$$\mathbf{v}_{i,j} = \Phi(\mathbf{o}_{k-1}, \mathbf{o}_k), \tag{4}$$

where function $\Phi(\cdot)$ is to compute the relative continuity between previous and current observations $\mathbf{o}_{k-1}$ and $\mathbf{o}_k$.

Now velocity $\mathbf{v}_{i,j+1}$ can be deterministically updated by

$$\mathbf{v}_{i,j+1} = \omega \underbrace{\Phi(\mathbf{o}_{k-1}, \mathbf{o}_k)}_{Observation} + \Delta_p + \Delta_g, \tag{5}$$

which is effective to control the population's propagation.

II.B.3. Evolutionary parameters analysis

Evolutionary parameters λ_1, λ_2, and ω heavily influence the optimization performance. Most of current improved PSO algorithms do not consider the relative or temporal continuity and spatial constraint information. This might result in a lack of systematic treatment of evolutionary states and expose PSO to a dangerous level of swarm explosion and divergence. To handle this limitation, we here adaptively control acceleration factors λ_1 and λ_2 relative to the particle's fitness value:

$$\lambda_1 = \frac{2f(\mathbf{p}_{i,j})}{f(\mathbf{p}_{i,j}) + f(\mathbf{g}_{i,j})}, \quad \lambda_2 = \frac{2f(\mathbf{g}_{i,j})}{f(\mathbf{p}_{i,j}) + f(\mathbf{g}_{i,j})}, \tag{6}$$

where fitness $f(\cdot)$ is defined as observation probability $\pi(\cdot)$ of each particle relative to current observation $\mathbf{o}_k$:

$$f(\mathbf{p}_{i,j}) = \pi(\mathbf{o}_k|\mathbf{p}_{i,j}), \quad f(\mathbf{g}_{i,j}) = \pi(\mathbf{o}_k|\mathbf{g}_{i,j}). \tag{7}$$

For adaptively calculating ω, we utilize both fitness value $f(\mathbf{x}_{i,j}) \in [0.0\ 1.0]$ and particle spatial distribution information $\gamma_{i,j}$ among all the particles in the population. We first compute average distance $d_{i,j}$ from one particle to all other particles:

$$d_{i,j} = \frac{1}{M-1} \sum_{i=1, i \neq i'}^{M} \left\| \mathbf{x}_{i,j} - \mathbf{x}_{i',j} \right\|. \tag{8}$$

where symbol $||\cdot||$ means the Euclidean norm. After finding the largest distance d_{max} and the smallest distance d_{min} from $\{d_{i,j}\}_{i=1}^{M}$, we normalize distance $d_{i,g}$ between one particle and the global best particle and obtain $\gamma_{i,j}$ for each particle:

$$\gamma_{i,j} = (d_{i,g} - d_{min})/(d_{max} - d_{min}), \quad \gamma_{i,j} \in [0.0\ 1.0]. \tag{9}$$

Finally, since inertia weight ω was suggested within the interval [0.4 0.9] for weighting the global and the local searching abilities[24], we can adaptively calculate it by

$$\omega(f(\mathbf{x}_{i,j}), \gamma_{i,j}) = \frac{2}{2 + 3\exp(-1.28(f(\mathbf{x}_{i,j}) + \gamma_{i,j}))}, \tag{10}$$

which shows a novel strategy to automatically control ω in our proposed EPSO algorithm. Note that spatial constraint $\gamma_{i,j}$ uses the particle distribution information to limit the particle to be moved far from the optimal solution and enhances the exploitation ability of each particle in the population.

II.B.4. Population update

After the $(j+1)$-th iteration, we obtain population $\mathcal{P}_{j+1}$ where $\mathbf{p}_{i,j+1}$ and $\mathbf{g}_{i,j+1}$ are updated using fitness $f(\mathbf{x}_{i,j+1})$:

$$\mathbf{p}_{i,j+1} = \begin{cases} \mathbf{x}_{i,j+1} & \text{if} f(\mathbf{x}_{i,j+1}) > f(\mathbf{p}_{i,j}) \\ \mathbf{p}_{i,j} & \text{otherwise} \end{cases}, \tag{11}$$

$$\mathbf{g}_{i,j+1} = \arg \max_{\mathbf{p}_{i,j+1} \in \mathcal{P}_{j+1}} f(\mathbf{p}_{i,j+1}). \tag{12}$$

Based on Eqs. 1∼12, all the particles are updated iteratively to approximate the optimal solution during our EPSO procedure.

II.C. Application to endoscopic guidance

This section applies the EPSO algorithm to fuse the multimodal information for guiding the endoscopic procedure. Fig. 3 shows the flowchart of our EPSO method that integrates the multimodal information for the robust electromagnetically endoscopic guidance during interventions. We first interpolate EM sensor measurements.

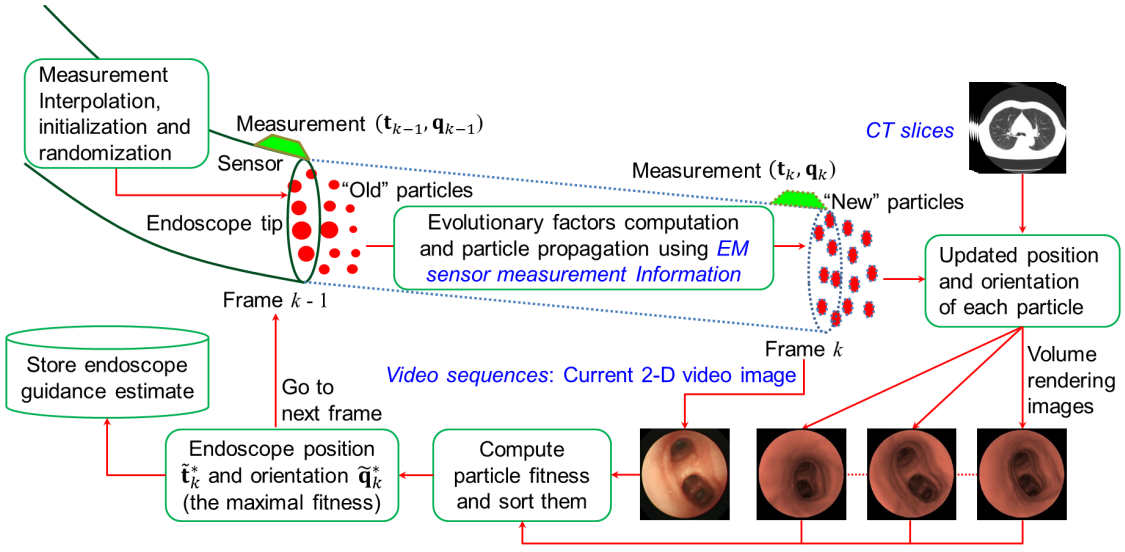


FIG. 3 Our endoscopic guidance framework using EPSO for the multimodal information integration.

II.C.1. Sensor measurement interpolation

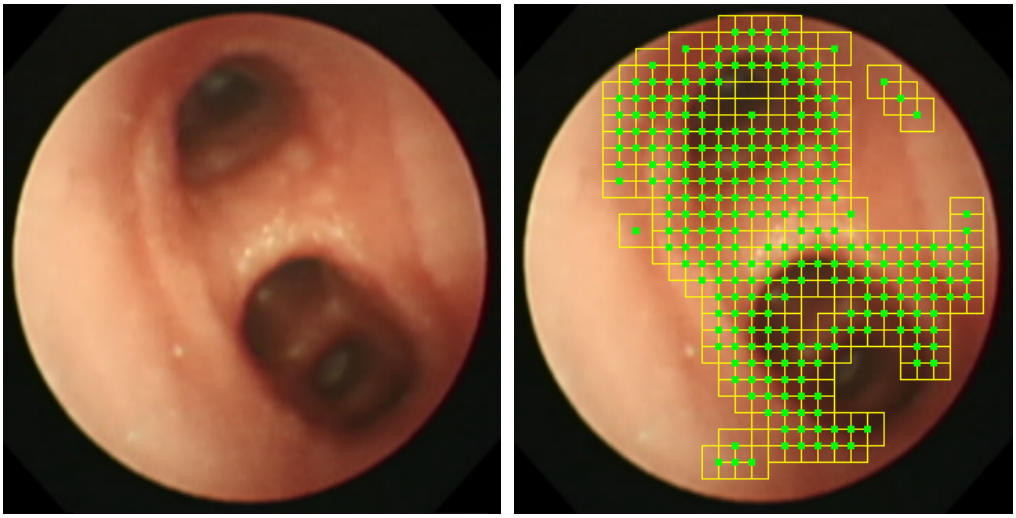
Since the endoscope movement is spatially continuous, we perform the Catmull-Rom spline interpolation for endoscope position $\mathbf{t}_k$ at time k (or at frame k) and the spherical linear interpolation (SLERP) for endoscope orientation that is described by quaternion $\mathbf{q}_k$ to smoothly approximate endoscope real movement[27]. Position $\mathbf{t}_k$ is interpolated by:

$$\mathbf{t}_k = \Omega \begin{pmatrix} 0 & 1 & 0 & 0 \\ -\beta & 0 & \beta & 0 \\ 2\beta & \beta-3 & 3-2\beta & -\beta \\ -\beta & 2-\beta & \beta-2 & \beta \end{pmatrix} \begin{pmatrix} \mathbf{t}_{v-1} \\ \mathbf{t}_v \\ \mathbf{t}_{v+1} \\ \mathbf{t}_{v+2} \end{pmatrix}, \tag{13}$$

where $\Omega = (1\ \alpha\ \alpha^2\ \alpha^3)$, the interpolation ratio $\alpha = k/u - \lfloor k/u \rfloor$, $\mathbf{t}_{v-1} \cdots \mathbf{t}_{v+2}$ are positions of continuous controlled points ($v = \lfloor k/u \rfloor$, floor operator $\lfloor \cdot \rfloor$, time spacing u), and tension parameter β impacts on the curvature at the control points (usually, $\beta = 0.5$). Quaternion $\mathbf{q}_k$ is interpolated by:

$$\mathbf{q}_k = \begin{cases} \frac{\sin(1-\alpha)\phi}{\sin\phi}\mathbf{q}_v + \frac{\sin\alpha\phi}{\sin\phi}\mathbf{q}_{v+1} & \phi \geq 0 \\ \frac{\sin(1-\alpha)\phi}{\sin\phi}\mathbf{q}_v - \frac{\sin\alpha\phi}{\sin\phi}\mathbf{q}_{v+1} & \phi < 0 \end{cases}, \tag{14}$$

$$\phi = \arccos \frac{< \mathbf{q}_v,\ \mathbf{q}_{v+1} >}{\| \mathbf{q}_v \| \| \mathbf{q}_{v+1} \|}, \tag{15}$$



(a)Input camera image (b)Detected image patches

FIG. 4 Detect image patches with specific structures from camera image $\mathbf{I}_k$ for the similarity or the fitness computation.

where $<,>$ is the dot product and $\|\cdot\|$ is the Euclidean norm.

Since a population of particles in the EPSO algorithm denotes the potential solutions in a dynamic system, we define the i-th particle $\mathbf{p}_{i,j}$ at iteration j as a seven-dimensional (7-D) vector with respect to the endoscopic camera 6DoF pose with position $\tilde{\mathbf{t}}_k$ and orientation (quaternion) $\tilde{\mathbf{q}}_k$ at time k:

$$\mathbf{p}_{i,j} = \begin{pmatrix} \tilde{\mathbf{t}}_k & \tilde{\mathbf{q}}_k \end{pmatrix}_{7\times 1}, \tag{16}$$

where $\tilde{\mathbf{t}}_k = (\tilde{t}_k^x, \tilde{t}_k^y, \tilde{t}_k^z)_{3\times 1}$ and $\tilde{\mathbf{q}}_k = (\tilde{q}_k^0, \tilde{q}_k^1, \tilde{q}_k^2, \tilde{q}_k^3)_{4\times 1}$.

Particle $\mathbf{p}_{0,0}$ is initialized by the first sensor measurement: $\mathbf{p}_{0,0} = (\mathbf{t}_0 \ \mathbf{q}_0)_{7\times 1}$. So, the endoscopic camera motion is parameterized as a 7-D vector with the position and orientation.

II.C.2. EPSO-based guidance framework

The EPSO algorithm performs several steps, as discussed in Section II.B, to fuse endoscopic video images, the CT images, and the EM sensor measurements. It estimates the endoscope's 6DoF position and orientation to locate it during endoscopic examinations.

Let $\mathbf{I}_k$ and $\mathbf{I}_V$ be the k-th endoscopic camera image and 2-D virtual rendering image generated from the CT images, respectively. In our endoscopic guidance, we have two live observations: endoscopic camera image $\mathbf{I}_k$ and EM sensor measurement $\mathbf{m}_k = (\mathbf{t}_k, \mathbf{q}_k)$.

Based on current EM sensor observation $\mathbf{m}_k$, the velocity in Eq. 5 can be calculated by

$$\mathbf{v}_{i,j+1} = \omega \underbrace{\Phi(\mathbf{m}_{k-1}, \mathbf{m}_k)}_{Observation} + \Delta_p + \Delta_g. \tag{17}$$

Based current endoscopic camera observation $\mathbf{I}_k$, we define the fitness of particle $\mathbf{p}_k^{i,j}$ in accordance with Eq. 7:

$$f(\mathbf{p}_{i,j}) = \pi(\mathbf{I}_k|\mathbf{I}_V(\mathbf{p}_{i,j})), \tag{18}$$

where $\mathbf{I}_V(\mathbf{p}_{i,j})$ denotes the virtual image corresponding to particle $\mathbf{p}_{i,j}$. It is natural to define probability $\pi(\mathbf{I}_k|\mathbf{I}_V(\mathbf{p}_{i,j}))$ as the similarity between images $\mathbf{I}_k$ and $\mathbf{I}_V(\mathbf{p}_{i,j})$:

$$\pi(\mathbf{I}_k|\mathbf{I}_V(\mathbf{p}_{i,j})) = S(\mathbf{I}_k, \mathbf{I}_V(\mathbf{p}_{i,j})). \tag{19}$$

Before computing similarity $S(\mathbf{I}_k, \mathbf{I}_V(\mathbf{p}_{i,j}))$, we detect the specific patches with structural information from the camera video image $\mathbf{I}_k$ (Fig 4). Based on the mean square error measure[18,28], we calculate similarity $S(\mathbf{I}_k, \mathbf{I}_V(\mathbf{p}_{i,j}))$ in the detected patches from endoscopic camera image $\mathbf{I}_k$:

$$S(\mathbf{I}_k, \mathbf{I}_V(\mathbf{p}_{i,j})) = \frac{1}{U}\sum_{U}\frac{1}{W}\sum_{(a,b)\in W} M, \tag{20}$$

$$M = \big((\mathbf{I}_k(a,b) - \bar{\mathbf{I}}_k) - (\mathbf{I}_V(a,b) - \bar{\mathbf{I}}_V)\big)^2 \tag{21}$$

where U is the number of the detected patches, W is the number of pixels in one patch, $\bar{\mathbf{I}}_k$ and $\bar{\mathbf{I}}_V$ are the average intensities of all the detected patches from images $\mathbf{I}_k$ and $\mathbf{I}_V$.

We perform N iterations in EPSO at frame k and obtain best solution set $\mathcal{B}_k = \{\mathbf{g}_{i,j+1}\}_{j=1}^{N}$ by storing global best solution $\mathbf{g}_{i,j+1}$ at each iteration. We select optimal solution $\mathbf{g}_k^*$ from $\mathcal{B}_k$ on the basis of fitness $f(\mathbf{g}_{i,j+1})$ (Eqs. 18∼21). Eventually, the output of EPSO, which fuses the CT images, endoscopic videos, and EM sensor measurements for the endoscopic guidance, is the endoscope's position $\tilde{\mathbf{t}}_k^*$ and orientation $\tilde{\mathbf{q}}_k^*$:

$$(\tilde{\mathbf{t}}_k^* \quad \tilde{\mathbf{q}}_k^*) \longleftrightarrow \mathbf{g}_k^* = \arg\max_{\mathbf{g}_{i,j+1}\in\mathcal{B}_k} f(\mathbf{g}_{i,j+1}). \tag{22}$$

II.D. Validation

II.D.1. Information acquisition

We validated our proposed method on a dynamic phantom with an adjustable motion: $0 \sim 24$ mm. The CT images of our phantom were acquired with parameters of

512×512×1021 voxels and 0.68×0.68×0.5 mm^3. A 3-D Guidance medSAFE tracker (Ascension Technology Corporation, USA) was used as an EMT system, which includes a 9-coil
190 at flat-type transmitter as a magnetic field generator and an EMT sensor (1.5 mm, 6DoF). Endoscopic video images of size 362×370 pixels were collected at 30 frames per second using an endoscope (BF Type P260F, Olympus, Tokyo).

We investigate five approaches for endoscope location: (1) M1: reported by Schwarz et al.[10], (2) M2: a hybrid method by Mori et al.[14], (3) M3, a method to selectively use
195 EMT sensor measurements by Luo et al.[29], (4) M4, a sequential Monte Carlo-based method introduced by Luo et al.[18], (5) M5, our proposed method, as discussed in Section II.

II.D.2. Evaluation criteria

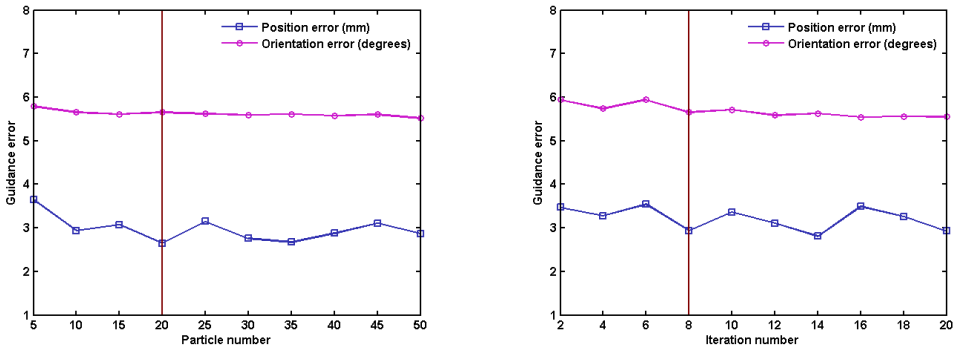
To evaluate the guidance accuracy of different methods, we generated five ground truth datasets (GTDs) by manually adjusting the position and orientation of the virtual camera to qualitatively register the video and virtual camera viewing points by hand. Two observers independently and repeatedly collected these GTDs. We clarify that intra-observer consistency was 1.81 mm and 5.9°, 1.76 mm and 4.9°, and 1.93 mm and 4.8° from the two observers, respectively; inter-observer consistency was 1.71 mm and 5.6°. Based on GTDs, position and orientation errors, ζ and ψ, are computed by:

$$\zeta = ||\mathbf{t} - \mathbf{t}_G||, \quad \epsilon = \psi(\mathbf{q}, \mathbf{q}_G), \tag{23}$$

where $\mathbf{t}$ and $\mathbf{t}_G$ are the estimated and ground-truth positions, and $\mathbf{q}$ and $\mathbf{q}_G$ are the estimated and ground-truth orientations.

Even though the accuracy and runtime of the endoscopic guidance are much significant for clinical applications, it is necessary to evaluate the smoothness of the guidance procedure, since the endoscope movement is usually a continuous procedure where the camera trajectory should be a smooth curve. Smooth guidance implies little jitter or jump errors might be involved in the endoscope movement estimation. We define the guidance smoothness as the average inter-frame distance for a series of endoscope position $\mathbf{t}_k$ and orientation $\mathbf{q}_k$ that are estimated by the five approaches:

$$\frac{1}{K-1}\sum_{k=1}^{K-1}||\mathbf{t}_{k+1} - \mathbf{t}_k||, \quad \frac{1}{K-1}\sum_{k=1}^{K-1}\psi(\mathbf{q}_{k+1}, \mathbf{q}_k). \tag{24}$$



(a)Particles (b)Iterations

FIG. 5 Experimentally determined particle and iteration numbers: $M = 20$ and $N = 8$.

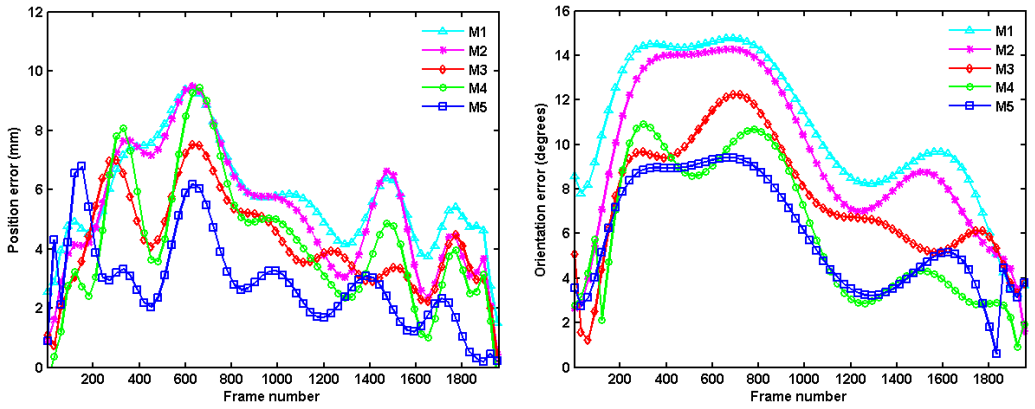
TABLE I Comparison of guidance accuracy (position, orientation) in terms of the ground truth datasets and the estimated results of using the five different methods of M1[10], M2[14], M3[18], M4[22], and M5 (our EPSO-based method).

Methods	Data 1	Data 2	Data 3	Data 4	Data 5	**Average**
M1	(4.2 mm, 6.7°)	(5.3 mm, 8.8°)	(5.6 mm, 7.9°)	(6.0 mm, 9.6°)	(7.2 mm, 13.5°)	**(5.7 mm, 9.3°)**
M2	(3.8 mm, 6.1°)	(4.9 mm, 7.6°)	(5.4 mm, 6.8°)	(5.8 mm, 8.8°)	(6.8 mm, 12.9°)	**(5.3 mm, 8.4°)**
M3	(3.1 mm, 4.8°)	(3.9 mm, 5.8°)	(4.1 mm, 6.2°)	(4.6 mm, 9.5°)	(5.6 mm, 12.9°)	**(4.3 mm, 7.8°)**
M4	(2.4 mm, 4.3°)	(3.4 mm, 5.3°)	(4.2 mm, 5.6°)	(4.4 mm, 8.2°)	(5.1 mm, 11.5°)	**(3.9 mm, 7.0°)**
M5	(2.1 mm, 3.2°)	(2.8 mm, 3.8°)	(3.0 mm, 4.1°)	(3.2 mm, 7.9°)	(4.1 mm, 9.1°)	**(3.0 mm, 5.6°)**

To evaluate the visual quality of 2-D virtual rendering images that were generated from the estimated position and orientation of the five methods, we compute normalized cross correlation (NCC) coefficient $C_{k,V}$ between each video image $\mathbf{I}_k$ and its estimated virtual image $\mathbf{I}_V$ at time k in an experiment. The average visual quality $\mathcal{Q}$ of one endoscopic video sequence for testing different methods can be defined as:

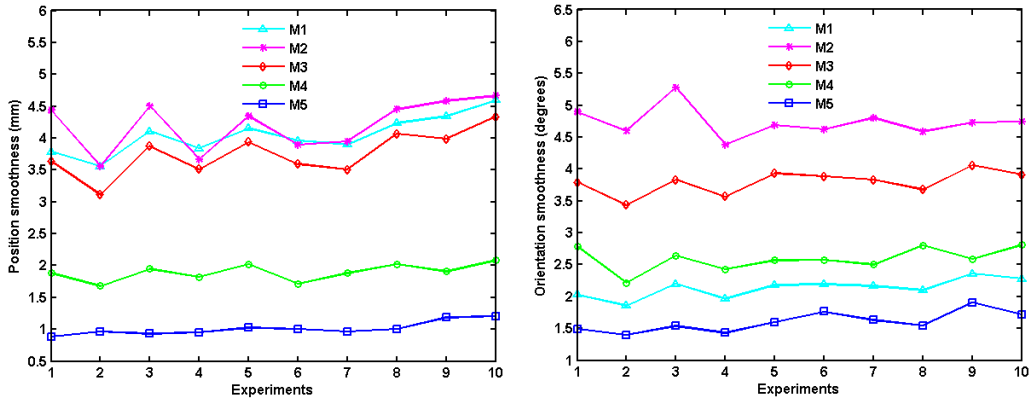
$$\mathcal{Q} = \frac{1}{K}\sum_{k=1}^{K}\frac{(C_{k,V}+1)}{2}, \tag{25}$$

where visual quality $\mathcal{Q}$ has a dynamic range of $[0,\ 1]$.



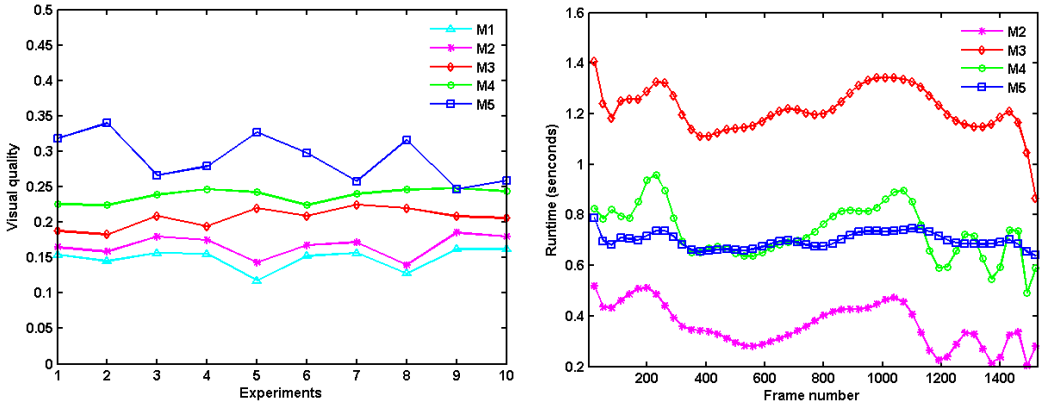
(a)Position error (b)Orientation error

FIG. 6 Plotted the guidance (position and orientation) accuracy of the five methods on Data 3.



(a)Position smoothness (b)Orientation smoothness

FIG. 7 Position and orientation smoothness of the five methods on ten experiments.



(a)Visual quality (b)Runtime

FIG. 8 Comparison of visual quality and computational time of the five methods.

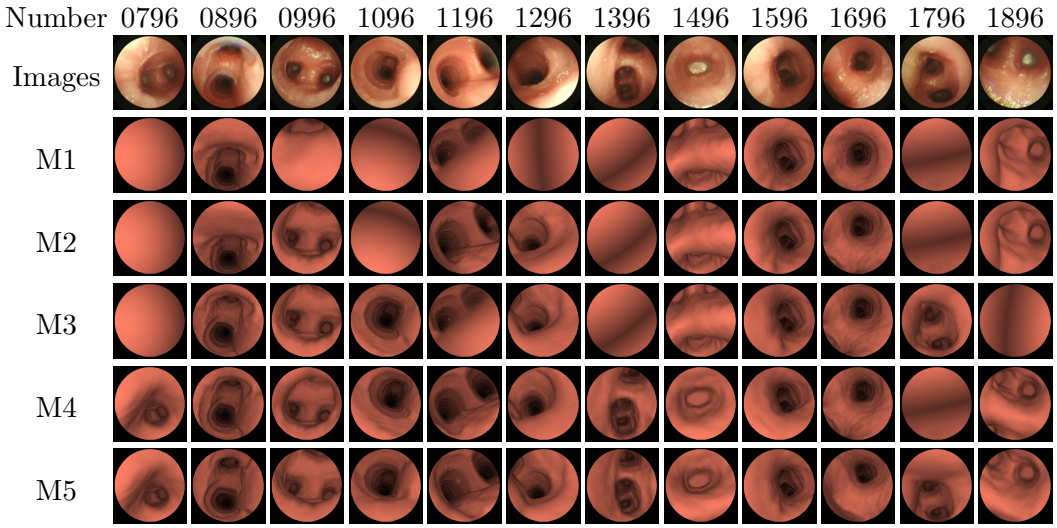


FIG. 9 Visualized video and virtual images to investigate the visual quality of different methods. Top row shows frame numbers selected uniformly at every 100 frames, and second row shows their corresponding real images. Other rows display virtual images generated from tracking results using the methods discussed above. Our method displays much better quality.

III. RESULTS

We determined iteration and particle numbers experimentally. Fig. 5 shows the position and orientation errors under different iterations and particles. We chose particle and iteration numbers, $M = 20$ and $N = 8$, since they can balance the accuracy and computational time of our proposed method.

Table I summarizes the guidance accuracy of the five methods. The average accuracy or error of our proposed EPSO-based method (M5) was about 3.0 mm and 5.6° while the previous methods (M1, M2, M3, and M4) show at least 3.9 mm and 7.0°. Fig. 6 plots the guidance accuracy of different methods on Data 3. Beyond using GTDs for assessment, we further performed ten experiments where we did not generate ground truth for them. Fig. 7 shows the smoothness of the five methods evaluated on ten experiments. The average position and orientation smoothness of our method was 1.0 mm and 1.6°, which is significantly better than other methods M1 (4.0 mm and 2.1°), M2 (4.2 mm and 4.7°), M3 (3.7 mm and 3.8°), and M4 (2.0 mm and 2.6°). Fig. 8 (a) compares the visual quality of the five methods on ten experiments. The average quality of methods of M1, M2, M3, M4, and M5 were 0.15, 0.17, 0.23, 0.25, and 0.29, respectively. Our proposed method shows better visual quality

than others. Fig. 9 further demonstrates that our method can obtain much better visual quality since the virtual images from our proposed approach resemble video images greatly better than other methods.

Additionally, Fig. 8 (b) displays the computational times of the five different methods. The current runtime of our proposed method was 0.71 seconds per frame (spf). The methods of M2, M3, and M4 were 0.45, 1.21, and 0.78 spf. The method of M1 can meet the real-time requirement during the endoscopic guidance.

IV. DISCUSSION

IV.A. Effectiveness

Basically, the objective of this work is to effectively and accurately fuse multimodal information of pre-operative images, endoscopic video sequences, and positional sensor measurements to determine endoscope position and orientation during endoscopic surgery. Our proposed EPSO algorithm provides a more accurate and robust framework to guide the endoscope than previous approaches. We attribute such an advantageous performance of EPSO to several aspects.

First, we believe that EPSO is partly an association of PSO iterations and sequential Monte Carlo (SMC) sampling procedures, and hence it outperforms SMC sampling algorithms in endoscope navigation[18]. During SMC sampling procedures, a successful particle sampling depends heavily on the proposal distribution function[30]. Particles with large fitnesses or weights located in the useful area of the proposal distribution are usually sampled. In fact, the proposal distribution is suggested to be the dynamic transition distribution, which may incur particles with larger fitness that are not sampled when the useful area of the transition distribution stays at the tail of the observation distribution[30]. However, PSO performs more like a hierarchical sampling strategy which propagates the particles integrated with the newest observations[22], possibly resolving the particle impoverishment problem. Next, automatically or adaptively controlling evolutionary parameters is greatly helpful to update particles in iterations. The two acceleration factors, which were calculated based on the fitness value from the image intensity information, are more reasonable than setting them to 2.0 in standard PSOs[31]. Moreover, the inertia weight is also adaptively

determined by both spatial distance constraint and image intensity information, resulting in more flexibly balancing the global and local search abilities and providing a reasonable velocity limitation to move particles. Finally, without any *resampling process* in our method, compared to SMC sampling or particle filtering, it is helpful to reduce the runtime of our method.

IV.B. Potential Limitations

We must clarify the potential limitations of our proposed method. The major limitation is that the particle fitness might be incorrectly computed and evaluated during particle propagation since it depends somewhat on the quality of endoscopic camera images. However, image artifacts, e.g., complex reflectance and motion blurring, which possibly occur in endoscopic video sequences, might collapse the correct computation of the fitness value. On the other hand, the particle fitness value also indicates the particle's exploitation capacity that means sufficient fitness values to obtain even better solutions from good particles: the higher the particle fitness, the more powerful particle exploitation. To tackle the incorrect fitness computation problem and obtain the powerful exploitation capacity, a more robust image similarity measure, which should be greatly insensitive to illumination changes or other image artifacts, is expected to develop in the future. Additionally, another open issue, which also remains very challenging, is the computational efficiency of our proposed method that is currently about 0.71 seconds per frame. In our proposed EPSO approach, the main computational effort lies in the calculation of each particle's fitness in the population, which is vastly time-consuming since 2-D virtual rendering images need to be generated by performing the volume rendering procedure. Note that we did neither speed optimizations nor multi-threading in our implementation. Furthermore, it is also possible to use graphics processing unit (GPU) techniques to accelerate our implementation to a real-time processing.

V. CONCLUSIONS

This article proposed an enhanced particle swarm optimization algorithm. It combines the current observation into the population propagation and refreshes its evolutionary parameters on the basis of the spatially distributed constraint of particles and the observation

information during iterations. Therefore, it is more accurate and robust to approximate the optimal solution for a stochastic problem. With its application to the multimodal information fusion for the endoscopic guidance, the experimental results demonstrate that our proposed method provides a more advantageous guidance performance than state-of-the-art methods. The average position and orientation errors were greatly reduced from (4.3, 7.8) to (3.0 mm, 5.6°). The average position and orientation smoothnesses were also significantly improved from (3.7, 3.8) to (1.0 mm and 1.6°), which makes our guidance less jitter or jumps during endoscopic surgery. Future work includes improvement of the particle fitness calculation to correctly evaluate particles and reduction of the computational time of our proposed method.

REFERENCES

[a)] Author to whom correspondence should be addressed; Electronic mail: xluo@robarts.ca

[1] F. J. F. Herth, R. Eberhardt, and A. Ernst, "The future of bronchoscopy in diagnosing, staging and treatment of lung cancer," Respiration **73**, 399–409 (2006).

[2] A. J. Chung, F. Deligianni, P. Shah, A. Wells, and G. Z. Yang, "Patient-specific bronchoscopy visualization through BRDF estimation and disocclusion correction," IEEE Transactions on Medical Imaging **25**, 503–513 (2006).

[3] J. P. Helferty, A. J. Sherbondy, A. P. Kiraly, and W. E. Higgins, "Computer-based system for the virtual-endoscopic guidance of bronchoscopy," Computer Vision and Image Understanding **108**, 171–187 (2007).

[4] X. Luo, M. Feuerstein, D. Deguchi, T. Kitasaka, H. Takabatake, and K. Mori, "Development and comparison of new hybrid motion tracking for bronchoscopic navigation," Medical Image Analysis **16**, 577–596 (2012).

[5] L. Y. Ching, K. Moller, and J. Suthakorn, "Non-radiological colonoscope tracking image guided colonoscopy using commercially available electromagnetic tracking system," in *Proc. IEEE RAM'10* (2010) pp. 62–67.

[6] A. Joshi, D. Scheinost, R. Globinsky, K. P. Vives, D. D. Spencer, L. H. Staib, and X. Papademetris, "Augmented inline-based navigation for stereotactic image guided neurosurgery," in *Proc. IEEE ISBI'11* (2011) pp. 1869–1872.

[7]D.J.Mirota, H. Wang, R. Taylor, M. Ishii, G. Gallia, and G. Hager, "A system for video-based navigation for endoscopic endonasal skull base surgery," IEEE Transactions on Medical Imaging **31**, 963–76 (2012).

[8]Http://www.ndigital.com/.

[9]H. D. Becker, F. Herth, A. Ernst, and Y. Schwarz, "Bronchoscopic biopsy of peripheral lung lesions under electromagnetic guidance: a pilot study," Journal of Bronchology and Interventional Pulmonology **12**, 9–13 (2005).

[10]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 **129**, 988–994 (2006).

[11]D. Makris, A. Scherpereel, S. Leroy, B. Bouchindhomme, J. B. Faivre, J. Remy, P. Ramon, and C. H. Marquette, "Electromagnetic navigation diagnostic bronchoscopy for small peripheral lung lesions," European Respiratory Journal **29**, 1187–1192 (2007).

[12]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 **138**, 1316–1321 (2010).

[13]I. Khan, R. Chin, N. Adair, A. Chatterjee, E. Haponik, and J. Conforti, "Electromagnetic navigation bronchoscopy in the fiagnosis of peripheral lung lesions," Clinical Pulmonary Medicine **18**, 42–45 (2011).

[14]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. MICCAI'05*, Vol. 3750 (2005) pp. 543–550.

[15]T. Reichl, X. Luo, M. Menzel, H. Hautmann, K. Mori, and N. Navab, "Deformable registration of bronchoscopic video sequences to ct volumes with guaranteed smooth output," in *Proc. MICCAI'11*, Vol. 6891 (2011) pp. 17–24.

[16]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 MI'10*, Vol. 7625 (2010) p. 76250W.

[17]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 Transactions on Biomedical Engineering **57**, 736–745 (2010).

[18]X. Luo, T. Reichl, M. Feuerstein, T. Kitasaka, and K. Mori, "Modified hybrid bronchoscope tracking based on sequential monte carlo sampler: dynamic phantom validation," in *Proc. ACCV'10*, Vol. 6494 (2010) pp. 409–421.

[19]I. Gergel, J. Hering, R. Tetzlaff, H.-P. Meinzer, and I. Wegner, "An electromagnetic navigation system for transbronchial interventions with a novel approach to respiratory motion compensation," Medical Physics **38**, 6742–6753 (2011).

[20]J. Kennedy and R. C. Eberhart, "Particle swarm optimization," in *Proc. IEEE Int. Conf. on Neural Networks'95* (1995) pp. 1942–1948.

[21]S. Mostaghim and J. Teich, "Strategies for finding good local guides in multi-objective particle swarm optimization," in *Proc. IEEE ISI'03* (2003) pp. 26–33.

[22]X. Zhang, W. Hu, W. Qu, and S. Maybank, "Multiple object tracking via species-based particle swarm optimization," IEEE Transactions on Circuits and Systems for Video Technology **22**, 1590–1602 (2010).

[23]B. Xue, M. Zhang, and W. N. Browne, "Particle swarm optimization for feature selection in classification: A multi-objective approach," IEEE Transactions on Cybernetics **43**, 1656–1671 (2013).

[24]D. Parrott and X. Li, "Locating and tracking multiple dynamic optima by a particle swarm model using speciation," IEEE Transactions on Evolutionary Computation **10**, 440–458 (2006).

[25]I. Bricault, G. Ferretti, and P. Cinquin, "Registration of real and CT-derived virtual bronchoscopic images to assist transbronchial biopsy," IEEE Transactions on Medical Imaging **17**, 703–714 (1998).

[26]L. Zhang, T. Mei, Y. Liu, D.Tao, and H. Q. Zhou, "Visual search reranking via adaptive particle swarm optimization," Pattern Recognition **44**, 1811–1820 (2011).

[27]D. Salomon, *The computer graphics manual* (Springer, 2011).

[28]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," Medical Image Analysis **13**, 621–633 (2009).

[29]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'10*, Vol. 7625 (2010) p. 76251B.

[30]M. S. Arulampalam, S. Maskell, and N. Gordon, "A tutorial on particle filters for nonlinear/non-gaussian Bayesian tracking," IEEE Transaction on Singal Processing **50**, 174–188 (2002).

[31]J. Kennedy and R. C. Eberhart, "Particle swarm optimization," in *Proc. IEEE ICNN'95* (1995) pp. 1942–1948.