NAVIGATION SYSTEM FOR BRONCHOFIBEROSCOPIC PROCEDURES BASED ON IMAGE REGISTRATION WITH SCALE ADAPTIVE IMAGE SIMILARITY MEASURE

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

Navigation system for bronchofiberoscopic procedures establishes the location of the tip of the endoscope through the registration of two images: the real image from endoscope camera and the virtual bronchoscopy image obtained in real-time from computer tomography data. Because of the difference between images modality multiscale image registration algorithm with appropriate choice of similarity measure at each level of registration algorithm is proposed. The motivation for such an approach is observation that at a lowest resolution level difference between images modality (resulting in difference in images appearance) have much less impact on similarity measure value than mismatch in virtual and real camera position. Presented method is robust to local minima, guarantee faster convergence and enables real-time implementation.

1. INTRODUCTION

One of the most important bronchofiberoscopy procedure is transbronchial needle aspiration (TBNA). It allows for neoplasmatic tissue sampling as well as lymph nodes situated outside trachea and bronchial tree for staging of lung cancer. Many efforts were undertaken to increase the sensitivity of this procedure including the application of endobronchial ultrasonography. The most suitable points for performation of needle aspiration during bronchofiberoscopy 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 [3]. A promising complex navigation system, exploiting virtual bronchoscopy presentation and image registration based on mutual information (MI) maximization concept [8], has been developed by Helferty & Higgins [4, 5, 6].

Novelty of this paper relies on introduction of some modifications to the Helferty & Higgins navigation system that aim reduction of computational complexity and real-time implementation as well as increasing range of registration parameters. They involve faster two-stage optimization strategy of image registration and scale adaptive similarity measure. Application of faster two-stage optimization strategy of image registration significantly reduces the number of images generated by virtual bronchoscopy system and in this way decreases the computational complexity of the whole navigation system. Further reduction in computational complexity is obtained by applying image registration with multi-resolution pyramid and appropriate choice of similarity measure and optimiser parameters at each pyramid level.

Registration is stared at the lowest resolution level where the difference between images modality (resulting in difference in images appearance) has much less impact on similarity measure value than misalignment in virtual and real camera position. Therefore, at this level as a similarity measure normalised correlation is used. Such an approach helps to avoid many local minima and is computationally efficient. At the higher levels of image pyramid where images are already quite well registered similarity measure based on mutual information is used. Initial verification of the proposed solutions is presented in the paper.

2. NAVIGATION SYSTEM FOR BRONCHOFIBEROSCOPY PROCEDURES

2.1. General information

The block diagram of the proposed and discussed navigation system for bronchofiberoscopy based on image registration is presented in the figure 1. The system consists of: 1) the video frame grabber acquiring endoscopic camera images, 2) computer system generating virtual images from CT data, 3) optimizer which is responsible for determining generation (visualization) and transformation parameters' values for virtual images in such a way so as to make them more similar to endoscopic images according to measure M, 4) module for virtual image transformation (scaling, rotation, shifting), 5) module for interpolation of the transformed virtual image or for determination of its intensity values in non-grid points.

Through last decades methods enabling registration of images from the same or different sources have been extensively developed [1] [2]. Image registration is the process of determination parameters μ of the transform T that brings into spatial correspondence two images I_A and I_B. As a measure of their spatial correspondence function M called similar-

This work has been supported by Polish Committee for Scientific Research, grant no. T11E 039 27

ity measure is used. The image registration process can be described as a task of function F minimization

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

In case of registration of images obtained with different diagnostic methods (multi-modality), 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.

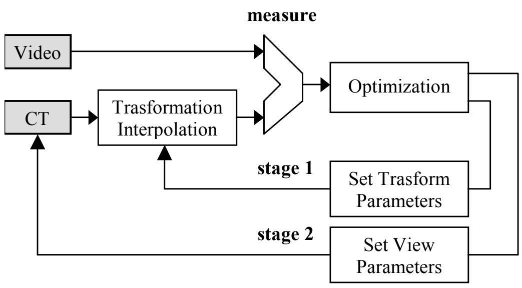

Figure 1. The block diagram of new navigation system

General measure of image similarity based on information theory has been proposed by Viola and Wells [7]:

$$M(u,v) = H(u) - H(u \mid v)$$

where $H(u)$ denotes the measure of uncertainty about the value of random variable u, and $H(u|v)$ denotes the same measure but determined with the assumption that value of random variable v is known. In this way $M(u,v)$ 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 $M(u|v)$ increases. Using the Bayesian theorem: $P(A,B)=P(A|B)P(B)$ and the definition of Shannon entropy the equation expressing mutual information (MI) may be rewritten into the form

$$M(u,v) = H(u) + H(v) - H(u,v)$$

where

$$H(u) = -\sum_i p_u(i) \log p_u(i)$$

$$H(u,v) = -\sum_i p_{uv}(i) \log p_{uv}(i)$$

This equation includes joint entropy $H(u,v)$, which is determined on the basis of joint probability distribution which in turn, can be inferred from the joint histogram $h(u,v)$ after appropriate normalization.

If the differences between images modality do not causes significant differences in images intensities normalized correlation as a similarity measure can be use.

$$M(I_A, I_B) = \frac{\sum_{i,j} I_A(i) I_B(j)}{\sqrt{\sum_i I_A^2(i) \sum_j I_B^2(j)}}$$

Such a measure has lower implementation cost than MI and its "not-random" nature results in faster convergence of registration process.

2.2. New navigation (optimization) strategy

The main difference between the presented (figure 1) and earlier systems [3, 4, 5, 6] is the method used for the estimation of real position of endoscopic camera (bronchofiberoscope tip) in the bronchial tree. In previous solutions the position of virtual camera in respect to virtual bronchial tree has been iteratively changed in optimization loop so as to minimize the difference between the real and virtual image (according to the chosen similarity measure). Such a strategy suffers from high computational complexity associated with generation of a great number of images by virtual bronchoscopy system.

In the proposed navigation system, the computational complexity of bronchofiberoscope tip localization algorithm may be decreased through introduction of a new two-stage optimization strategy.

In the first stage, new incoming image (t_i) from endoscope camera is registered to the previous image (t_{i-1}) obtained from the virtual bronchoscopy system. As a results we get the values of transformation parameters (scale, rotation, translation) which yields the registration of the images (t_i) and (t_{i-1}).

If the calculated parameters' values differ from the parameters of identity transformation less than a given value ε, we assume that the actual position of virtual camera corresponds with the position of real camera in respect to the virtual/real bronchial tree. Otherwise, the second stage of the algorithm is performed: the position of virtual camera is changed and a new virtual CT-based image is generated. Afterwards we return to the first stage and the comparison of the images is repeated. Such a strategy reduces significantly the number of images generated by virtual bronchoscopy system and in this way decreases the computational complexity of the whole navigation system.

2.3. Multiresolution image registration

Applications of the multi-resolution pyramids [10] for image analysis speed-up has a long history. Application of this technique to virtual endoscopy navigation system results in its computational complexity reduction and faster image registration. Additionally such a solution offers the following advantages:

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

In our implementation 5-level multi-resolution decomposition is applied. Registration starts at the lowest resolution level where normalized correlation as a similarity measure is used. 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

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

where λ is scaling factor, $\mathbf{R}$ is rotation matrix and $\mathbf{t}$ represents translation's vector. At each level of image pyramid similarity measure function M is maximized with the help of gradient based optimizer

$$\boldsymbol{\mu}_{k+1} = \boldsymbol{\mu}_k - s_k \frac{\nabla M(\boldsymbol{\mu}_k)./\Gamma_k}{\left\| \nabla M(\boldsymbol{\mu}_k)./\Gamma_k \right\|}$$

$\nabla M(\mu_k)$ represents gradient of similarity measure function M computed at μ_k point, $\|.\|$ denotes vector norm and ./ is element-by-element division. Vector of transform T parameters in k-th step is defined as

$$\boldsymbol{\mu}_k = (\lambda, \varphi_z, t_x, t_y)$$

In order to make registration process faster and stable elements of gradient vector are scaled by factor

$$\Gamma_k = (\alpha - 1)\Gamma_{k-1} + \alpha \left| \Delta \left[\nabla M(\boldsymbol{\mu}_k) \right] ./ \Delta \boldsymbol{\mu}_k \right|$$

where $\Delta\mu_k = \mu_k - \mu_{k-1}$ and α=0.2. The step size s_k is set to 0.2 at the beginning of registration process and is halved at each change of search direction.

At higher levels of image pyramid influence of difference between images modality on similarity measure value can not be neglected. At those levels mutual information based similarity measure [9] is used.

2.4. Mutual information based similarity measure

Further reduction of computational complexity can be achieved in the Helferty & Higgins navigation system [6] by the use of different method for estimation of joint and marginal probability distributions that are required for determination of MI measure value. In [6] these probability distributions are derived from marginal and joint histogram as follows

$$p(k) = \frac{1}{NM} \sum_{i=1}^{N} \sum_{j=1}^{M} \delta\left(k - I(i,j)\right)$$

$$p_{AB}(k,l) = \frac{1}{NM} \sum_{i=1}^{N} \sum_{j=1}^{M} \delta\left(k - I_A(i,j)\right) \delta\left(l - I_B(i,j)\right)$$

where δ denotes the Kronecker delta function. Following [7, 9], the author of this paper propose usage of the Parzen window method to their estimation:

$$p(k) = \frac{1}{N_A} \sum_{a_i \in I} w\left(k - a_i\right)$$

$$p_{AB}(k,l) = \frac{1}{N_A} \sum_{\substack{a_i \in I_A \\ b_i \in I_B}} w\left(k - a_i\right) w\left(l - b_i\right)$$

where w is a properly selected smooth, non-negative, symmetric, zero mean and integrate to one function (cubic spline is used for this purpose), and N_A denotes the number of randomly chosen points a_i , b_i from images I_A and I_B.

Computational complexity of the probability distribution estimator used in [6] is of the order of NxM, where N denotes number of rows, and M number of columns of the image. On the contrary, the estimator used in [9] requires only about 20 % randomly selected image points.

The second benefit from application of the estimator based on Parzen window is the existence of analytical expression of the first (gradient) and the second derivative (Hessian) of mutual information function. Values of this derivatives are necessary during optimization process and have to be approximated numerically if can not be expressed analytically. However numerical approximation not only increases the number of transformations and calculations of the similarity measure but also can lead to the lack of stability of optimization and therefore the whole registration process.

3. EXPERIMENTS

The system described above has been implemented in C++ language making use of the open source National Library of Medicine Insight Segmentation and Registration Toolkit [11]. In order to verify the efficiency of the proposed solutions 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 (see figure 2) and spatially deformed:

$$x_{r,c} = 0.5\left(1 + \frac{0.2r}{\max(r)} - \frac{0.2c}{\max(c)}\right)\left(1 + \sin\left(1.3\frac{\pi}{2}(x_{r,c} - 1)\right)^3\right)$$

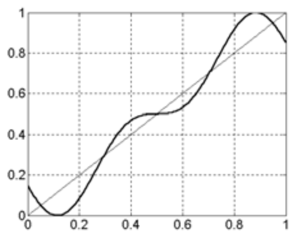

Figure 2. Nonlineal part of image intensity deformation formula.

Then on both images, the original and the modified one (playing the role of a CT-based image), the proposed two-stage multi-resolution 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. Spline based interpolation and 5-level image decomposition have been used.

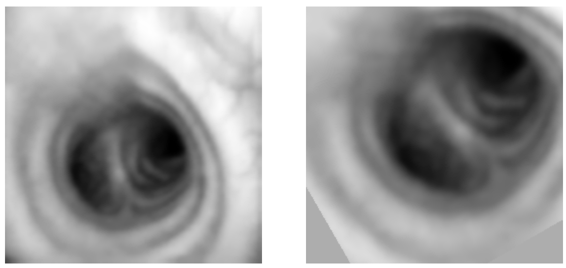

Figure 3. Real endoscopic image (left) and its initially deformed version (right).

4. RESULTS

History of the transform parameters values adaptation is presented in figures 4 and 5. Recovered images on consecutive decomposition levels (from 0 to 4) are presented in figure 6. In this figure also number of iterations at each registration level is denoted. One can observe that smaller number is required for registration of larger images. Despite of initial severe deformation the image registration algorithm works well.

5. CONCLUSIONS

In the paper some algorithmic modifications of the Helferty & Higgins virtual bronchoscopy navigation system have been proposed and successfully verified. They regarded reduction of its computational complexity since the real-time application of the system is planned in the near future at the Jagiellonian University Medical College, Krakow, Poland.

REFERENCES

[1] D.L.G. Hill, P.G. Batchelor, M. Holden, D.J. Hawkes, "Topical review, Medical image registration," *Physics in Medicine and Biology* vol. 46, pp. R1–R45, 2001.

[2] B. Zitova, J. Flusser, "Image registration methods: a survey," *Image and Vision Computing* vol. 21, pp. 977–1000, 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] J.P. Helferty, W.E. Higgins, "Technique for registering 3D virtual CT images to endoscopic video," *IEEE Int. Conf. on Image Processing* 2001, pp.893-896, Oct. 7-10 2001.

[5] 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.

[6] 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.

[7] 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.

[8] C. Studholme, D.L.G. Hill, D.J. Hawkes, "An overlap invariant entropy measure of 3D medical image alignment," *Pattern Recognition* vol. 32, pp. 71-86, 1999.

[9] D. Mattes, D. R. Haynor, H. Vesselle, T. K. Lewellen, and W. Eubank. Non-rigid multi-modality image registration. In *Medical Imaging 2001: Image Processing*, pages 1609–1620, 2001.

[10] S. Mallat, *A Wavelet Tour of Signal Processing*. San Diego, Academic Press 1998.

[11] www.itk.org

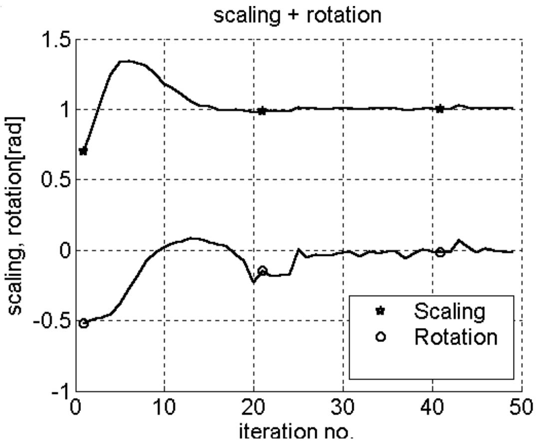

Figure 4. Rotation angle and scale value at each iteration of the registration.

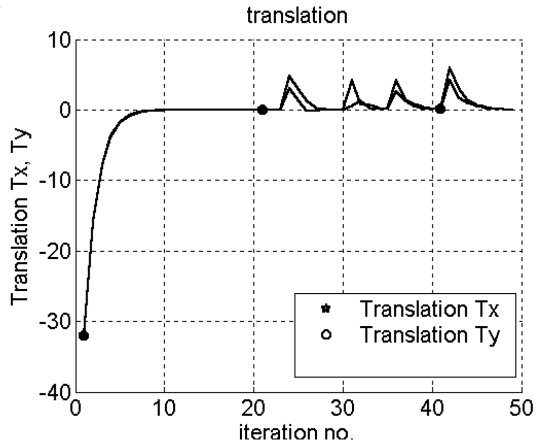

Figure 5. Translation value in x and y at each iteration of the registration.

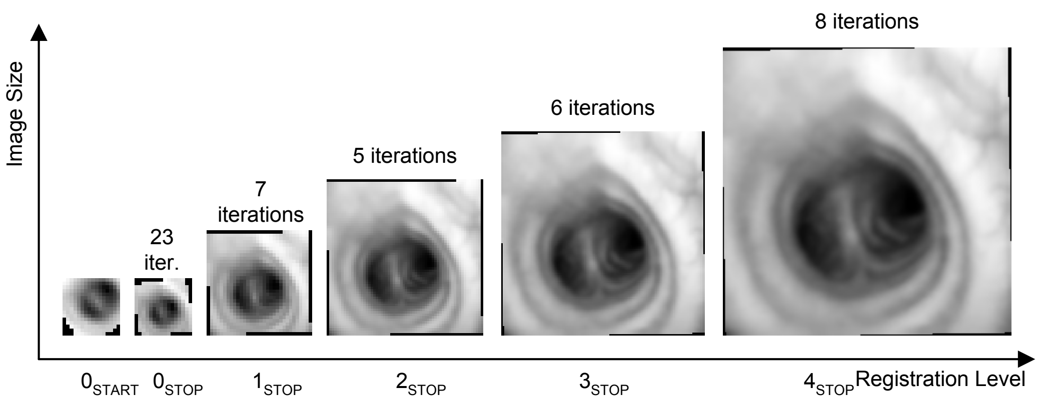

Figure 6. Image registration on consecutive decomposition levels (relative sizes of images have been changed in order to improve the clarity of presentation).