

Diversity-Enhanced Condensation Algorithm and Its Application for Robust and Accurate Endoscope Three-Dimensional Motion Tracking

Xiongbiao Luo^{1*} Ying Wan^{2†} Xiangjian He² Jie Yang³ Kensaku Mori¹
Nagoya University¹ University of Technology, Sydney² Shanghai Jiaotong University³
Xiongbiao.Luo@gmail.com Ying.Wan@student.uts.edu.au Xiangjian.He@uts.edu.au

Abstract

The paper proposes a diversity-enhanced condensation algorithm to address the particle impoverishment problem which stochastic filtering usually suffers from. The particle diversity plays an important role as it affects the performance of filtering. Although the condensation algorithm is widely used in computer vision, it easily gets trapped in local minima due to the particle degeneracy. We introduce a modified evolutionary computing method, adaptive differential evolution, to resolve the particle impoverishment under a proper size of particle population. We apply our proposed method to endoscope tracking for estimating three-dimensional motion of the endoscopic camera. The experimental results demonstrate that our proposed method offers more robust and accurate tracking than previous methods. The current tracking smoothness and error were significantly reduced from (3.7, 4.8) to (2.3 mm, 3.2 mm), which approximates the clinical requirement of 3.0 mm.

1. Introduction

Stochastic filtering methods are widely introduced in dynamic state estimation problems, e.g., target tracking, video surveillance, and camera 3-D motion tracking [9, 15, 7]. As one of stochastic filtering methods, the condensation algorithm (CA), which was originally proposed to detect and track the contour of objects moving in a cluttered environment [6], is widely used in the field of computer vision and pattern recognition to solve nonlinear state estimation problems. It employs the sampling importance resampling (SIR) strategy to estimate dynamic system states, i.e., the original part of CA is the application of particle filtering estimation techniques. Compared to Kalman filtering, CA has the ability to tackle multimodal probability density functions and tractably solve nonlinear non-Gaussian problems [1]. Unfortunately, *particle degeneracy* or *impoverishment*, which

frequently occurs in the SIR step of CA, constrains its ability at finding the optimal solution for dynamic systems since less of particle modes hardly approximates the probability density function (PDF) of the current state. In an extreme situation, a number of different particles might collapse to a set of particles with same state weights and parameters after a sequence of updates. In an usual situation, even though several particles were updated to be different from each other, it is somewhat impossible to approach the expected PDF. Although various work to address such a problem has published in the literature [2, 4, 5, 13, 12, 8], it still remains challenging in CA.

Currently, as one of powerful evolutionary computation algorithms, differential evolution (DE), which was primarily proposed by Storn et al. [14], is widely introduced to address multidimensional complex optimization problems. It has the ability to deal with non-differentiable, nonlinear and multimodal optimization problems under easily selecting few parameters to control the optimization procedure [14]. It was demonstrated to be the best evolutionary algorithm for tackling real parameter optimization problems [3]. The attractive properties of DE are easy implementation, better convergence performance, few control parameters, and low space complexity [3], resulting in its popularity in global optimization over continuous dynamic state estimation.

This work aims to address the particle impoverishment problem of stochastic filtering methods. By inspired these unique properties of DE, our basic idea is to use the DE algorithm to deal with the particle impoverishment in CA. We propose a diversity-enhanced condensation algorithm (DECA) that first differentially evolves each particle to enhance the diversity of particle population before the SIR step, and then transmits each particle by a state propagation model to approximate the posterior density distribution.

The main contribution of this work is summarized as follows. First, since the DE's performance depends heavily on evolutionary parameters of the mutation factor and crossover rate, we modified DE to be adaptive differential evolution (ADE) that automatically determines the evolutionary parameters on the basis of the current observation

*indicates corresponding author.

†denotes equal contribution with corresponding author.

information. More importantly, an ADE-based strategy was introduced to successfully handle the particle impoverishment problem. Finally, from an application point of view, we apply DECA to endoscope 3-D motion tracking and our experimental results prove that it can provide a more accurate and robust tracking method than previous approaches. We also believe that DECA are definitely useful for other applications, e.g., object tracking and video surveillance.

The rest of this paper consists of six parts. We first recall the standard CA in Section 2. Section 3 details the proposed DECA on the basis of the ADE method to handle the particle impoverishment. Section 4 applies DECA to endoscope 3-D motion tracking. Experimental setups are described in Section 5, followed by experimental results and discussion in Section 6, before concluding this work in Section 7.

2. Condensation Algorithm

Basically, CA is one of sequential Monte Carlo methods to solve the recursive Bayesian filtering problem. It uses a set of weighted particles to approximate the posterior density distribution and recursively search for the optimal estimation at each state of one stochastic dynamic system on the basis of noisy partial observations.

Let \mathbf{x}_i and \mathbf{y}_i be the current state and observation at time i of a dynamic system. The history observations are $\mathcal{Y}_i = \{\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_i\}$ ($i = 1, \dots, N$, N is the number of observations). CA seeks to approximate posterior probability distribution $p(\mathbf{x}_i|\mathcal{Y}_i)$ of current state \mathbf{x}_i . It first generates a set of particles $\mathcal{X}_i = \{(\mathbf{x}_i^j, \omega_i^j, c_i^j), j = 1, \dots, M\}$ (M is the number of particles) with particle weight ω_i^j and accumulative weight c_i^j . After that, $p(\mathbf{x}_i|\mathcal{Y}_i)$ is approximated by these particles with respect to \mathbf{x}_i^j and ω_i^j [1]:

$$p(\mathbf{x}_i|\mathcal{Y}_i) \approx \sum_{j=1}^M \omega_i^j \delta(\mathbf{x}_i - \mathbf{x}_i^j), \quad (1)$$

where $\delta(\cdot)$ is the Dirac delta function; ω_i^j is described by:

$$\omega_i^j \propto \omega_{i-1}^j \frac{p(\mathbf{y}_i|\mathbf{x}_i^j)p(\mathbf{x}_i^j|\mathbf{x}_{i-1}^j)}{q(\mathbf{x}_i^j|\mathbf{x}_{i-1}^j, \mathbf{y}_i)}, \quad (2)$$

where proposal $q(\cdot)$ is an importance density function relative to the degree of the particle degeneracy. It is convenient to select $q(\cdot)$ as prior $p(\mathbf{x}_i^j|\mathbf{x}_{i-1}^j)$: $q(\mathbf{x}_i^j|\mathbf{x}_{i-1}^j, \mathbf{y}_i) = p(\mathbf{x}_i^j|\mathbf{x}_{i-1}^j)$; we then obtain: $\omega_i^j \propto \omega_{i-1}^j p(\mathbf{y}_i|\mathbf{x}_i^j)$ [1].

Essentially, a pseudo-code description of CA using the SIR scheme can be summarized in **Algorithm 1**.

3. Proposed Method – DECA

Although CA works well in realistic state estimation problems, it suffers from the particle impoverishment that

Algorithm 1: Condensation algorithm [6]

```

if  $i = 0$  then
  Generate  $M$  particles  $\mathcal{X}_0 = \{(\mathbf{x}_0^j, \omega_0^j, c_0^j)\}_{j=1}^M$ ;
  Set initial importance density  $q(\mathbf{x}_0|\mathbf{x}_0, \mathbf{y}_0) = p(\mathbf{x}_0)$ ;
  for  $j = 1$  to  $M$  ( $M$ : the number of particles) do
    Draw particle  $\mathbf{x}_0^j \sim p(\mathbf{x}_0)$ ;
    Compute particle weight  $\omega_0^j = 1/M$ ;
    Set accumulative weight  $c_0^j = 0$ ;
  end
else
  for  $i = 1$  to  $N$  ( $N$ : the number of observation  $\mathbf{y}_i$ ) do
    ❶ for  $j = 1$  to  $M$  ( $M$ : the number of particle  $\mathbf{x}_i$ ) do
      1. Choose particle  $\hat{\mathbf{x}}_i^j$ ;
      Generate uniform random number  $r \in [0, 1]$ ;
      Find the smallest  $j$  to satisfy:  $c_{i-1}^j \geq r$ ;
      Choose  $\hat{\mathbf{x}}_i^j = \mathbf{x}_{i-1}^j$ ;
      2. Propagate  $\hat{\mathbf{x}}_i^j$  to  $\mathbf{x}_i^j$  by a transition model;
      3. Compute weight  $\omega_i^j = p(\mathbf{y}_i|\mathbf{x}_{i-1} = \mathbf{x}_i^j)$ ;
    end
    ❷ Normalize  $\omega_i^j = \omega_i^j / \sum_{j=1}^M \omega_i^j, \sum_{j=1}^M \omega_i^j = 1$ ;
    ❸ Update accumulative weight  $c_i^j = c_{i-1}^j + \omega_i^j$ ;
    ❹ Calculate the expectation of  $\mathbf{x}_i$  using Eq. 1;
    ❺ Store the estimate and go to the next iteration;
  end
end

```

easily collapses the procedure of state estimations since few particle modes are impossible to approximate the required PDF. In this work, we introduce an ADE strategy to handle the particle degeneracy problem that is originated from the SIR step (1. Choose particle $\hat{\mathbf{x}}_i^j$ in **Algorithm 1**), observation noise increasing, and estimation error accumulation.

Basically, DECA consists of several steps: (1) ADE-based particle diversification, (2) particle transition on the basis of a dynamic model, and (3) observation model to construct the probabilistic density, as described as follows.

3.1. ADE-Diversified Particles

The section uses ADE to diversify the particles as different as possible to avoid the particle impoverishment. Given particles $\mathcal{X}_{i-1} = \{(\mathbf{x}_{i-1}^j, \omega_{i-1}^j, c_{i-1}^j)\}_{j=1}^M$ ($\mathbf{x}_{i-1}^j \in \mathbb{R}^D$, D is the dimension of state \mathbf{x}_{i-1}^j) ADE performs three operations: (1) mutation, (2) crossover, and (3) selection, to enhance the diversity of particle population \mathcal{X}_{i-1} .

3.1.1 Mutation

Let \mathbf{x}_{i-1}^b denotes the particle with the best or maximal weight in particle population \mathcal{X}_{i-1} . For each particle \mathbf{x}_{i-1}^j , its mutant vector \mathbf{v}_{i-1}^j can be calculated by a frequently

used mutation scheme in DE algorithms [3]:

$$\mathbf{v}_{i-1}^j = \mathbf{x}_{i-1}^j + \mu_{i-1}^b \Delta_{i-1}^b + \mu_{i-1}^r \Delta_{i-1}^r, \quad (3)$$

$$\Delta_{i-1}^b = (\mathbf{x}_{i-1}^b - \mathbf{x}_{i-1}^j), \Delta_{i-1}^r = (\mathbf{x}_{i-1}^{r^1} - \mathbf{x}_{i-1}^{r^2}), \quad (4)$$

where μ_{i-1} is a mutation factor, vector Δ_{i-1}^b perturbs base state \mathbf{x}_{i-1}^j , and Δ_{i-1}^r represents the difference vector. Index number r^1 and r^2 are mutually exclusive integers chosen randomly from set $\{1, \dots, j-1, j+1, \dots, M\}$.

The mutation operation is the core of DE algorithms. The advantages and limitations of different mutation strategies were discussed in [3]. Unfortunately, none of these strategies takes the current observation into consideration, possibly resulting in the loss of the population diversity. To address such a limitation, we modify the mutation strategy in Eq. 3 by introducing the current observation information:

$$\mathbf{v}_{i-1}^j = \mathbf{x}_{i-1}^j + \lambda_i^j \Delta_{i-1}^i + \mu_{i-1}^b \Delta_{i-1}^b + \mu_{i-1}^r \Delta_{i-1}^r, \quad (5)$$

where $\Delta_{i-1}^i = \mathbf{o}_i - \mathbf{o}_{i-1}$, \mathbf{o}_i and \mathbf{o}_{i-1} are the observations at times i and $i-1$. Factor λ_i^j determines how much the current observation \mathbf{o}_i to be reserved and we set it to be an uniformly distributed random number: $\lambda_i^j \in [0, 1]$.

On the other hand, the mutation performance also depends on mutation factor μ_{i-1} . In this work, we introduce two factors μ_{i-1}^b and μ_{i-1}^r in Eq. 5. We adaptively calculate μ_{i-1}^b and μ_{i-1}^r on the basis of particle weights ω_{i-1}^b and ω_{i-1}^j of particle states \mathbf{x}_{i-1}^b and \mathbf{x}_{i-1}^j :

$$\mu_{i-1}^b = \frac{2\omega_{i-1}^b}{(\omega_{i-1}^b + \omega_{i-1}^j)}, \quad \mu_{i-1}^r = \frac{2\omega_{i-1}^j}{(\omega_{i-1}^b + \omega_{i-1}^j)}. \quad (6)$$

Mutant factors μ_{i-1}^b and μ_{i-1}^r automatically control best particle state \mathbf{x}_{i-1}^b and stochastic difference vector Δ_{i-1}^r .

3.1.2 Crossover

After the mutation step, a binomial crossover operation is performed to generate trial state $\mathbf{u}_{i-1}^j = \{u_{i-1}^{j,1}, \dots, u_{i-1}^{j,D}\}$ on the basis of target state $\mathbf{x}_{i-1}^j = \{x_{i-1}^{j,1}, \dots, x_{i-1}^{j,D}\}$ and mutant state $\mathbf{v}_{i-1}^j = \{v_{i-1}^{j,1}, \dots, v_{i-1}^{j,D}\}$:

$$u_{i-1}^{j,a} = \begin{cases} v_{i-1}^{j,a} & \text{if } (a_c \leq C_r) \text{ or } (a = a_r) \\ x_{i-1}^{j,a} & \text{otherwise} \end{cases}, \quad (7)$$

where $a = 1, 2, \dots, D$, random number a_c yields an uniform distribution, crossover rate or probability C_r determines whether or not $u_{i-1}^{j,a} \in \mathbf{u}_{i-1}^j$ is copied from $v_{i-1}^{j,a} \in \mathbf{v}_{i-1}^j$, and a_r is randomly selected from set $\{1, 2, \dots, D\}$.

The crossover rate during the binomial crossover operation is also adaptively determined in terms of weights ω_{i-1}^j and v_{i-1}^j of states \mathbf{x}_{i-1}^j and \mathbf{v}_{i-1}^j . Since C_r was suggested

Algorithm 2: Particle diversification using ADE

Input: Particles $\mathcal{X}_{i-1} = \{(\mathbf{x}_{i-1}^j, \omega_{i-1}^j, c_{i-1}^j)\}_{j=1}^M$;

Output: Particles $\hat{\mathcal{X}}_{i-1} = \{(\hat{\mathbf{x}}_{i-1}^j, \hat{\omega}_{i-1}^j, \hat{c}_{i-1}^j)\}_{j=1}^M$;

for $j = 1$ **to** M (M : the number of particles) **do**

1 Mutation:

 Randomly select $\mathbf{x}_{i-1}^{r^1}$ and $\mathbf{x}_{i-1}^{r^2}$ from \mathcal{X}_{i-1} ;

 Generate λ_i^j and compute μ_{i-1}^b and μ_{i-1}^r by Eq. 6;

 Calculate mutant vector \mathbf{v}_{i-1}^j by Eq. 5;

2 Crossover:

 Calculate crossover rate C_r by Eq. 8;

 Compute trial state \mathbf{u}_{i-1}^j using \mathbf{x}_{i-1}^j and \mathbf{v}_{i-1}^j ;

3 Selection:

 Evaluate \mathbf{x}_{i-1}^j and \mathbf{u}_{i-1}^j to determine $\hat{\mathbf{x}}_{i-1}^j$;

4 Normalize $\hat{\omega}_{i-1}^j = \frac{\hat{\omega}_{i-1}^j}{\sum_{j=1}^M \hat{\omega}_{i-1}^j}$, $\sum_{j=1}^M \hat{\omega}_{i-1}^j = 1$;

5 Accumulative weight $\hat{c}_{i-1}^j = c_{i-1}^j + \hat{\omega}_{i-1}^j - \omega_{i-1}^j$;

6 Store $(\hat{\mathbf{x}}_{i-1}^j, \hat{\omega}_{i-1}^j, \hat{c}_{i-1}^j)$ and go to the next iteration;

end

to range in the interval $[0, 1]$ for balancing the global and local searching abilities [3], it can be adaptively updated by:

$$C_r = \frac{\omega_{i-1}^j + v_{i-1}^j}{2}, \quad (8)$$

which gives an adaptive strategy to control crossover rate C_r relative to the particle weight that will be computed on the basis of the current observation information.

3.1.3 Selection

The selection operation seeks to choose the better particle or individual $\hat{\mathbf{x}}_{i-1}^j$ from \mathbf{u}_{i-1}^j in accordance with its weight u_{i-1}^j . Such an operation is performed by:

$$\hat{\mathbf{x}}_{i-1}^j = \begin{cases} \mathbf{u}_{i-1}^j & \text{if } u_{i-1}^j \geq \omega_{i-1}^j \\ \mathbf{x}_{i-1}^j & \text{otherwise} \end{cases}. \quad (9)$$

We also renew weight $\hat{\omega}_{i-1}^j$ of particle $\hat{\mathbf{x}}_{i-1}^j$:

$$\hat{\omega}_{i-1}^j = \begin{cases} u_{i-1}^j & \text{if } u_{i-1}^j \geq \omega_{i-1}^j \\ \omega_{i-1}^j & \text{otherwise} \end{cases}. \quad (10)$$

Each particle weight $\hat{\omega}_{i-1}^j$ is further normalized: $\hat{\omega}_{i-1}^j = \hat{\omega}_{i-1}^j / \sum_{j=1}^M \hat{\omega}_{i-1}^j$, $\sum_{j=1}^M \hat{\omega}_{i-1}^j = 1$. Accumulative weight \hat{c}_{i-1}^j is also updated: $\hat{c}_{i-1}^j = c_{i-1}^j + \hat{\omega}_{i-1}^j - \omega_{i-1}^j$. Finally, a set of diversified particles $\hat{\mathcal{X}}_{i-1} = \{(\hat{\mathbf{x}}_{i-1}^j, \hat{\omega}_{i-1}^j, \hat{c}_{i-1}^j)\}_{j=1}^M$ is obtained. Particle population $\hat{\mathcal{X}}_{i-1}$ not only has the better diversity but also includes the better particles to be resampled in SIR. The implementation of ADE to enhance the particle diversity is generalized in **Algorithm 2**.

3.2. Particle Transition

After obtaining particle set $\hat{\mathcal{X}}_{i-1}^j$, we resample from $\hat{\mathcal{X}}_{i-1}^j$ and get particle $\hat{\mathbf{x}}_i^j = \hat{\mathbf{x}}_{i-1}^j$. We transmit $\hat{\mathbf{x}}_i^j$ to new state or particle \mathbf{x}_i^j on the basis of current observation \mathbf{o}_i :

$$\mathbf{x}_i^j = \Delta_{i-1}^i \hat{\mathbf{x}}_i^j + \Pi(n_i^j), \quad (11)$$

where $\Pi(n_i^j)$ is a stochastic noise term with an independent variable n_i^j that is assume to yield a normal distribution.

Therefore, the conditional density $p(\mathbf{x}_i | \mathbf{x}_{i-1} = \mathbf{x}_{i-1}^j)$ of dynamic state \mathbf{x}_i can be computed by:

$$p(\mathbf{x}_i | \mathbf{x}_{i-1} = \mathbf{x}_{i-1}^j) \propto \exp\left(-\frac{(\Delta_{i-1}^i)^{-1}(\Pi(n_i^j) - \mathbf{x}_i^j)}{2}\right). \quad (12)$$

3.3. Observation Model

Since CA that is used for dynamic state estimation problems involves the history and current observations to build the PDF of the current state, the observation probability $p(\mathbf{o}_i | \mathbf{x}_i)$ needs to be determined after getting \mathbf{x}_i^j during the particle transition. The particle weights are used to approximate $p(\mathbf{o}_i | \mathbf{x}_i)$ [6]:

$$p(\mathbf{o}_i | \mathbf{x}_i = \mathbf{x}_i^j) \propto \frac{\omega_i^j}{\sum_{j=1}^M \omega_i^j}. \quad (13)$$

The processing within several steps of DECA was discussed above. The performance of DECA will be evaluated by applying it to endoscope 3-D motion tracking. DECA will be proved to outperform CA in realistic applications.

4. Application to Endoscope Tracking

This section shows utilizing DECA for endoscope 3-D motion tracking. An endoscope, which is inserted into the body by physicians to directly observe hollow organs in the operating room, is usually integrated with a video camera at its distal tip. Endoscope tracking seeks to find endoscope position and orientation with six degrees of freedom (6DoF) in a reference coordinate system, e.g., a computed tomography (CT) image coordinate system in our case. To track the endoscope, an electromagnetic sensor is usually attached at the endoscope distal tip. Therefore, during endoscope tracking, the input information includes CT images and two kinds of observation: (1) endoscopic video images and (2) electromagnetic sensor measurements, and the outputs are the endoscope 6DoF position and orientation in the CT space. We first parameterize endoscope 3-D motion.

4.1. Motion Parameterization

Our task is to estimate endoscopic camera position and orientation in the CT space, i.e., transformation ${}^{CT}\mathbf{T}_C$ from

Algorithm 3: Endoscope tracking using DECA

Input: CT images to generate virtual rendering image \mathbf{m}_i^j , electromagnetic sensor measurement \mathbf{s}_i , and endoscopic video image \mathbf{m}_i ;

Output: Endoscope 3-D motion estimation ${}^{CT}\mathbf{T}_C^i$;

if $i = 0$ **then**

 Initialize ${}^{CT}\mathbf{T}_C^0 = \mathbf{x}_0 = [{}^{CT}\mathbf{t}_C^0, {}^{CT}\mathbf{k}_C^0]$;

 Generate M particles $\mathcal{X}_0 = \{(\mathbf{x}_0^j, \omega_0^j, c_0^j)\}_{j=1}^M$;

for $j = 1$ **to** M (M particles) **do**

$\mathbf{x}_0^j = \mathbf{x}_0$, $\omega_0^j = 1/M$, $c_0^j = 0$;

end

else

for $i = 1$ **to** N (N observations of \mathbf{s}_i and \mathbf{m}_i) **do**

 ① Call **Algorithm 2** and obtain $\hat{\mathcal{X}}_{i-1}$;

 ② Particle resample and propagation;

for $j = 1$ **to** M (M particles) **do**

 ① Select particle $\hat{\mathbf{x}}_i^j = \hat{\mathbf{x}}_{i-1}^j$ from $\hat{\mathcal{X}}_{i-1}$;

 ② Transmit particle $\hat{\mathbf{x}}_i^j$ to \mathbf{x}_i^j using Eq. 18;

 ③ Calculate particle weight using Eq. 19;

end

 ③ Normalize $\omega_i^j = \omega_i^j / \sum_{j=1}^M \omega_i^j$, $\sum_{j=1}^M \omega_i^j = 1$;

 ④ Update accumulative weight $c_i^j = c_{i-1}^j + \omega_i^j$;

 ⑤ Find the optimal estimation $\hat{\mathbf{x}}_i$ using Eq. 20;

 ⑥ Store $\hat{\mathbf{x}}_i \leftrightarrow {}^{CT}\mathbf{T}_C^i$ and go to the next iteration;

end

end

the endoscopic camera to CT coordinate systems:

$${}^{CT}\mathbf{T}_C = \begin{pmatrix} {}^{CT}\mathbf{R}_C & {}^{CT}\mathbf{t}_C \\ \mathbf{0}^T & 1 \end{pmatrix}, \quad (14)$$

where translation or position vector ${}^{CT}\mathbf{t}_C$ has three components: ${}^{CT}\mathbf{t}_C = [{}^{CT}t_C^x, {}^{CT}t_C^y, {}^{CT}t_C^z]$. Orientation ${}^{CT}\mathbf{R}_C$ is characterized by a quaternion ${}^{CT}\mathbf{k}_C$ with four elements $[{}^{CT}k_C^0, {}^{CT}k_C^x, {}^{CT}k_C^y, {}^{CT}k_C^z]$ in the CT x -, y -, and z -axes:

$${}^{CT}\mathbf{R}_C \longleftrightarrow {}^{CT}\mathbf{k}_C, \quad (15)$$

$$({}^{CT}k_C^0)^2 + ({}^{CT}k_C^x)^2 + ({}^{CT}k_C^y)^2 + ({}^{CT}k_C^z)^2 = 1. \quad (16)$$

Therefore, endoscope 3-D motion ${}^{CT}\mathbf{T}_C^i$ at time i and particle \mathbf{x}_i can both be parameterized as a seven-dimensional vector ($D = 7$) with respect to ${}^{CT}\mathbf{t}_C^i$ and ${}^{CT}\mathbf{k}_C^i$ in tracking:

$${}^{CT}\mathbf{T}_C^i \longleftrightarrow \mathbf{x}_i = \left[{}^{CT}\mathbf{t}_C^i \quad {}^{CT}\mathbf{k}_C^i \right]_{7 \times 1}^T, \quad (17)$$

Hence, endoscope 3-D motion tracking aims to determine transformation ${}^{CT}\mathbf{T}_C^i$ or the optimal particle \mathbf{x}_i using the proposed DECA in our case.

4.2. Optimal Estimation Selection

At time i , observation \mathbf{o}_i involves two kinds of outputs: $\mathbf{o}_i = \{\mathbf{s}_i, \mathbf{m}_i\}$, where \mathbf{s}_i denotes the current electromagnetic sensor output and \mathbf{m}_i is the current video image. We diversify $\mathcal{X}_{i-1} = \{(\mathbf{x}_{i-1}^j, \omega_{i-1}^j, c_{i-1}^j)\}_{j=1}^M$ and obtain $\hat{\mathcal{X}}_{i-1} = \{(\hat{\mathbf{x}}_{i-1}^j, \hat{\omega}_{i-1}^j, \hat{c}_{i-1}^j)\}_{j=1}^M$. We randomly select $\hat{\mathbf{x}}_i^j$ from $\hat{\mathcal{X}}_{i-1}$ and propagate $\hat{\mathbf{x}}_i^j$ deterministically and stochastically to new state or particle \mathbf{x}_i^j :

$$\mathbf{x}_i^j = \mathcal{F}(\Delta_{i-1}^i \hat{\mathbf{x}}_i^j, \Pi(n_i^j)), \quad (18)$$

where $\mathcal{F}(\cdot)$ is a transform function. Note that Δ_{i-1}^i is equal to $(\mathbf{s}_i - \mathbf{s}_{i-1})$ in particle diversification (Eq. 5) and propagation (Eq. 18) during endoscope 3-D motion tracking.

After particle propagation (Eq. 18), we obtain new particle set $\mathcal{X}_i = \{\mathbf{x}_i^j\}_{j=1}^M$. Weight ω_i^j and accumulative weight c_i^j of each particle $\mathbf{x}_i^j \in \mathcal{X}_i$ must be determined. We here use another observation information, endoscopic video image \mathbf{m}_i , to compute these weights.

We define the particle weight as an image similarity value between current video image \mathbf{m}_i and virtual rendering image \mathbf{m}_i^j [16]: $\omega_i^j = s(\mathbf{m}_i, \mathbf{m}_i^j)$, where \mathbf{m}_i^j was generated from 3-D motion state \mathbf{x}_i^j using volume or surface rendering techniques in terms of CT data. Based on such a definition, the similarity value can be calculated by:

$$s(\mathbf{m}_i, \mathbf{m}_i^j) = \frac{1}{2} \left(1 + \frac{4\vartheta_{1,2}\kappa_1\kappa_2}{(\vartheta_1^2 + \vartheta_2^2)(\kappa_1^2 + \kappa_2^2)} \right), \quad (19)$$

where $\vartheta_{1,2}$ is the correlation between images \mathbf{m}_i and \mathbf{m}_i^j ; ϑ_1 and ϑ_2 are the covariances of \mathbf{m}_i and \mathbf{m}_i^j , and κ_1 and κ_2 are the average values of \mathbf{m}_i and \mathbf{m}_i^j . Note that $s(\mathbf{m}_i, \mathbf{m}_i^j)$ ranges within the interval $[0, 1]$. We also update the accumulative weight: $c_i^j = c_{i-1}^j + \omega_i^j$.

Each particle \mathbf{x}_i^j in $\mathcal{X}_i = \{(\mathbf{x}_i^j, \omega_i^j, c_i^j)\}_{j=1}^M$ is a potential solution or estimation for current endoscope 3-D motion parameters ${}^{CT}\mathbf{T}_C^i$. We search for the optimal particle that has the maximal particle weight:

$$\tilde{\mathbf{x}}_i = \arg \max_{\omega_i^j \in \mathcal{X}_i} \{\mathbf{x}_i^j \leftrightarrow \omega_i^j\}, \quad (20)$$

which is the optimal estimation for current 3-D pose ${}^{CT}\mathbf{T}_C^i$:

$$\tilde{\mathbf{x}}_i = \left[{}^{CT}\tilde{\mathbf{t}}_C^i \quad {}^{CT}\tilde{\mathbf{k}}_C^i \right]^T \mapsto {}^{CT}\mathbf{T}_C^i. \quad (21)$$

5. Experiments

We apply our proposed DECA to endoscope 3-D motion tracking and evaluate on seven datasets of endoscopic videos and electromagnetic sensor outputs (in total 10468 video images or sensor measurements). We manually generated ground truth for these datasets. Hence we can compute

tracking position and orientation errors by:

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

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.

Theoretically, endoscope movement is a continuous procedure, i.e., its motion path should be a smoothing curve. Since it is impossible for a camera to record all continuous motion information without any interval, it is necessary to evaluate the smoothness of tracking results. We define the smoothness as the average Euclidean distance and standard deviation of estimated positions among continuous frames as well as orientations and compute them by :

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

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

where N is the number of video images. The larger of τ and ψ , the more unsmoothed of estimated motion results, and the more jitter or jump occurred in tracking results.

We also introduce a visual quality measure to evaluate the tracking results estimated from different approaches. The 3-D motion estimation parameters $\{{}^{CT}\mathbf{T}_C^i\}_{i=1}^M$ can be used to generate virtual rendering images on the basis of CT images. It is very important to inspect whether or not these virtual rendering images resemble to endoscopic video images to provide an augmented visualization environment in the operating room. After getting all 3-D motion estimations, visual quality $\xi(\mathbf{m}_r, \mathbf{m}_v)$ between real video image \mathbf{m}_r and its corresponding virtual rendering image \mathbf{m}_v generated from its estimation is defined in terms of Eq. 19:

$$\xi(\mathbf{m}_r, \mathbf{m}_v) = \frac{1}{2} \left(1 + \frac{4\vartheta_{r,v}\kappa_r\kappa_v}{(\vartheta_r^2 + \vartheta_v^2)(\kappa_r^2 + \kappa_v^2)} \right). \quad (25)$$

The larger $\xi(\mathbf{m}_r, \mathbf{m}_v)$, the better visual quality of the tracking results, and the more accurate tracking method.

We compare our proposed DECA to several endoscope 3-D motion tracking methods (1) Schwarz et al. [11], directly using absolute electromagnetic sensor measurements; (2) Mori et al. [10], a hybrid tracking method; (3) Luo et al. [7], using CA to track endoscope movements.

6. Results and Discussion

Table 1 summarize the tracking position and orientation errors of using the four approaches. The average tracking errors were significantly reduced at least from (4.8, 8.0) to (3.2 mm, 6.5°). Figure 1 plots the tracking errors against

Table 1: Quantitative comparison of average position and orientation errors, ζ , ϕ , of using the different methods (Eq. 22).

Methods Experiments	Schwarz et al. [11]		Mori et al. [10]		Luo et al. [7]		DECA	
	ζ	ϕ	ζ	ϕ	ζ	ϕ	ζ	ϕ
A	6.5 mm	7.7°	5.2 mm	3.7°	4.4 mm	4.2°	3.3 mm	3.7°
B	5.3 mm	11.6°	5.0 mm	10.2°	4.4 mm	11.6°	2.7 mm	10.2°
C	5.4 mm	3.9°	5.2 mm	9.9°	4.7 mm	4.3°	3.6 mm	3.8°
D	4.5 mm	10.0°	3.9 mm	8.2°	3.5 mm	10.0°	2.9 mm	8.6°
E	9.1 mm	10.4°	5.2 mm	4.0°	4.5 mm	4.5°	3.6 mm	4.0°
F	5.1 mm	12.4°	4.2 mm	11.2°	3.6 mm	12.4°	2.8 mm	10.5°
G	10.6 mm	9.5°	10.3 mm	10.2°	8.7 mm	9.6°	3.5 mm	4.9°
Average	6.6 mm	9.4°	5.6 mm	8.2°	4.8 mm	8.0°	3.2 mm	6.5°

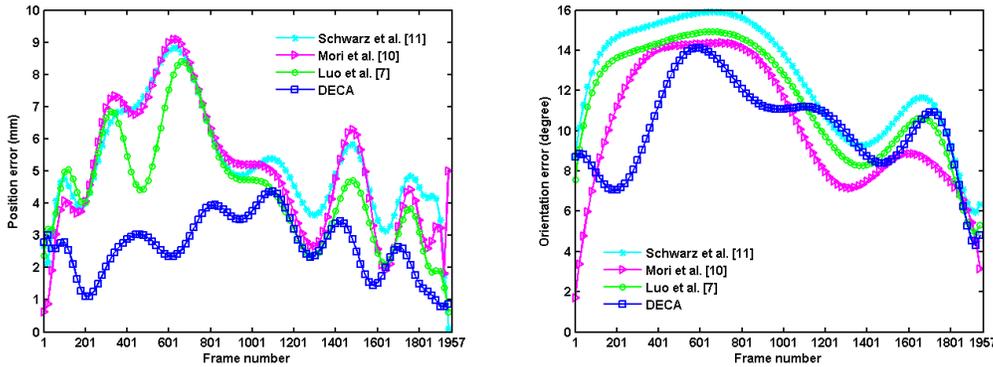


Figure 1: An example (Experiment B) of plotted tracking position and orientation errors of using the four methods.

ground truth frame by frame of Experiment B. Figure 2 illustrates an example of the tracking position and orientation smoothnesses of Experiment D. Table 2 quantifies the smoothness of 3-D motion estimates of using the compared methods. The position smoothness was greatly improved from 3.7 to 2.3 mm, which implies that our proposed DECA involves much less jitter or jump errors in its motion estimates than other methods. Figure 3 displays the visual quality value of estimates using each method. A visual comparison of video and virtual rendering images from the four approaches is shown in Figure 4, which further demonstrates that DECA outperforms other tracking methods.

The purpose of this work is to tackle the particle impoverishment problem in the standard CA. In general, we successfully achieved such a purpose by exploring DECA. From the experimental results of its application to endoscope motion estimation, the tracking accuracy, smoothness, and visual quality of using DECA are significantly better than other methods. We attribute our success in much better tracking to the particle diversity that was significantly improved using the proposed ADE strategy. Figure 5 compares the diversity of CA and DECA with 120 particles at one estimation and demonstrates that DECA has the ability to obtain the much better particle diversity than CA. DECA

distributes particles much more diversely than CA. The majority of particles in CA has the same weight, easily getting trapped in particle impoverishment. Not only the diversity of DECA is much better than CA but also the particle weights in DECA are much larger than CA, i.e., the more diverse of the particles, the more potential solutions could be provided in a multidimensional optimization space; the larger of the particle weight, the more powerful of the particle’s searching ability in solution space.

The potential limitations are clarified as follows. The computation of particle weights is possibly incorrect since it relates to the observation of endoscopic videos that might involve image artifacts or uninformative sequences. Another open problem of DECA is the computational time that can not reach a real-time processing in current experiments.

7. Conclusions

This paper proposed a diversity-enhanced condensation algorithm and its application to endoscope 3-D motion tracking. The proposed approach uses an adaptive differential evolution strategy to successfully address the particle impoverishment problem that frequently happens in the standard condensation algorithm. In the task of endoscope

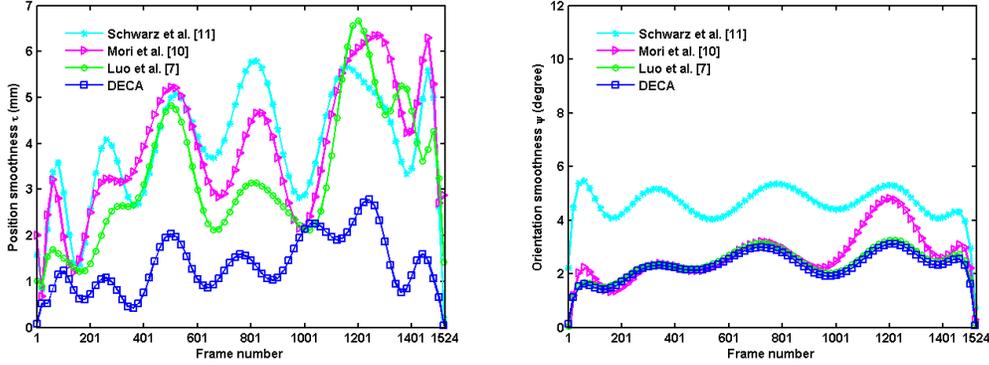


Figure 2: Plotted position and orientation smoothnesses of using the four methods validated on Experiment D.

Table 2: Quantitative comparison of position and orientation smoothnesses, τ , ψ (Eqs. 23, 24), of using the different methods.

Methods Experiments	Schwarz et al. [11]		Mori et al. [10]		Luo et al. [7]		DECA	
	τ	ψ	τ	ψ	τ	ψ	τ	ψ
A	4.4 mm	4.2°	3.3 mm	3.9°	3.2 mm	2.3°	2.8 mm	2.2°
B	4.8 mm	2.6°	4.9 mm	5.4°	4.7 mm	2.7°	1.5 mm	3.3°
C	4.6 mm	6.8°	3.9 mm	5.0°	3.8 mm	3.7°	3.5 mm	3.7°
D	4.8 mm	2.6°	4.5 mm	5.4°	3.8 mm	2.7°	1.4 mm	2.7°
E	5.6 mm	6.2°	3.6 mm	5.1°	3.6 mm	3.3°	3.0 mm	3.3°
F	3.9 mm	2.2°	3.9 mm	4.6°	3.3 mm	2.3°	1.3 mm	2.7°
G	5.1 mm	6.5°	4.0 mm	5.2°	3.5 mm	3.5°	3.0 mm	3.5°
Average	4.7 mm	4.4°	4.0 mm	4.9°	3.7 mm	2.9°	2.3 mm	3.0°

3-D motion tracking, the experimental results demonstrate that the proposed method provides more accurate and robust tracking than previous approaches. The current tracking position error were significantly reduced from 4.8 to 3.2 mm, which approaches the clinical requirement of 3.0 mm.

References

- [1] M. S. Arulampalam and et al. A tutorial on particle filters for nonlinear/non-gaussian Bayesian tracking. *IEEE TSP*, 50(2):174–188, 2002.
- [2] J. Carpenter and et al. Improved particle filter for nonlinear problems. In *IEEE Proc. RSN*, volume 1, pages 2–7, 1999.
- [3] S. Das and et al. Differential evolution: A survey of the state-of-the-art. *IEEE TEVC*, 15(1):4–31, 2011.
- [4] E. Eade and et al. Scalable monocular slam. In *IEEE Proc. CVPR*, volume 1, pages 469–476, 2006.
- [5] G. Grisetti and et al. Improved techniques for grid mapping with rao-blackwellized particle filters. *IEEE Trans. Robotics*, 23(1):34–46, 2007.
- [6] M. Isard and et al. Condensation - conditional density propagation for visual tracking. *IJCV*, 29:5–28, 1998.
- [7] X. Luo and et al. Modified hybrid bronchoscope tracking based on sequential monte carlo sampler: Dynamic phantom validation. In *Proc. ACCV*, volume 6494, pages 409–421, 2011.
- [8] V. Maroulas and et al. Improved particle filters for multi-target tracking. *J. Comp. Phy.*, 231(2):602–611, 2012.
- [9] M. R. Morelande and et al. Maneuvering target tracking in clutter using particle filters. *IEEE TAES*, 41(1):252–270, 2005.
- [10] K. Mori and et al. Hybrid bronchoscope tracking using a magnetic tracking sensor and image registration. In *Proc. MICCAI*, volume 3750, pages 543–550, 2005.
- [11] Y. Schwarz and et al. Real-time electromagnetic navigation bronchoscopy to peripheral lung lesions using overlaid CT images: the first human study. *Chest*, 129(4):988–994, 2006.
- [12] C. Shan and et al. Real-time hand tracking using a mean shift embedded particle filter. *Patt. Rec.*, 40:1958–970, 2007.
- [13] R. Sim and et al. A study of the rao-blackwellised particle filter for efficient and accurate vision-based slam. *IJCV*, 74(3):303–318, 2007.
- [14] R. Storn and et al. Differential evolution - A simple and efficient heuristic for global optimization over continuous spaces. *J. Global Optim.*, 11(4):341–359, 1997.
- [15] X. Wang and et al. An improved particle filter for target tracking in sensor systems. *Sensors*, 7(1):144–156, 2007.
- [16] Z. Wang and et al. A universal image quality index. *IEEE Signal Processing Letters*, 9(3):81–84, 2002.

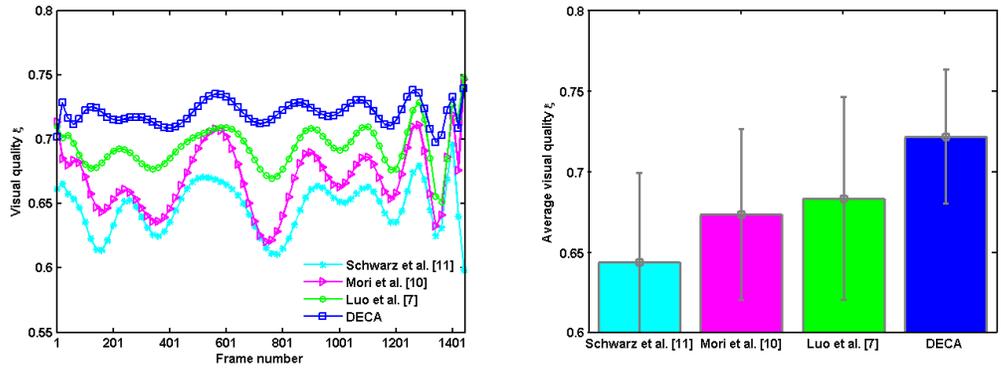


Figure 3: Comparison of the visual quality of using the different methods: An example (Experiment G) of the visual quality (left) and the average visual quality of the total experiments (right).

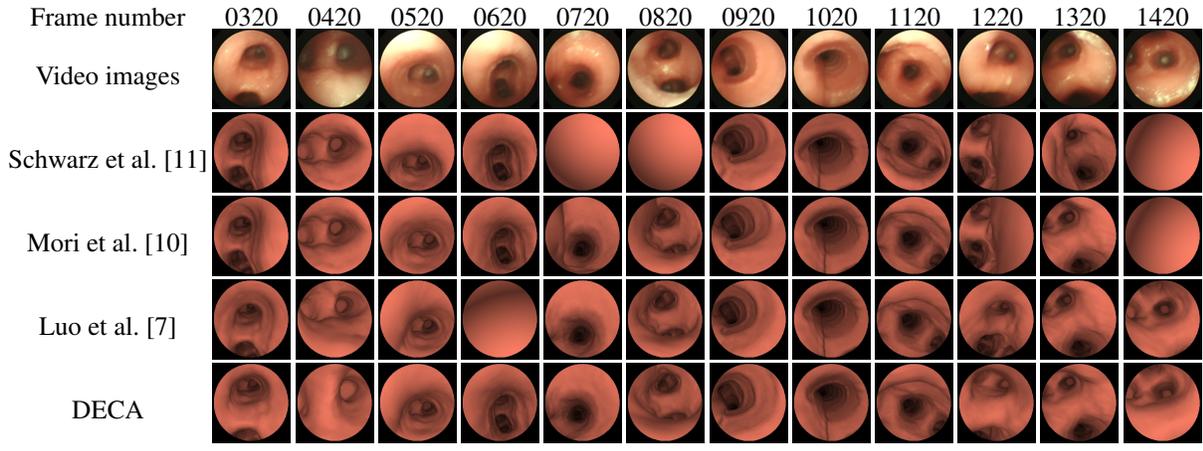


Figure 4: An example (Experiment F) of visually comparing virtual rendering images that were generated from the tracking results of using the four approaches. Top row shows the frame numbers, and second row shows their corresponding real images that were uniformly selected by every 100 frames from endoscope video sequences of Experiment F. Other rows display virtual rendering images. Our proposed DECA framework for endoscope 3-D motion tracking shows the best performance.

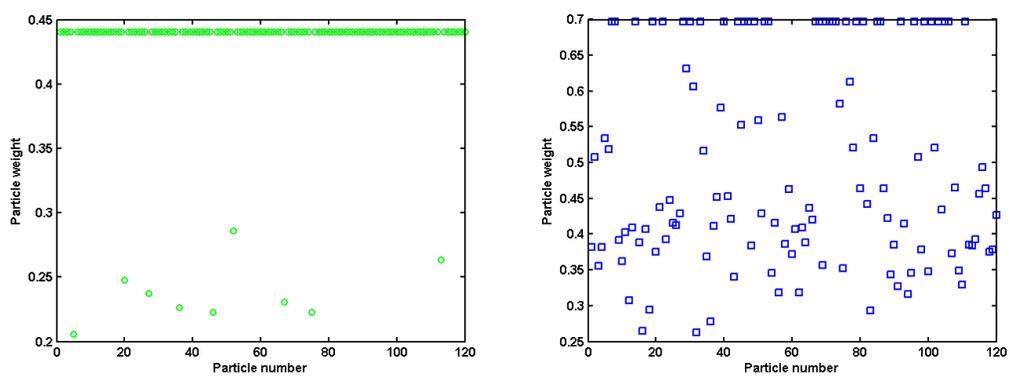


Figure 5: Comparison of the particle diversity of CA (left) and DECA (right). The more different weights that exist, the more potential modes in solution space and less particle impoverishment.