Hybrid Bronchoscope Tracking Using a Magnetic Tracking Sensor and Image Registration*

Kensaku Mori[1], Daisuke Deguchi[1], Kenta Akiyama[1], Takayuki Kitasaka[1], Calvin R. Maurer Jr.[2], Yasuhito Suenaga[1], Hirotsugu Takabatake[3], Masaki Mori[4], and Hiroshi Natori[5]

[1] Graduate School of Information Science, Nagoya University, Nagoya, Japan
{kensaku@is., ddeguchi@suenaga.m.is.}nagoya-u.ac.jp
[2] Department of Neurosurgery, Stanford University, Stanford, U.S.A
[3] Sapporo Minami-Sanjyo Hospital, Sapporo, Japan
[4] Sapporo Kosei Hospital, Sapporo, Japan
[5] School of Medicine, Sapporo Medical University, Sapporo, Japan

Abstract. In this paper, we propose a hybrid method for tracking a bronchoscope that uses a combination of magnetic sensor tracking and image registration. The position of a magnetic sensor placed in the working channel of the bronchoscope is provided by a magnetic tracking system. Because of respiratory motion, the magnetic sensor provides only the approximate position and orientation of the bronchoscope in the coordinate system of a CT image acquired before the examination. The sensor position and orientation is used as the starting point for an intensity-based registration between real bronchoscopic video images and virtual bronchoscopic images generated from the CT image. The output transformation of the image registration process is the position and orientation of the bronchoscope in the CT image. We tested the proposed method using a bronchial phantom model. Virtual breathing motion was generated to simulate respiratory motion. The proposed hybrid method successfully tracked the bronchoscope at a rate of approximately 1 Hz.

1 Introduction

A videobronchoscope consists of a flexible tube and a tiny camera installed at the tip. It is controlled by a physician who watches the camera image through a headpiece or on a video monitor. Bronchoscopy is currently the most commonly employed invasive procedure in the practice of pulmonary medicine and is used for a variety of diagnostic and therapeutic procedures. The physician guides the bronchoscope using only what he sees on the monitor and his knowledge of anatomy. The long-term goal of this research is to develop a bronchoscopic navigation system that provides information such as the planned path of the

* This research was supported in part by a Grant-In-Aid for Scientific Research from the Ministry of Education; the 21st-Century COE program; a Grant-In-Aid for Scientific Research from the Ministry of Education, Culture, Sports, Science, and Technology; the Japan Society for Promotion of Science; and a Grant-In-Aid for Cancer Research from the Ministry of Health and Welfare.

J. Duncan and G. Gerig (Eds.): MICCAI 2005, LNCS 3750, pp. 543–550, 2005.

bronchoscope, the current position of the bronchoscope, and augmented reality visualization of anatomical structures that lie beyond the surface of the bronchial airway. Tracking a bronchoscope is one of the fundamental functions required to build a bronchoscopic navigation system and is the short-term goal of this work.

One approach is image-based tracking in which the position of the bronchoscope is determined by registration of real bronchoscopic (RB) video images and virtual bronchoscopic (VB) images generated from a CT image acquired before the examination. Bricault *et al.* [1] reported the first such work. Their method, which uses the structure of the bronchial tree extracted from the CT image, has difficulty estimating the position of the bronchoscope in areas where no bifurcation appears and considers only the static registration of RB and VB images. Mori *et al.* [2] reported continuous image-based bronchoscope tracking. Good starting points for the image registration process are provided by epipolar geometry analysis of sequential video images. They later improved the tracking process using a sub-block matching algorithm [3].

Image-based tracking generally works very well, but one limitation is that when mistracking occurs in one frame, tracking of subsequent frames is difficult and the method often fails. The two most common causes of failure are quick motion of the bronchoscope and the appearance of bubbles. Another situation that can cause difficulty and failure is when the bronchoscope is both near the bronchial surface and pointing at the surface: the light from the source at the bronchoscope tip can saturate the image. To address this issue and make tracking more robust, we propose a hybrid method for tracking a bronchoscope that uses a combination of magnetic sensor tracking and image-based tracking (image registration). Rigid endoscopes have been tracked with magnetic and optical sensors, and one report uses sensor-based tracking to compare real and virtual endoscopy images [4]. Our proposed method is the first we are aware of to use magnetic sensor tracking for a flexible endoscope, in particular a bronchoscope, and to combine sensor-based tracking with image-based tracking. The position of a magnetic sensor placed in the working channel of the bronchoscope is provided by a magnetic tracking system. Because of respiratory motion, the magnetic sensor provides only the approximate position and orientation of the bronchoscope in the coordinate system of the CT image. The sensor position and orientation is used as the starting point for the registration of RB and VB images. We define coordinate systems and transformations, present the hybrid tracking method, and present preliminary experimental results using a bronchial phantom model with simulated respiratory motion.

2 Background

Bronchoscope tracking: The inputs of the proposed hybrid tracking method are: 1) CT image acquired before the examination, 2) RB video image, and 3) magnetic sensor tracking data. The output is the position and orientation of the bronchoscope, specifically the viewpoint and view direction of the bronchoscope camera model. Tracking consists in generating a sequence of outputs from a sequence of inputs. Synthetic VB images are generated from the CT image using the viewpoint and view direction of the camera. We use volume rendering of the CT image in this work rather than surface rendering of a bronchial surface model extracted from the CT image as in previous work [2]. We denote the k-th frame of a RB video sequence by $\mathbf{B}^{(k)}$ and the position (viewpoint) and orientation (view direction) of the camera by the homogeneous transformation matrix $\mathbf{Q}^{(k)}$ constructed from the rotation matrix $\mathbf{R}^{(k)}$ and the translation vector $\mathbf{t}^{(k)}$.

Magnetic tracking system and sensor: We use the Aurora magnetic tracking system (Northern Digital, Inc., ON, Canada) and a miniature magnetic sensor that is placed in the working channel of the bronchoscope. The measurement volume of the magnetic tracking system is $500 \times 500 \times 500\,\text{mm}$. The sensor has a cylindrical shape with diameter 0.8 mm and length 8 mm. With this sensor, the tracking system provides only five of the six parameters required to fully describe position and orientation in 3D space. The missing parameter is the rotation angle about the longitudinal sensor axis, which means that the sensor does not provide the twist angle of the bronchoscope. Preliminary tests showed that the bronchoscope we use (BF-200, Olympus, Tokyo, Japan) has negligible effect on tracking accuracy when the sensor is placed in the working channel.

Coordinate systems and transformations: Figure 1 illustrates the various coordinate systems and transformations. We represent the position of a point in the bronchoscope camera coordinate system C as $\mathbf{p}^C$. The position of this point in the CT image coordinate system is represented as $\mathbf{p}^{CT}$. We consider the reference frame defined by a sensor attached to the patient (or phantom) as the world coordinate system W and the position of this point is denoted as $\mathbf{p}^W$. The transformation between $\mathbf{p}^C$ and $\mathbf{p}^{CT}$ is given by

$$\begin{aligned}\mathbf{p}^{CT} &= {}^{CT}_{W}\mathbf{M}\,\mathbf{p}^W = {}^{CT}_{W}\mathbf{M}\,{}^{W}_{F}\mathbf{M}^{(k)}\,{}^{F}_{S}\mathbf{M}^{(k)}\,{}^{S}_{E}\mathbf{M}\,{}^{E}_{C}\mathbf{M}\,\mathbf{p}^C \\ &= {}^{CT}_{W}\mathbf{M}\begin{pmatrix} {}^{W}_{S}\mathbf{R}^{(k)} & {}^{W}\mathbf{t}_S^{(k)} \\ \mathbf{0}^T & 1 \end{pmatrix}{}^{S}_{E}\mathbf{M}\,{}^{E}_{C}\mathbf{M}\,\mathbf{p}^C. \end{aligned} \quad (1)$$

The coordinate system at the tip of the sensor in the working channel of the bronchoscope is denoted as E and the transformation between C and E is ${}^{E}_{C}\mathbf{M}$. The coordinate system at the actual sensing point of the sensor is denoted as S and the transformation between E and S is ${}^{S}_{E}\mathbf{M}$. The coordinate system of the magnetic field generator is denoted as F and the transformations between S and F and between F and W for the k-th frame are ${}^{F}_{S}\mathbf{M}^{(k)}$ and ${}^{W}_{F}\mathbf{M}^{(k)}$. The transformation between S and W is ${}^{W}_{S}\mathbf{M}^{(k)} = {}^{W}_{F}\mathbf{M}^{(k)}\,{}^{F}_{S}\mathbf{M}^{(k)}$ and is obtained from the magnetic tracking system; Eq. (1) shows the rotation and translation components ${}^{W}_{S}\mathbf{R}^{(k)}$ and ${}^{W}\mathbf{t}_S^{(k)}$ of this transformation. The transformation ${}^{E}_{C}\mathbf{M}$ describes the relation between the bronchoscope camera and the tip of the sensor.

3 Hybrid Tracking Method

The tracking method consists of several steps: 1) magnetic sensor calibration, 2) image-to-physical registration, 3) estimation of camera pose for the first frame,

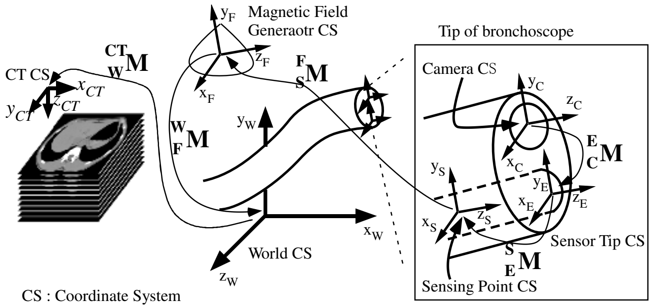

Fig. 1. Coordinate systems and transformations

and 4) updating of camera pose for additional frames. Steps 1 and 2 are performed once before a bronchoscopic examination. Step 3 is performed for the first frame. Continuous tracking is achieved by iterating step 4.

Magnetic sensor calibration: We need to determine the transformation ${}^{S}_{C}\mathbf{M} = {}^{S}_{E}\mathbf{M}\ {}^{E}_{C}\mathbf{M}$. If the sensor provided all six parameters necessary to describe pose in 3D space, the transformation ${}^{S}_{C}\mathbf{M}$ could easily be determined by recording sensor tracking data and RB images of a calibration grid while moving the bronchoscope to various positions and orientations. Since the sensor we use does not provide the rotation angle about the longitudinal sensor axis, we set ${}^{E}_{C}\mathbf{M} = \mathbf{I}$ and construct ${}^{S}_{E}\mathbf{M}$ by measuring the offset between the tip of the sensor and the actual sensing point of the sensor. The error of the identity transformation assumption is corrected during the image registration process.

Image-to-physical registration: We need to estimate the transformation ${}^{CT}_{W}\mathbf{M}$ between the coordinate systems W and CT. This is accomplished by point-based registration between corresponding points on the patient (or phantom) and the CT image. A set of N points $\mathbf{p}_i^{CT}$ $(i = 1, \ldots, N)$ are identified in the CT image. If we had a magnetically tracked probe, we could measure corresponding points on the patient by touching them with the probe and recording their positions. Since we do not have a tracked probe, we use the tip of our five-parameter sensor. The transformation ${}^{CT}_{W}\mathbf{M}$ is found by calculating the rigid transformation that minimizes

$$\sum_{i=1}^{N} \left| \mathbf{p}_i^{CT} - {}^{CT}_{W}\mathbf{M} \left({}^{W}_{F}\mathbf{M}_i\ {}^{F}_{S}\mathbf{M}_i\ {}^{S}_{E}\mathbf{M}\ \mathbf{p}_i^{E} \right) \right|^2, \tag{2}$$

where $\mathbf{p}_i^{E} = (0, 0, 0, 1)^t$ is the tip of the sensor.

Initial estimation of camera pose: Since the sensor does not provide the twist angle about the bronchoscope camera's viewing direction, we estimate the twist angle using a 1D image registration method for the first frame ($k = 0$). We modify Eq. (1) as

$$\mathbf{p}^{CT} = {}^{CT}_{W}\mathbf{M}\ {}^{W}_{F}\mathbf{M}^{(0)}\ {}^{F}_{S}\mathbf{M}^{(0)}\ {}^{S}_{E}\mathbf{M}\ {}^{E}_{C}\mathbf{M} \begin{pmatrix} \mathbf{r}(\theta) & \mathbf{0} \\ \mathbf{0} & 1 \end{pmatrix} \mathbf{p}^{C} = {}^{CT}_{C}\mathbf{M}(\mathbf{r}(\theta))\ \mathbf{p}^{C} \tag{3}$$

where $\mathbf{r}(\theta)$ is a rotation matrix that represents rotation about the z-axis by angle θ. The pose (viewpoint and view direction) of the bronchoscope camera is $\mathbf{Q}^{(0)} = {}^{CT}_{C}\mathbf{M}(\mathbf{r}(\theta))$. The twist angle $\theta_{\max}$ is found as the parameter generating the VB image $\mathbf{V}(\mathbf{Q}^{(0)})$ that is most similar to the RB image $\mathbf{B}^{(0)}$. This 1D image registration process is formulated as:

$$\theta_{\max} = \arg\max_{\theta} \ S\left(\mathbf{B}^{(0)}, \mathbf{V}\left({}^{CT}_{C}\mathbf{M}(\mathbf{r}(\theta))\right)\right), \tag{4}$$

where $S(\mathbf{B}, \mathbf{V})$ denotes the image similarity between $\mathbf{B}$ and $\mathbf{V}$ calculated by the method presented in Ref. [3]. The search is performed using Brent's algorithm [5]. Finally, the initial camera pose is $\mathbf{Q}^{(0)} = {}^{CT}_{C}\mathbf{M}(\mathbf{r}(\theta_{\max}))$.

Continuous tracking of camera pose: The sensor position and orientation is used as the starting point for the registration of RB and VB images. The output transformation of the image registration process is the pose of the bronchoscope camera in the CT image. For the k-th frame, we rewrite Eq. (1) as

$$\begin{aligned} \mathbf{p}^{CT} &= {}^{CT}_{W}\mathbf{M} \begin{pmatrix} {}^{W}_{S}\mathbf{R}^{(0)} & {}^{W}\mathbf{t}^{(k)}_{S} \\ \mathbf{0}^{T} & 1 \end{pmatrix} {}^{S}_{E}\mathbf{M} \, {}^{E}_{C}\mathbf{M} \begin{pmatrix} \mathbf{R}^{(k)} \ldots \mathbf{R}^{(0)} & \mathbf{t}^{(k)} \\ \mathbf{0}^{T} & 1 \end{pmatrix} \mathbf{p}^{C} \\ &= {}^{CT}_{C}\mathbf{M}(\mathbf{R}^{(k)}, \mathbf{t}^{(k)}) \ \mathbf{p}^{C}, \end{aligned} \tag{5}$$

where $\mathbf{R}^{(k)}$ represents rotation between frames $\mathbf{B}^{(k-1)}$ and $\mathbf{B}^{(k)}$, $\mathbf{R}^{(0)} = \mathbf{r}(\theta_{\max})$, and ${}^{W}_{S}\mathbf{R}^{(0)}$ is the rotation component of ${}^{W}_{S}\mathbf{M}^{(0)} = {}^{W}_{F}\mathbf{M}^{(0)} \, {}^{F}_{S}\mathbf{M}^{(0)}$. The updated camera pose is obtained by performing the following image registration process:

$$(\mathbf{R}^{(k)}_{\max}, \mathbf{t}^{(k)}_{\max}) = \arg\max_{\mathbf{R}^{(k)}, \mathbf{t}^{(k)}} \ S\left(\mathbf{B}^{(k)}, \mathbf{V}\left({}^{CT}_{C}\mathbf{M}(\mathbf{R}^{(k)}, \mathbf{t}^{(k)})\right)\right). \tag{6}$$

The search is performed using Powell's algorithm [5]. The camera pose is $\mathbf{Q}^{(k)} = {}^{CT}_{C}\mathbf{M}(\mathbf{R}^{(k)}_{\max}, \mathbf{t}^{(k)}_{\max})$. This process is iterated for frames $k \geq 1$.

4 Experiments and Results

Configurations: We tested the hybrid tracking method using a bronchial phantom model made of rubber. The phantom was fixed in a plastic box. We placed 24 acrylic tubes in the box and used them as fiducial markers. The tips of these tubes are used for estimating the image-to-physical transformation ${}^{CT}_{W}\mathbf{M}$. CT images were acquired using a multi-detector CT scanner. Image acquisition parameters of the CT image are: 512×512 pixels, 341 slices, 0.68 mm pixel size, and 1.25 mm slice thickness. The locations of the tips of the acrylic tubes were manually identified on the CT images. The distance between the magnetic field generator and the phantom was 200 mm. The computer used in these experiments was a PC workstation with dual 3.6 GHz Intel Xeon processors and 2 GB memory running Windows XP operating system. Intrinsic camera parameters of the bronchoscope camera were determined using a calibration grid and standard camera calibration software. Synthetic VB images are generated using a highly optimized software-based volume rendering method. The actual sensing point of the sensor was found to be 3.3 mm from the sensor tip.

Virtual breathing motion (VBM): The bronchoscope and the RB and VB images are in the coordinate system of the lung or the CT image of the lung. The lung of a living person moves during respiration and thus the lung and the bronchoscope move relative to the coordinate system of the magnetic field generator, which is fixed in the physical space of the treatment room. Even if the bronchoscope is stationary with respect to the lung (and the RB image does not change with time), the bronchoscope position will move relative to the magnetic field generator. Since the phantom we employed here does not simulate breathing motion, we added virtual breathing motion (VBM) to the sensor output. Here, we simplify breathing motion as motion in the axial direction of a human body. We assume that breathing motion in the axial direction is zero at the carina of the trachea and is sinusoidal at the diaphragm. We define the VBM as $\Delta\mathbf{p}(t) = (\Delta\mathbf{p}_x(t), \Delta\mathbf{p}_y(t), \Delta\mathbf{p}_z(t))^T$ and add this motion to Eq. (5) as

$$\mathbf{p}^{CT} = \begin{pmatrix} \mathbf{I} & \Delta\mathbf{p}(t) \\ \mathbf{0}^T & 1 \end{pmatrix} {}^{CT}_{W}\mathbf{M} \begin{pmatrix} {}^{W}_{S}\mathbf{R}^{(k)} & {}^{W}\mathbf{t}_S^{(k)} \\ \mathbf{0}^T & 1 \end{pmatrix} {}^{S}_{E}\mathbf{M}\, {}^{E}_{C}\mathbf{M}\, \mathbf{p}^{C}. \quad (7)$$

Since the z-axis of CT images is aligned with the cranial-caudal direction, we add VBM after the transformation between W and CT (${}^{CT}_{W}\mathbf{M}$). The VBM at positions between the carina and diaphragm is interpolated according to fractional axial position. Thus the VBM $\Delta\mathbf{p}(t)$ is defined as $\Delta\mathbf{p}_x(t) = \Delta\mathbf{p}_y(t) = 0$, $\Delta\mathbf{p}_z(t) = A_d \frac{z_c - z_{carina}}{z_d - z_{carina}} \times \sin(2\pi\frac{t}{T})$, where A_d is amplitude of breathing motion at the bottom of the diaphragm, z_c is the z-coordinate of the real camera position before adding VBM, z_{carina} is the z-coordinate of the carina of the trachea, and the z_d is the z-coordinate of the bottom of the diaphragm. These z-coordinates are represented in the CT coordinate system.

Effectiveness of image registration: We measured the changes of estimated RB camera motion caused by VBM. We fixed a bronchoscope at some position and recorded estimated camera positions. As a gold standard of the camera position, we used the estimated position obtained by only image registration. The breathing cycle was set to $T = 6$ s. The mean position errors and their standard deviations are measured. The results are shown in Table 1. Examples of VBM are shown in Fig. 2.

Continuous tracking: We performed bronchoscopic examination on the phantom model using the hybrid tracking method. Camera motion was continuously

Table 1. Position estimation error by adding virtual breathing motion

Amp. of VBM at diaphragm (mm)	Amp. of VBM at camera position (mm)	Avg. error of estimated position (mm)	Standard deviation(mm)
±5.0	±2.4	0.54	0.3
±10.0	±4.8	1.22	0.57
±15.0	±7.2	2.26	1.29

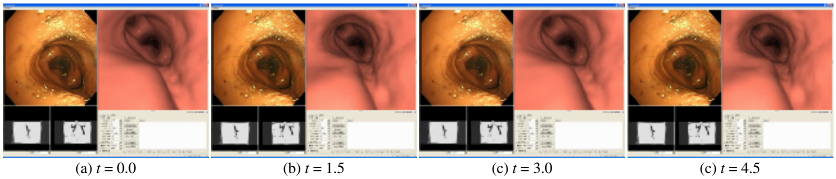

Fig. 2. Effect of virtual breathing motion (VBM). VB views are rendered by using only the sensor's outputs.

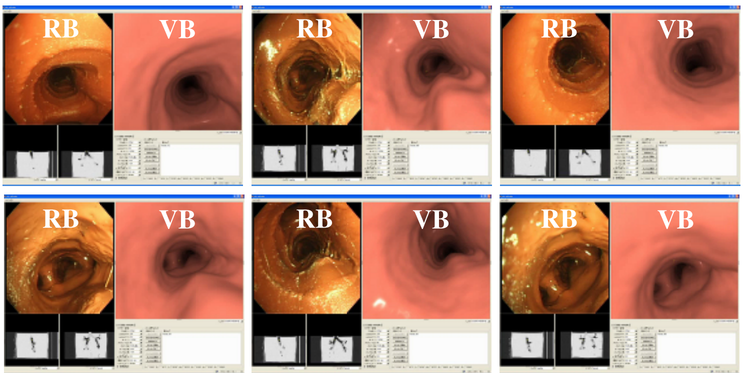

Fig. 3. Tracking results. VB views were rendered by using estimated RB camera position and orientation. VBM was generated during tracking.

tracked and estimated positions were displayed on CT images. VB images were generated using estimated results. In these experiments, VBM was added to the outputs of the magnetic sensor. Figure 3 shows examples of the tracking results. It was possible to update RB camera position and orientation every 1.2 s.

5 Discussion

In the experimental results shown in Table 1, the position estimation errors are much smaller than the VBM; the RB camera positions are satisfactorily estimated by image registration. Figure 3 illustrates that the proposed system successfully estimated RB camera motion in the presence of the VBM shown in Fig. 2. This is the main advantage of hybrid tracking. Although the sensor gives us only a rough estimation, precise estimation is done by image registration. Since the sensor's outputs are used for initial estimation of image registration, it is possible to make the tracking system robust and to reduce computation time. In previous work, e.g., the method in Ref. [2], camera motion is tracked only by image registration. The estimated position of the previous frame is used as the starting point of the iterative search for the current frame. When significant estimation failures occurred, tracking failed in subsequent frames. However, the proposed method in this paper continues to track after such failures. The proposed system tracked the RB camera at 0.8 frames per second. Because of the limited processing speed, the system failed in tracking when an operator inserted or pulled out a bronchoscope quickly. However, in such situations, when the operator kept the bronchoscope stationary at a bifurcation point, the system recovered to the correct estimated position.

We assumed that virtual breathing motion is motion in the axial direction. However, actual breathing motion is much more complicated. Further experiments are required to validate the proposed method. It would be useful to use a bronchial phantom that includes physical breathing motion. Because the Aurora sensor outputs only five pose parameters, we estimated only translation terms of ${}^{S}_{E}\mathbf{M}$ and assumed that the tip of the sensor is aligned with the optical center and direction of the bronchoscope camera (${}^{E}_{C}\mathbf{M} = \mathbf{I}$). The image registration process compensated for the error inherent in these assumptions and produced good results in actual tracking.

The work described in this paper is another step towards developing a fast, accurate, and robust method for tracking a bronchoscope. The results are preliminary but promising. Future work includes: 1) incorporating a magnetic tracking system and sensor that fits in the working channel of a bronchoscope and provides all six parameters necessary to describe pose in 3D space (e.g., Ascension microBird magnetic tracking system), 2) development of a better sensor calibration method, 3) development of better methods for estimating and correcting for breathing motion, 4) improvement of processing speed, and 5) testing with a bronchial phantom model that includes physical breathing motion and with data collected from patients.

References

1. I Bricault, G Ferretti, P Cinquin. Registration of real and CT-derived virtual bronchoscopic images to assist transbronchial biopsy. *IEEE Trans Med Imaging*, 17: 703–714, 1998.
2. K Mori, D Deguchi, J Sugiyama, et al. Tracking of a bronchoscope using epipolar geometry analysis and intensity-based image registration of real and virtual endoscopic images. *Med Image Anal*, 6: 321–336, 2002.
3. D Deguchi, K Mori, Y Suenaga, et al. New image similarity measure for bronchoscope tracking based on image registration. *Proc MICCAI 2003*, LNCS 2878: 309–406, 2003.
4. R Shahidi, MR Bax, CR Maurer Jr, et al. Implementation, calibration and accuracy testing of an image-enhanced endoscopy system. *IEEE Trans Med Imaging*, 21: 1524–1535, 2002.
5. WH Press, SA Teukolsky, WT Vetterling, BP Flannery. *Numerical Recipes in C: The Art of Scientific Computing*, 2nd ed. Cambridge University Press, 1992.