Multi-Scale Optimization Strategy in Image Registration Based Navigation System for Bronchofiberoscopic Procedures

Paweł Turcza

Department of Instrumentation and Measurement, AGH University of Science and Technology
Al. Mickiewicza 30, 30-059 Kraków, Poland, email: turcza@agh.edu.pl

Abstract- Image processing based navigation systems establish the position of bronchofiberoscopy tip by registering images from endoscope camera to images obtained from virtual camera observing virtual objects generated by the computer from CT data. The paper propose application of multi-scale optimization strategy in image registration algorithm in such a navigation system. Proposed strategy intends to enable real-time operation of navigation system by reduction of computational complexity, increase of robustness to local minima and improvement registration process convergence. Some initial verification results are presented.

I. Introduction

One of the most important bronchofiberoscopy procedure is transbronchial needle aspiration (TBNA). It allows for sampling neoplasmatic tissue and lymph nodes for staging of lung cancer. Because lymph nodes are situated outside bronchial tree wall they are invisible during bronchofiberoscopy procedure. The most suitable points for performation of needle aspiration may be shown in virtual bronchoscopy view. Furthermore, the location of the tip of the real endoscope may be established through the registration of two images: the real image from endoscope camera and the virtual bronchoscopy image, obtained in real-time, from computer tomography (CT) data. A promising complex navigation system, exploiting virtual bronchoscopy presentation and image registration has been developed by Helferty & Higgins [1,2].

In this paper we propose combining recently proposed, faster two-stage optimization strategy of image registration [4], with more flexible multi-scale image matching. Application of multi-scale optimization technique to virtual endoscopy navigation system results in computational complexity reduction and faster image matching. Additionally it offers the following advantages:

- reduction of sensitivity to noise,
- avoiding local minima during optimization,
- increasing range of registration parameters,
- possibility of changing degree of freedoms on different image decomposition levels.

Initial verification of the proposed solution is presented in the paper.

II. Bronchofiberoscopy navigation system

The block diagram of discussed navigation system for bronchofiberoscopy is presented in the Figure 1. The system consists of: 1) the video frame grabber connected to endoscopic camera, 2) computer system generating images from CT data, 3) the optimizer which is responsible for determining virtual camera position and parameters for virtual image transformation in such a way so as to makes images from virtual camera more similar to real endoscopic images according to a chosen measure *M*,

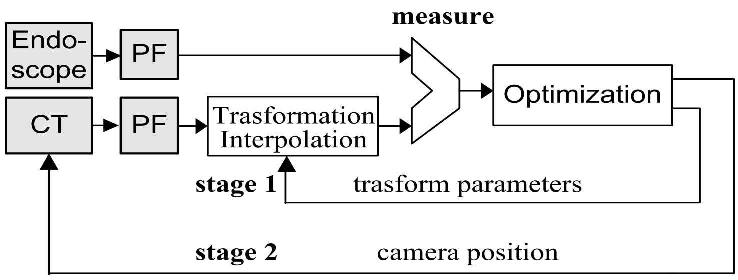

Figure 1. The block diagram of navigation system

4) module for image transformation (scaling, rotation, shifting) and interpolation of image intensity values in non-grid points, 5) two image prefilters (PF), which shift and scale images so that the pixels have a zero mean and unit variance. Additionally, the noise from images is removed by low pass filtering with gaussian filter (standard deviation set to 2.0).

In earlier navigation systems [1, 2, 3] the position of virtual camera in respect to virtual bronchial tree was iteratively changed in optimization loop in order to minimize the difference between the real and the virtual image according to chosen similarity measure. A system working in such a way requires generation of many images by virtual bronchoscopy system, so its computational complexity is very high. To reduce computational complexity requirement and enabling real-time operation two-stage optimization strategy has been proposed [4].

In the first stage, new incoming image (t_i) from real, endoscope camera is registered to the previous image (t_{i-1}) obtained from the virtual bronchoscopy system. As a results the values of transformation parameters (scale, rotation, translation) which yields the registration of the images t_i and t_{i-1} is obtained. If the calculated parameters' values differ less than assumed value ε from the parameters of identity transformation, one can assume that the actual position of virtual camera corresponds with the position of real camera in respect to the virtual/real bronchial tree. In the other case, the second stage of the algorithm is performed: the position of virtual camera is shifted and and new virtual CT-based image is generated t_i. Then the first stage, the comparison of the images, is repeated. Because such a strategy reduces significantly the number of images generated by virtual bronchoscopy system it decreases the computational complexity of the whole navigation process. Additional complexity reduction is achieved by application of multi-scale optimisation in image registration algorithm.

A. Image registration problem

In general, image registration is the process of determination parameters of the transform T that brings into spatial correspondence two images I_A and I_B. As a measure of spatial correspondence function M called similarity measure is used. The image registration process can be described as a task of function F minimization [5]

$$F : R^N \to R_+ : F(T) = M(I_A(p), I_B(T(p))) .$$

where N is a number of transform T parameters. If images undergoing registration are obtained with different diagnostic methods, e.g. CT and MRI or, as in our case, virtual/real camera, their intensities are quite different. In such a case similarity measures based on mutual information concept known from information theory are frequently used.

B. Image similarity measure

Image similarity measure based on mutual information concept, has been proposed by Viola and Wells [6].

$$I(u,v) = H(u) - H(u \mid v) \tag{1}$$

In (1) entropy $H(u)$ of a random variable "u" expressed as an expectation of the negative logarithm of the probability density $f_u(u)$

$$H(u) = -E\left[\ln f_u(u)\right] = -\int_{-\infty}^{\infty} f_u(u) \ln f_u(u) du$$

denotes the measure of uncertainty about the value of random variable u. $H(u|v)$ denotes the same measure but determined with the assumption that value of random variable v is known. In this way $I(u,v)$ expresses how much the uncertainty about value of u decreased after getting to know the value v. It is obvious that if value of conditional entropy $H(u|v)$ decreases, the value of mutual information $I(u|v)$ increases. Using the Bayesian theorem

$$f_{uv}(u,v) = f_u(u|v)\, f_v(v)$$

and the definition of entropy, the equation expressing mutual information (1) may be rewritten into the form

$$I(u,v) = H(u) + H(v) - H(u,v) \tag{2}$$

where joint entropy

$$H(u,v) = -E\left[\ln f_{uv}(u,v)\right]$$

is determined on the basis of joint probability density $f_{uv}(u,v)$.
In image processing there is no direct access to probability density function, so it has to be estimated. In [6] Parzen window method has been proposed as differentiable estimator of probability density function

$$f_x(\mathbf{x}) \approx \frac{1}{|A|} \sum_{\mathbf{x}_j \in A} G_{\psi x}(\mathbf{x} - \mathbf{x}_j)$$

where $G_{\psi x}$ is properly selected smooth, non-negative, symmetric, zero mean and integrate to one function (e.g. Parzen window), A is set of elements drawn from X (e.g. randomly chosen points from image) while $|A|$ is cardinality of A (usually few hundred). If $\mathbf{x}$ is a vector, usually as a Parzen window multidimensional (N) Gaussian density function

$$G_{\psi x}(\mathbf{x}) = \frac{1}{\sqrt{(2\pi)^N |\boldsymbol{\psi}_x|}} \exp(-\frac{1}{2}\mathbf{x}^T \boldsymbol{\psi}_x^{-1} \mathbf{x})$$

is used, where $\boldsymbol{\Psi}_\mathbf{x}$ is the covariance matrix. Process of estimation of *pdf* for test image from Figure 3 is presented on Figure 2.

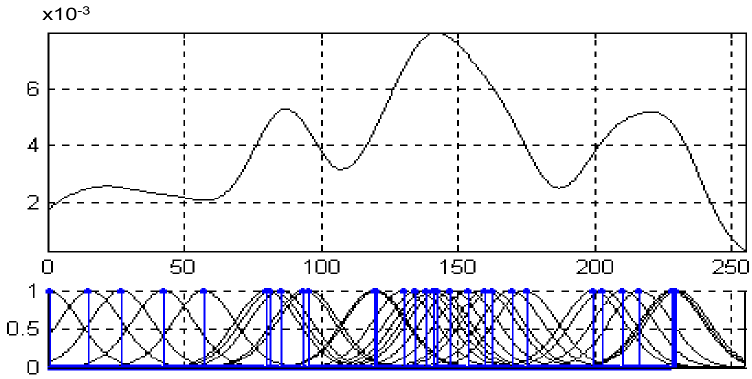

Figure 2. Process of estimation of *pdf* for test image (from Figure 3) by superimposing Parzen window centered on the samples obtained from the image.

Having probability density estimator $f_x(\mathbf{x})$ entropy $H(\mathbf{x})$ can be approximated as a mean of samples from set B [6] (e.g. another set randomly chosen points from the same image):

$$H(\mathbf{x}) = -E\left[\ln f_x(\mathbf{x})\right] = \frac{-1}{|B|} \sum_{\mathbf{x}_i \in B} \ln\left(\frac{1}{|A|} \sum_{\mathbf{x}_j \in A} G_{\psi}(\mathbf{x}_i - \mathbf{x}_j)\right)$$

In image registration one seek for such transform T parameters which maximize chosen similarity measure. In order to find maximum of mutual information (2) the derivative of mutual information with respect to transform T parameters has to be computed.

$$\frac{d}{dT} I(u(x), v(T(x))) = \frac{d}{dT} H(u(x)) + \frac{d}{dT} H(v(T(x))) - \frac{d}{dT} H(w(x))$$

where

$$w(x) = \left[u(x), v(T(x))\right]^T$$

Denoting $u_i = u(x_i)$, $v_i=v(T(\mathrm{x}_i))$ as intensity value of samples with coordinate x_i and $T(\mathrm{x}_i)$ respectively taken from image u and v and defining vector w_i=[$u(x_i)$, $v(T(\mathrm{x}_i))]^{\mathrm{T}}$ one gets [6]

$$\frac{dI}{dT} = \frac{1}{|B|} \sum_{x_i \in B} \frac{1}{|A|} \sum_{x_j \in A} (v_i - v_j) \sigma_v^{-2} \left[W_v(v_i, v_j) - W_w(w_i, w_j)\right] \frac{d}{dT}(v_i - v_j)$$

where

$$W_v(v_i, v_j) = \frac{G_{\psi v}(v_i - v_j)}{\sum_{x_k \in A} G_{\psi v}(v_i - v_k)}, \qquad W_w(w_i, w_j) = \frac{G_{\psi w}(w_i - w_j)}{\sum_{x_k \in A} G_{\psi w}(w_i - w_k)}$$

For simplicity, during estimation of joint entropy $H(u(x), v(T(x)))$ separable Parzen window is used. This is equivalent to assumption that covariance matrix Ψ_w is diagonal.

$$\boldsymbol{\psi}_w = diag(\sigma_u^2, \sigma_v^2), \quad \psi_v = \sigma_v^2$$

This assumption can be justified by the low correlation of samples $u(x_i)$ and $v(T(x_i))$, when images are unregistered. In our implementation, covariance σ_u^2 and σ_v^2 of images' samples from real and virtual cameras have been set to 0.4.

C. Multi-scale image registration

In proposed implementation 5-level multi-scale decomposition is used. Registration starts at the lowest resolution level (the highest approximation level) and is iteratively repeated on larger ones. Adaptation speed of optimization procedure is different on each image decomposition level. We assume that the images from real and virtual cameras show structures or tissues which behave as rigid body objects so the transform T can be described by the formula

$$T(\mathbf{p})=\lambda\mathbf{R}\mathbf{p} + \mathbf{t},$$

where λ is scaling factor, $\mathbf{t}=[t_x, t_y]$ represents translation's vector and

$$\mathbf{R} = \begin{bmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{bmatrix}$$

is rotation matrix. At each level of image pyramid similarity measure function is maximized with the help of gradient based optimizer.

$$T \leftarrow T + \mu S \frac{dI}{dT}$$

Because scale of units in translation (**t**), rotation (**R**) and scaling coefficient (λ) is quite different elements of gradient similarity measure vector are additionally scaled (S). At the lowest resolution level, where translation is supposed to be large, elements of gradient vector corresponding to the translation are scaled 10 times higher then others. At higher resolution levels where images are already quite well registered scaling factors decrease. In order to make registration process faster and stable the length of step size (μ) is halved with the changes of image pyramid level.

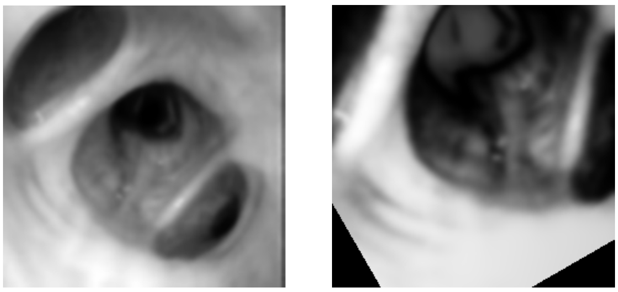

Figure 3. Real endoscopic image (left) and its initially deformed version (right).

III. Results

The described above system has been implemented in C++ language making use of the open source National Library of Medicine Insight Segmentation and Registration Toolkit [7]. In order to verify the efficiency of the proposed solution and to calibrate the system the following test has been performed. An exemplary endoscopic camera image has been acquired and its luminance has been nonlinearly deformed (see Figure 3). Then on both images, the original and the modified one (playing the role of a CT-based image), the proposed multi-scale image registration procedure has been started. Initial transform parameters have been set to be far away from *identity transformation* (scaling = 0.7, rotation = 0.52 radians, translation in both directions = -32 pixels). Convergence speed of the parameters values to *identity* ones has been observed on consecutive levels. The starting images are presented in figure 3. History of the transform parameters values adaptation is presented in table 1. One can observe that smaller number of iterations is required for registration of larger images. Recovered images on consecutive decomposition levels (from 0 to 4) are presented in Figure 4. Despite of initial severe deformation the image registration algorithm works well.

IV. Conclusions

In the paper some algorithmic enhancements of virtual bronchoscopy navigation system [1,2,4] have been proposed and successfully verified. The real-time application of the system is planned in the near future at the Jagiellonian University Medical College, Krakow, Poland.

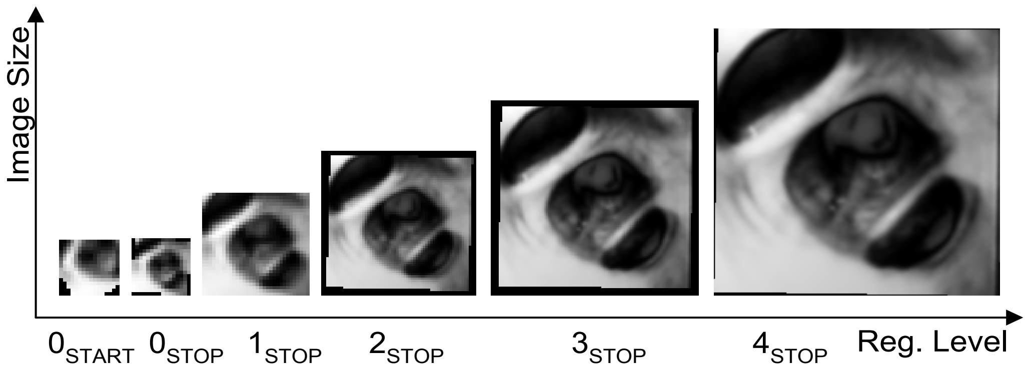

Figure 4. Image registration on consecutive decomposition levels.

Table. 1. History of the transform parameters values adaptation.

Action	Level	No of iterations	Step size (μ)	Scale coefficient (λ)	Rotation coefficient (φ) [rad]	X translation coefficient (t_x) [pixels]	Y translation coefficient (t_y) [pixels]
Start	--	--		**0.7**	**0.52**	**-32**	**-32**
Stop	0	50	0.2	0.9770	0.1133	-1.824	-2.4336
Stop	1	25	0.1	0.9179	0.0571	-1.4248	-1.0936
Stop	2	12	0.05	1.0708	0.0276	-1.444	-1.0132
Stop	3	6	0.025	1.0772	0.0151	-1.4428	-1.1011
Stop	4	3	0.00625	1.0010	0.0138	-1.4427	-1.0111

References

[1] J.P. Helferty, W.E. Higgins, "Combined Endoscopic Video Tracking and Virtual 3D CT Registration for Surgical Guidance," *IEEE Int. Conf. Image Processing,* pp. II-961-964, 2002.

[2] W.E. Higgins, J.P. Helferty, D.R. Padfield, "Integrated Bronchoscopic Video Tracking and 3D CT Registration for Virtual Bronchoscopy," *SPIE Medical Imaging 2003,* SPIE Proc., A. Clough and A. Amini, eds. vol. 5031, pp. 80-89, 2003.

[3] K. Mori, Y. Suenaga, J. Toriwaki, et al., "Tracking of camera motion of real endoscope by using the Virtual Endoscope System," *Proc. of CARS2000*, pp. 85-90, 2000.

[4] P. Turcza, M. Duplaga, "Navigation Systems Based on Registration of Endoscopic and CT-drived Virtual Images for Bronchofiberoscopic Procedures". Transformation of Healthcare with Information Technologies M.Duplaga et al. (Eds), IOS Press, 2004, ISBN 1 58603 438 3

[5] B. Zitova, J. Flusser, "Image registration methods: a survey," *Image and Vision Computing* vol. 21, pp. 977–1000, 2003.

[6] P. Viola, W.M. Wells III, "Alignment by maximization of mutual information," in Int. Conf. on Computer Vision 1995 Eds.: E. Grimson et al, *IEEE Computer Society Press*, Los Alamitos, CA., pp. 16-23, 1995.

[7] www.itk.org

Acknowledgments

Author would like to thank Prof. Krzysztof Sładek, MD, PhD, Head of Department of Interventional Pulmonology, Jagiellonian University Medical College (Krakow, Poland) for video recordings of bronchofiberoscopic procedures.