# Robust Real-Time Image-Guided Endoscopy: A New Discriminative Structural Similarity Measure for Video to Volume Registration

Xiongbiao Luo[1], Hirotsugu Takabatake[2], Hiroshi Natori[3], and Kensaku Mori[1]

[1] Information and Communications Headquarters, Nagoya University, Japan
[2] Sapporo Minamisanjo Hospital
[3] Keiwakai Nishioka Hospital, Japan
xiongbiao.luo@gmail.com, kensaku@is.nagoya-u.ac.jp

**Abstract.** This paper proposes a fully automatic real-time robust image-guided endoscopy method that uses a new discriminative structural similarity measure for pre- and intra-operative registration. Current approaches are limited to clinical applications due to two major bottlenecks: (1) weak continuity, i.e., endoscopic guidance may be blocked since a similarity measure might incorrectly characterize video images and virtual renderings generated from pre-operative volume data, resulting in a registration failure; (2) slow computation, since volume rendering is a time-consuming step in the registration. To address the first drawback, we introduce a robust similarity measure, which uses the degradation of structural information and considers image correlation or structure, luminance, and contrast to characterize images. Moreover, we utilize graphics processing unit techniques to accelerate the volume rendering step. We evaluated our method on patient datasets. The experimental results demonstrated that we provide a promising method, which is possibly applied in the operating room, to accurately and robustly guide endoscopy in real time, particularly the average accuracy of position and orientation was improved from (14.6, 51.2) to (4.45 mm, 12.3°) and the runtime was about 32 frames per second compared to current image-guided methods.

**Keywords:** Image-Guidance Endoscopy, Endoscope Tracking and Navigation, Video-Volume Registration, Discriminative Structural Similarity.

## 1 Endoscopic Interventions

Endoscopic interventions are widely performed for cancer diagnosis, e.g., bronchoscopy and endoscopic sinus surgery. Such interventions use endoscopes to insert into the body through natural orifices and observe suspicious regions where biopsies may be performed. However, these interventions in the hands of different skilled endoscopists are the most sensitive procedure for locating tumors since endoscopic video cameras only provide two-dimensional (2-D) image information, which is not enough to determine six-degree-of-freedom (6DoF) position and orientation of an endoscope in a three-dimensional (3-D) space. Moreover, timing of endoscopy depends on physicians' skills; the more time of endoscopy

D. Barratt et al. (Eds.): IPCAI 2013, LNCS 7915, pp. 91–100, 2013.

being operated, the more high risk the patients have. An image-guided endoscopy is promising to address the problems of location and timing of endoscopy.

Image-guided endoscopy registers 2-D video images to 3-D pre-operative data, e.g., computed tomography (CT) or magnetic resonance (MR) images that are usually acquired before interventions, to navigate or locate the endoscope in a reference coordinate system in real time. It usually defines a similarity measure to compute image intensity difference between video and virtual rendering images and runs an optimizer to find the optimal corresponding virtual image [1,2,3]. Compared to commercially available electromagnetically navigated endoscopy [4,5], it has several interesting advantages including cost-efficient, without additional setups, little influence from respiratory motion, and without inherent system or dynamic errors. Unfortunately, two main weaknesses limit image-guided endoscopy to apply in operation rooms: (1) guidance discontinuity and (2) large amount of calculation. The former is caused by problematic endoscopic images (e.g., local luminance and contrast changes) that may easily collapse the registration since the similarity measure may not adapt itself to these changes. The latter results from volume rendering to generate virtual images, blocking a real-time guidance procedure where at lest 30 frames are processed in a second. Even though many papers have been published in the literature [1,3], more accurate and effective methods to tackle these weaknesses are still expected for the robust real-time image-guided endoscopy.

This work realized a robust real-time image-guided endoscopy. To accurately register 2-D video images and 3-D CT volume, we proposed a new discriminative structural similarity (DSSIM) measure. The similarity function is a key element that is expected to precisely characterize intensity difference under a dynamic environment. DSSIM can adapt itself successfully to image changes due to nonlinear illumination, specular- or inter-reflection, or collision with the organ walls in endoscopy. Moreover, since generating 2-D virtual images is time-consuming, we use graphics processing unit (GPU) techniques to accelerate our method up to 32 frames per second (fps), which meets the real-time requirement ($\geq$ 30 fps).

Several highlights of this work are summarized as follows. First, we modified a measure of structural similarity (SSIM) to DSSIM that is robust and accurate for a video-volume registration. We extended a new application of SSIM in computer assisted interventions. Furthermore, to best of our knowledge, no methods were published as real-time image-guided endoscopy using image registration methods before. We reported a fully automatic image-guided endoscopy in real time. Additionally, our method is suitable to other endoscopies (e.g., conchoscope).

## 2 Proposed Approaches

Our proposed approach to guide endoscopic interventions and determine endoscope 6DoF location information comprises of several main steps: (1) automatically initializing the guided procedure, (2) formulating the discriminative structural similarity measure, and (3) performing video-volume registration for continuous endoscope guidance. Fig. 1 shows the flowchart of our proposed method.

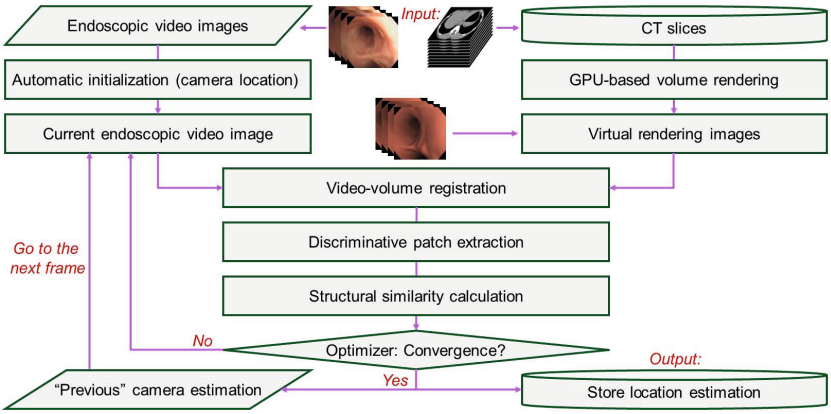



**Fig. 1.** The processing flowchart of our proposed method for endoscope guidance

### 2.1 Automatic Initialization

Endoscopic guidance must be initialized before continuous navigation. It is hard to perform a manual initialization that takes much time during examination. It is also somewhat difficult to use fiducials to align from patient to CT spaces. For surgical requirements, we here introduce a fully automatic initialization method on the basis of airway tree structures and manifold learning.

First, we segment CT images to obtain the centerlines of the trachea, the left main bronchus, and the right main bronchus with their start and end positions, $(\mathbf{s}_t, \mathbf{e}_t)$, $(\mathbf{s}_l, \mathbf{e}_l)$, and $(\mathbf{s}_r, \mathbf{e}_r)$, before an endoscopic intervention. The carina position should be either $\mathbf{e}_t$ or $\mathbf{s}_l$ or $\mathbf{s}_r$.

Next, we generate a set of virtual images by updating position $\mathbf{p}_i$ and orientation $\mathbf{o}_i(\mathbf{o}_i^x, \mathbf{o}_i^y, \mathbf{o}_i^z)$ of a virtual camera in the CT space ($\alpha \in [0.5\ 0.9]$):

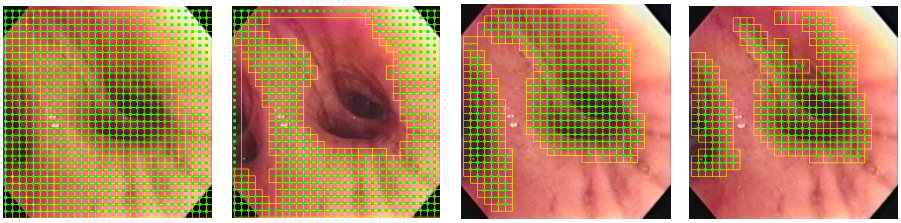
$$\mathbf{p}_i = \mathbf{s}_t + \frac{\alpha(\mathbf{e}_t - \mathbf{s}_t)}{\|\mathbf{e}_t - \mathbf{s}_t\|},\ \mathbf{o}_i^z = \frac{(\mathbf{e}_t - \mathbf{s}_t)}{\|\mathbf{e}_t - \mathbf{s}_t\|},\ \mathbf{o}_i^y = \frac{(\mathbf{e}_l - \mathbf{s}_l)}{\|\mathbf{e}_l - \mathbf{s}_l\|} \times \frac{(\mathbf{e}_r - \mathbf{s}_r)}{\|\mathbf{e}_r - \mathbf{s}_r\|}, \quad (1)$$

where, $\mathbf{o}_i^x = \mathbf{o}_i^z \times \mathbf{o}_i^y, \mathbf{o}_i^y$, and $\mathbf{o}_i^z$ are the direction vectors of the virtual camera.

Finally, we use a manifold learning method to construct the subspace for those generated virtual images with different camera poses (position and orientation parameters) [6]. During the intervention, the physician can initially locate the endoscope around the carina of the airways and embedded the current video image to the subspace and find the optimal initialization to start a navigation.

### 2.2 Discriminative Structural Similarity

The similarity measure is a core of image registration. It is supposed to accurately and robustly represent image changes (distortion), e.g., illumination and motion blurring. We propose a discriminative structural similarity measure that



(a) (b) (c) (d)

**Fig. 2.** Discriminative region extraction (a *yellow square* indicates one patch and a *green point* is one patch center): (a) all separated patches from an input image, (b) removed patches without structural information, (c) remained patches with structural information, (d) finally used patches during similarity computation.

takes incomplete correlation, luminance and contrast distortion into consideration to model image changes. *Discriminative* here means specific structures such as bifurcations and folds inside the airways. Since the structural information is very useful for the similarity calculation, we first extract discriminative regions.

**Discriminative Region Extraction.** For an image with $W \times H$ pixels, we divide it into $U \times V$ patches. One patch $P_{u,v}$ with $\frac{W}{U} \times \frac{H}{V}$ pixels is presented by:

$$P_{u,v} = \{(c_x, c_y), u \in U, v \in V\}, \tag{2}$$

where $c_x$ and $c_y$ are the patch center coordinates. We define two variables: intensity variance $\sigma_{u,v}$ and contrast $\omega_{u,v}$ that indicates the tone of the highlights and lighter areas, to check whether $P_{u,v}$ includes the structural information:

$$\sigma_{u,v}^2 = \frac{1}{|P_{u,v}|} \sum_{P_{u,v}} \left(P_{u,v}(x,y) - \bar{P}_{u,v}\right)^2, \ \omega_{u,v} = \frac{1}{|P_{u,v}|} \sum_{P_{u,v}} \Psi\left(P_{u,v}(x,y)\right), \tag{3}$$

where $(x, y)$, $|P_{u,v}|$, and $\bar{P}_{u,v}$ denote one pixel coordinates, the pixel number, and the average intensity in patch $P_{u,v}$, respectively. Function $\Psi\left(P_{u,v}(x,y)\right)$, which depends on the pixel color information of saturation $S(x,y)$ and lightness $L(x,y)$ in the hue-saturation-lightness (HSL) color model, is defined to evaluate whether pixel $(x, y)$ belongs to the highlights and lighter areas or not:

$$\Psi\left(P_{u,v}(x,y)\right) = \begin{cases} 1 & S(x,y) \leq \beta_S \; and \; L(x,y) \geq \delta_L \\ 0 & otherwise \end{cases}, \tag{4}$$

where $\beta_S$ and $\delta_L$ are two pre-determined thresholds. We remove the white patches without structural information by $\omega_{u,v} \geq \epsilon_\omega$ (a fixed constant), descendingly sort the remained patches in terms of $\sigma_{u,v}$, and choose $\lambda \cdot U \cdot V$ patches for the similarity calculation. Fig. 2 shows the discriminative patch detection.

**Structural Similarity Function.** A similarity function seeks to correctly depict pixel difference between distorted and reference images in the registration. Image distortion usually results from structure (correlation), luminance, and

contrast changes. Based on the work of SSIM [7], we introduce the similarity function $M$ into the volume-video registration for guided interventions as:

$$M = \underbrace{\frac{\sigma_{d,r} + C_1}{\sigma_d \sigma_r + C_1}}_{Structure} \cdot \underbrace{\frac{2\xi_d \xi_r + C_2}{\xi_d^2 + \xi_r^2 + C_2}}_{Luminance} \cdot \underbrace{\frac{2\sigma_d \sigma_r + C_3}{\sigma_d^2 + \sigma_r^2 + C_3}}_{Contrast}, \tag{5}$$

where $\sigma_{d,r}$ is the correlation between distorted and reference images; $\xi_d$ and $\xi_r$ are the intensity mean; $\sigma_d$ and $\sigma_r$ are the intensity variance, respectively (constants: $C_1$, $C_2$, and $C_3$). Three elements in Eq. 5 were demonstrated to successfully characterize image changes [7]. By $C_3 = 2C_1$, we rewrote Eq. 5 as:

$$M = \frac{(2\sigma_{d,r} + C_1)(2\xi_d \xi_r + C_2)}{(\sigma_d^2 + \sigma_r^2 + C_1)(\xi_d^2 + \xi_r^2 + C_2)}. \tag{6}$$

After choosing $\lambda \cdot U \cdot V$ discriminative regions, similarity $DSSIM(I_k, I_{CT})$ between $k$-th video sequence $I_k$ and CT-based virtual image $I_{CT}$ is computed by:

$$DSSIM(I_k, I_{CT}) = \frac{1}{\lambda \cdot U \cdot V} \sum_{P_{u,v} \in \lambda \cdot U \cdot V} \frac{1}{|P_{u,v}|} \sum_{P_{u,v}} \hat{M}_{u,v}, \tag{7}$$

$$\hat{M}_{u,v} = \frac{\left(2\sigma_{k,CT}^{u,v} + C_1\right)(2\xi_k^{u,v}\xi_{CT}^{u,v} + C_2)}{((\sigma_k^{u,v})^2 + (\sigma_{CT}^{u,v})^2 + C_1)((\xi_k^{u,v})^2 + (\xi_{CT}^{u,v})^2 + C_2)}. \tag{8}$$

The DSSIM measure will be demonstrated to very robust and accurate for registering video and CT-based virtual images from our experimental results.

**Remarks on the DSSIM Measure.** Image structural or discriminative information is very useful for the similarity calculation since it describes the pixel dependency that involves significant information about visual structures. Hence, a robust similarity measure should be able to characterize visual structural information in images. Moreover, image similarity should be computed locally but not globally, i.e., an image should be divided into many patches and the similarity of each patch is calculated and added up to the finial similarity. The similarity's locality is better than its globality since it yields several practical situations, e.g., dynamic of image statistical features, image distortion being independent or dependent of local characteristics, the human vision system being sensitive to local structures, and a variable image quality map related to local quality measurement. Additionally, a good measure should be insensitive to luminance and contrast changes. DSSIM can meet three requirements of a good similarity measure: (1) usage of structural information (2) locality, and (3) adaptation of luminance or contrast distortion. We extract discriminative structures (bifurcations or folds) in local regions and compute the local similarity of the patches whose luminance or contrast distortion was modeled.

### 2.3 Video-Volume Registration

For a continuous endoscopic navigation, we must perform the video-volume registration ($V^2R$) to determine the spatial transformation between the video and

CT volume coordinate systems during the image-guided endoscopic intervention. Such a spatial transformation involves with the 6DoF parameters of position and orientation of the endoscope located somewhere in the airways.

Suppose that ${}^{CT}\mathbf{T}_V^k$ with position ${}^{CT}\mathbf{t}_V$ and rotation matrix ${}^{CT}\mathbf{R}_V$ is the transformation matrix from video to volume at frame $k$. To estimate ${}^{CT}\mathbf{T}_V^{k+1}$, we formulate V$^2$R as an optimization process on the basis of the proposed DSSIM measure and determine the changeable transformation parameter $\Delta^{CT}\mathbf{T}_V^{k+1}$ by:

$$\Delta^{CT}\mathbf{T}_V^{k+1} = \arg max_{\Delta^{CT}\mathbf{T}_V^{k+1}} DSSIM\left(I_k, I_{CT}({}^{CT}\mathbf{T}_V^k \cdot \Delta^{CT}\mathbf{T}_V^{k+1})\right), \quad (9)$$

where virtual image $I_{CT}(\cdot)$ is generated on the basis of virtual camera parameters ${}^{CT}\mathbf{T}_V^k \cdot \Delta^{CT}\mathbf{T}_V^{k+1}$. By running an optimizer, we find optimal $\Delta^{CT}\breve{\mathbf{T}}_V^{k+1}$ to maximize the similarity between images $I_{k+1}$ and $I_{CT}({}^{CT}\mathbf{T}_V^k \cdot \Delta^{CT}\breve{\mathbf{T}}_V^{k+1})$.

Note that the initialization of $\Delta^{CT}\mathbf{T}_V^{k+1}$ is important to the optimizer, as discussed in [3]. It can be initialized as an identity matrix. Such an initialization will lose the temporal coherence between two consecutive video frames, possibly resulting in a guidance failure. Video image textures or features can be used to compensate such losing. However, such a compensation takes much time. In this work, we determine the initialization empirically. We clarify that typical translating and rotating speeds of an endoscope is 10.0 mm and 20 degrees per second. An endoscopic camera is usually at frame rate of 30 fps. Therefore, interframe speeds $\tau$ and $\phi$ of translation and rotation are about 0.33 mm and 0.66 degrees per frame ($\tau = 0.33$ mm and $\phi = 0.66$ degrees). Hence, we can initialize $\Delta^{CT}\mathbf{T}_V^{k+1}$ by the following equations:

$$\Delta^{CT}\mathbf{T}_V^{k+1} = \begin{pmatrix} \Delta^{CT}\mathbf{R}_V^{k+1} & \Delta^{CT}\mathbf{t}_V^{k+1} \\ \mathbf{0}^T & 1 \end{pmatrix}_{4\times 4}, \quad (10)$$

$$\Delta^{CT}\mathbf{t}_V^{k+1} = [\tau\ \tau\ \tau]^T,\ \Delta^{CT}\mathbf{R}_V^{k+1} = \begin{pmatrix} b^2 & a^2b - ab & ab^2 + a^2 \\ ab & a^3 + b^2 & a^2b - ab \\ -a & ab & b^2 \end{pmatrix}_{3\times 3}, \quad (11)$$

where the variables of matrix $\Delta^{CT}\mathbf{R}_V^{k+1}$ are defined as: $a = \sin\phi$ and $b = \cos\phi$.

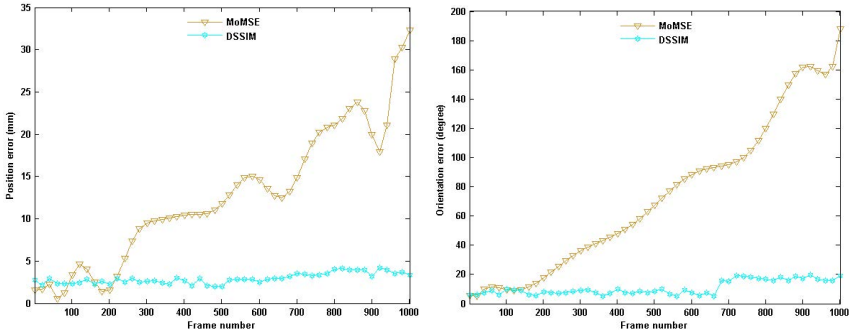
## 3 Experimental Settings

We validated our proposed method on six cases of patient datasets: (1) endoscopic video images, whose sizes were 360 × 370 and 256 × 263 pixels, were recorded at a frame rate of 30 fps, and (2) CT volumes were acquired by space parameters of 512 × 512 pixels, 72-351 slices, 2.0-5.0-mm slice thickness.

We implemented our method on a Dell Precision Workstation that was equipped with Intel (R) Xeon(R) CPU X5355 2.66 GHz × 8, NVIDIA GeForce 8800 GTX, and 16.0 GB memory and installed with the Windows 7 64-bit operating system and the NVIDIA CUDA 4.2 toolkit. We investigate two image-based

**Table 1.** Quantitative results of the guidance accuracy of the two methods in terms of position and orientation errors between the estimates and ground truth

| Patient data (Frames) | Comparison of (position, orientation) of the two methods | |
|---|---|---|
| | MoMSE | DSSIM |
| Case A (379) | (31.2±25.8 mm, 38.8±29.3°) | (9.08±6.88 mm, 12.4±8.00°) |
| Case B (1000) | (12.4±7.84 mm, 72.8±52.3°) | (2.88±1.62 mm, 10.8±6.53°) |
| Case C (449) | (4.75±2.99 mm, 10.0±5.80°) | (4.35±2.77 mm, 9.29±4.50°) |
| Case D (2650) | (10.4±5.70 mm, 66.6±35.4°) | (2.32±1.81 mm, 8.67±7.21°) |
| Case E (450) | (13.8±11.7 mm, 23.9±18.6°) | (4.64±2.75 mm, 17.7±14.7°) |
| Case F (2000) | (15.3±14.3 mm, 45.6±28.5°) | (3.42±3.07 mm, 14.2±12.3°) |
| **Average** | **(14.6±11.4 mm, 51.2±28.3°)** | **(4.45±3.15 mm, 12.3±8.88°)** |



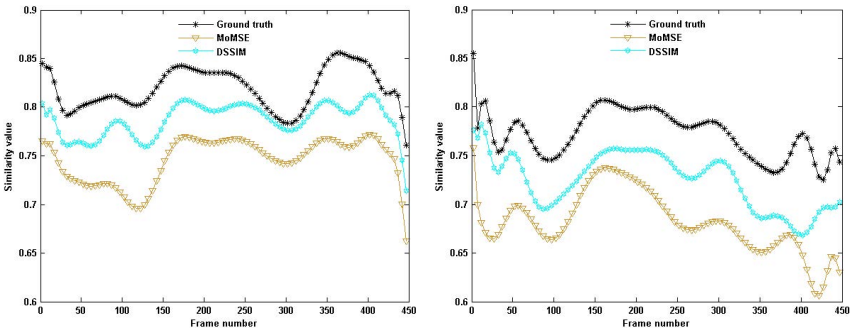

(a) Position error (b) Orientation error

**Fig. 3.** Navigation position and orientation errors of the two methods on Case B was plotted against ground truth by every 20 frames

methods: (1) MoMSE: a method using a modified mean square error similarity measure [1], (2) DSSIM: our method, as discussed in Section 2. To evaluate the guidance accuracy, we generate ground truth data by manually adjusting the position and orientation of the virtual camera to qualitatively align video and CT-driven virtual images. Additionally, we set parameters: $U = V = 30$, $\lambda = 0.3$, $\beta_S = 0.6$, $\delta_L = 0.7$, and $\epsilon_\omega = 0.9$ during discriminative region extraction.

## 4 Results

Table 1 lists the guidance accuracy by computing the position and orientation errors between ground truth and the estimates. The mean position and orientation errors of our approach were 4.45 mm and 12.3°, which are significantly better than 14.6 mm and 51.2° of the MoMSE-based method. Fig. 3 plots the guidance accuracy of the MoMSE- and DSSIM-based methods on Case B. Fig. 5 shows some video images of Case D and their corresponding virtual images generated from the estimated results. Fig. 4 compares the similarity between video and virtual images, demonstrating that the visualization quality of the DSSIM-based method is absolutely better than the MoMSE-based method (Fig. 5).




(a) Case C (b) Case E

**Fig. 4.** Comparison of the similarity value of the two methods.

**Table 2.** Comparison of iterations and computation time of volume rendering, similarity, and one frame with and without CUDA speed-up (ms: milliseconds)

| Computation comparison | Without CUDA | | With CUDA | |
|---|---|---|---|---|
| | MoMSE | DSSIM | MoMSE | DSSIM |
| Iterations | 77 | 52 | 67 | 49 |
| Rendering | 138 ms | 104 ms | 22 ms | 15 ms |
| Similarity | 38 ms | 68 ms | 6 ms | 10 ms |
| One frame | 246 ms | 219 ms | 38 ms | 31 ms |

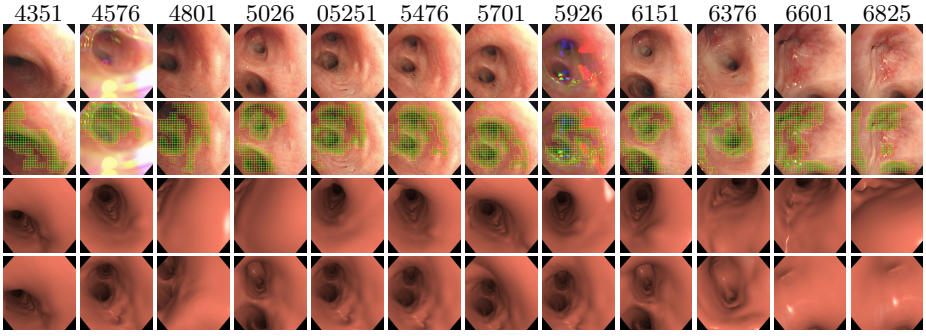
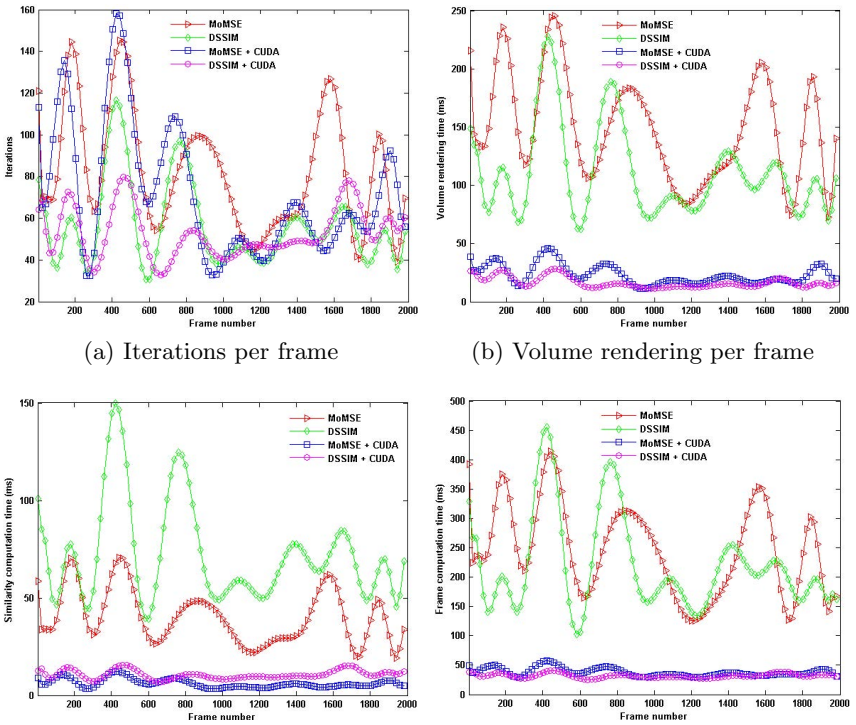



**Fig. 5.** Visual comparison of guidance results of Case D. Top row shows uniformly selected frame numbers, and second row shows their corresponding video images. Third row gives the results of discriminative region extraction. Fourth and fifth rows display virtual images based on the estimates from the MoMSE- and DSSIM-based methods, respectively. Our method shows better performance.

More interestingly, our approach can be implemented in real time using GPU techniques. After accelerating by GPU, the DSSIM-based approach needs about 31 milliseconds per frame (mpf), i.e., processing about 32 fps, which exceeds the




(a) Iterations per frame (b) Volume rendering per frame

(c) Similarity computation per frame (d) Runtime per frame

**Fig. 6.** Comparison of the computational times of the two methods on Case F

clinical requirement of 30 fps. The MoMSE-based method can process about 26 fps (38 mpf), slightly being lower than the real-time need (Table 2 and Fig. 6).

## 5 Discussion and Conclusion

We realized a real-time endoscope guidance with a more robust and accurate navigation. We believe that the effectiveness lies in the DSSIM's robustness. Sine the visualization quality of guidance results (i.e., virtual images generated from endoscope location parameters) depends on the human visual system (HVS) that is very sensitive to structural information in images, a good similarity measure should approximate structural information changes as accurate as possible. MoMSE computes pixel difference to approximate image distortion but hardly fits to HVS. DSSIM, which use structural information changes to characterize image distortion, follows HVS well. Moreover, DSSIM can adapt itself to luminance and contrast dynamics, as proved in our experimental results. Additionally, the runtime, which was improved to the real-time level, is mainly attributed to GPU techniques. We believe that the similarity measure that makes convergence fast can also reduce the runtime (Fig. 6). Even though DSSIM is computed by more time than MoMSE, its robustness makes iterations reduced in optimization.

Our method has one main potential limitation that is difficult to tackle problematic video images (e.g., bubbles), which possibly fail a continuous endoscope guidance. Future work includes recovering the continuous guidance by removing these ambiguous images. We also plan to revoke a re-initialization mechanism to tackle failure since an endoscope is usually operated back to where it has flied through. Additionally, since we current used a relatively simple processing method in discriminative region detection, we seek to use more robust functions to perform the patching and calculate the inter-pixel similarity among images.

To summarize our work, this article proposes a framework of a fully automatic, robust, and real-time image-guided endoscopy by a video-volume registration on the basis of a discriminative structural similarity measure and GPU acceleration techniques, without additional positional sensors (e.g., electromagnetic sensors). Current guidance accuracy and processing time were significantly improved up to position error 4.45 mm, orientation error 12.3°, and 32 fps.

**Acknowledgment.** This work was partly supported by the project "Development of Bedside Medical Devices for High Precision Diagnosis of Cancer in Its Preliminary Stage" (01-D-D0806) funded by the Aichi Prefecture, and the program "Development of Scale Seamless Endoscopy Navigation System for Diagnostic Surgery" funded by the Japan Society for the Promotion of Science, and the project "Computational Anatomy for Computer-aided Diagnosis and Therapy: Frontiers of Medical Image Sciences" (21103006) funded by Grant-in-Aid for Scientific Research on Innovative Areas, MEXT, Japan.

## References

1. Deguchi, D., et al.: Selective image similarity measure for bronchoscope tracking based on image registration. MedIA 13(4), 621–633 (2009)
2. Mirota, D.J., et al.: A system for video-based navigation for endoscopic endonasal skull base surgery. IEEE TMI 31(4), 963–976 (2012)
3. Luo, X., et al.: Development and comparison of new hybrid motion tracking for bronchoscopic navigation. MedIA 16(3), 577–596 (2012)
4. Schwarz, Y., et al.: Real-time electromagnetic navigation bronchoscopy to peripheral lung lesions using overlaid CT images: The first human study. Chest 129(4), 988–994
5. Luó, X., Reichl, T., Feuerstein, M., Kitasaka, T., Mori, K.: Modified hybrid bronchoscope tracking based on sequential Monte Carlo sampler: dynamic phantom validation. In: Kimmel, R., Klette, R., Sugimoto, A. (eds.) ACCV 2010, Part III. LNCS, vol. 6494, pp. 409–421. Springer, Heidelberg (2011)
6. Luo, X., Kitasaka, T., Mori, K.: ManiSMC: A new method using manifold modeling and sequential Monte Carlo sampler for boosting navigated bronchoscopy. In: Fichtinger, G., Martel, A., Peters, T. (eds.) MICCAI 2011, Part III. LNCS, vol. 6893, pp. 248–255. Springer, Heidelberg (2011)
7. Wang, Z., et al.: Image quality assessment: From error visibility to structural similarity. IEEE TIP 13(4), 600–612 (2004)