Towards registration of multimodal images of vocal folds based on mutual information.

Mouloud Ourak, Brahim Tamadazte, Nicolas Andreff, Christian Pieralli, Olivier Gaiffe, Bruno Wacogne

Towards registration of multimodal images of vocal folds based on mutual information

M. Ourak*[1], B. Tamadazte[1], N. Andreff[1], C. Pieralli[1], O. Gaiffe[2] and B. Wacogne[2]

[1]FEMTO-ST institute-UFC/CNRS/ENSMM/UTBM. 24, rue Alain Savary, 25000 Besançon, France.
[2]University Hospital Besançon (CHU). 3, boulevard Alexandre Fleming, 25000 Besançon, France.
(*)Corresponding author : mouloud.ourak@femto-st.fr

***Abstract*— This paper deals with mutual information-based registration of multimodal images for laser phonomicrosurgery of the vocal folds. The images to be registered are white light images (white light camera) versus fluorescence images. This work is carried out within the framework of the European project μRALP which involves the use of microrobotic system for endoluminal laser phonosurgery. The designed system includes two fiber bundles connected to a high speed camera and one fiber bundle used for fluorescence image. Using the mutual information based registration method, it will be possible to represent the visible information in the fluorescence image and use it in the other image.**

I. INTRODUCTION

A phonosurgery system uses a stereo-microscope and a semi-automatic system to control the laser source positioned at 400 mm outside the patient (see Fig. 1). This technique relies completely on the skills of surgeons who must operate through a microscope, control the laser directly by hand, and deals with the associated poor ergonomics of the operating setup [1], [2]. The μRALP project will enable to overcome such limitations by developing a more intuitive system (inserted through the mouth into the laryngo-pharyngeal cavity) that allows to position the laser close to the vocal chords (*ie.* 20 mm). Mainly, the proposed system includes a piezoelectric actuated mirror (two degrees of freedom), two fiber bundles for high frequency imaging, one fiber bundle for fluorescence image, and a Helium-Neon (HeNe) laser (see Fig. 2).

The developed system will work as follows: the surgeon defines a trajectory around the tumor to incise, in the fluorescence image. Thereby, the objective of μRALP system is to control automatically the displacement of the laser beam along the pre-defined trajectory. In order to ensure a better accuracy during automatic resection, it is required to close the control loop over an exteroceptive sensor (high speed camera in our case). Namely, this so-called visual servoing (vision-based robot control) is robust to robot calibration errors and to robot environment variations [3].

In general, visual servoing requires defining a reference features (points, lines, image, 3D pose, etc.) that allows the system to evolve from its current position to its target position. In our case, the reference can be defined (by the surgeon) in the fluorescence image frame R_f and the system will move in the high speed white light image frame R_s. Thus, it is necessary to compute the transformation $(\widehat{T}, \widehat{\rho})$ between both frames (*ie.* R_f and R_s). The mutual information-based method seems very appropriate for the computation of this transformation, especially when the images are of different modalities.

In the literature, the mutual information-based registration has been widely discussed. Zitova et *al.* [4] has classified the registration techniques for medical applications into two main categories: area-based and feature-based methods. Generally, a registration technique follows mainly these four steps: features detection, features matching, transformation estimation, and image resampling.

This paper focuses in the adaptation of the mutual information technique in the case of fluorescence images *vs.* high

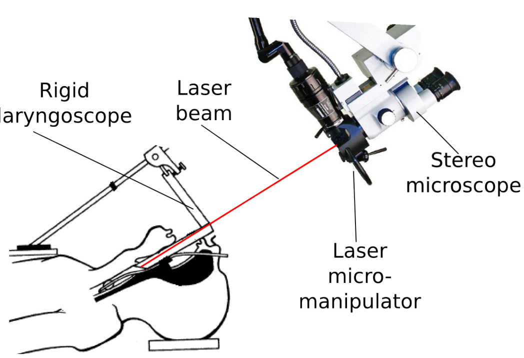

Fig. 1. Current laser phonosurgery setup.

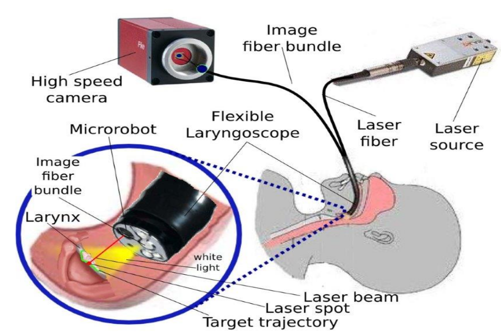

Fig. 2. New laser phonosurgery system proposed by the μRALP project.

speed white light images. Thereby, it contains some basic definitions about the mutual information, the context of use, some primary results and the future investigations of this work.

II. Mutual Information

The mutual information is inspired by the information theory domain. It is based on the measure of information, commonly called entropy, in a message. In the case of image processing, this allows to measure the similarity between two images I_1 and I_2.

A. Entropy

In information theory domain, the entropy was first introduced by Hartley [5]. There, he used it to measure the uncertainty of a signal s between a transmitter and a receiver. The first Hartley entropy function, is given by:

$$H_0(X) = log_b\,|X| \tag{1}$$

where X is a vector containing random values x_i.

Move over, the mean information quantity associated to each source value is the expectation of the own information of each event $X = x_i$. According to Shannon, adding a weight to the output of the Hartley entropy allows to rewrite the equation (1) as:

$$H = \sum_i p_i log_2(\frac{1}{p_i}) = -\sum_i p_i log_2(p_i) \tag{2}$$

where p_i is the associated probability of each event x_i.

Applied to the images, the measure of the entropy allows us to compute the degree of similarity information between two images (higher is the similarity between images, lower is the entropy).

B. Joint histogram

To estimate the joint probability distribution of the gray values between two images, it is possible to use the notion of joint histogram. This joint histogram is a 3D map ($\mathbb{N}^2$ in $\mathbb{N}$), whose 2D inputs are the grayscale value of each pixel of $I_1(i,j)$ and of the corresponding pixel in $I_2(i,j)$ and the output is the amount of such couple $(I_1(i,j), I_2(i,j))$ (see Fig. 3 and algorithm 1).

Thereby, the joint histogram allows illustrating the similarity between two images. In practice, lower the similarity between images, more the joint histogram is uniform (its

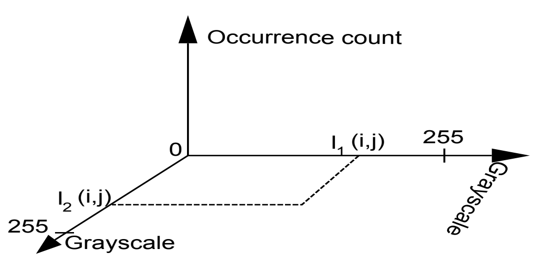

Fig. 3. Representation of the joint histogram.

Algorithm 1 Joint histogram algorithm

Require: I_1, I_2 two images
$init : binSize \leftarrow 255;$
$init : h(binSize, binSize);$
$u \leftarrow height(I_1);$
$v \leftarrow width(I_1);$
for i from 0 to u **do**
  for j from 0 to v **do**
    $h(I_1(i,j), I_2(i,j)) = h(I_1(i,j), I_2(i,j)) + 1;$
  end for
end for
return h

graph is a plane). Conversely, if two images are perfectly similar, the joint histogram is represented by a diagonal line.

C. Mutual information

Based on the entropy and the joint histogram principles, the mutual information can be defined following the three methods described in [6], [7].
However, for our work, we use the method using the marginal and joint entropy based on the value of the normalized joint histogram, given by (3).

$$MI(I_1, I_2) = H(I_1) + H(I_2) - H(I_1, I_2) \tag{3}$$

where $H(I_1)$ and $H(I_2)$ are the marginal entropy of I_1 and I_2, respectively, and $H(I_1, I_2)$ is joint entropy computed from the joint histogram.

Let us assume the rigid transformation $\widehat{T} \epsilon SE(3)$ between the images I_1 and I_2. Thereby, the transformation can be estimated by maximizing the mutual information $(MI)_i$.

$$\widehat{T} = arg \max_{t \in SE(3)} MI[I_1, t(I_2)] \tag{4}$$

where t is a possible rigid transformation. If we add a zoom factor, then we need to multiply the transformation matrix with the scaling factor:

$$zoom = \begin{pmatrix} \rho & 0 & 0 \\ 0 & \rho & 0 \\ 0 & 0 & 1 \end{pmatrix} \tag{5}$$

where ρ is the scale value. Hence, the equation (4) can evolve using the zoom parameter as:

$$(\widehat{T}, \widehat{\rho}) = arg \max_{t,\rho} MI[I_1, \rho(t(I_2))] \tag{6}$$

III. Optimization

Once the basics of the mutual information are addressed, it is necessary to establish an effective optimization method able to compute the maximum (respectively the minimum) of the joint information between both images. In this work, we opted to use a non-gradient-based optimization technique commonly known as *the simplex method* [8]. This choice is justified by the fact that the *simplex* is better fitted to an optimization problem with limited variables number [9].
The different steps of the presented mutual information based images registration are shown in Fig. 4.

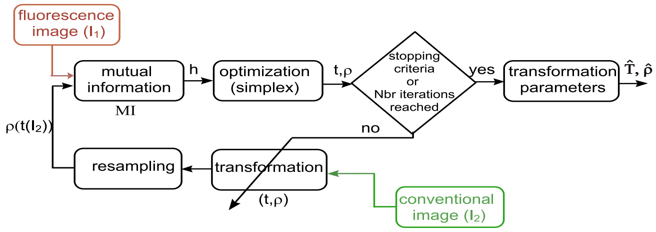

Fig. 4. Key steps of the proposed mutual information based registration.

IV. First Results

The experimental validation is performed with the objective to determine the estimated transformation $(\widehat{T}, \widehat{\rho})$ applied to image I_2 to be confounded with the image I_1. This transformation is here decomposed in translations (t_x and t_y), rotation θ_z and zoom ρ.

The validation tests were carried out using different images, between monomodal images acquired in various conditions (point of view, lighting illumination, adding objects onto the scene and scale changing), and between fluorescence images *vs.* white light image of vocal chords. This paper show only the results using fluorescence *vs.* white light images. In this experiment, we added a fictive incision mark on the fluorescence image where the tumor is visible. To display the difference between image I_1 and the transformed image $\rho(t(I_2))$, we combine them as their checkerboard masked sum (Fig. 5). Thus a quality criterion of the registration is the continuity of the combined image a long the check-board lines.

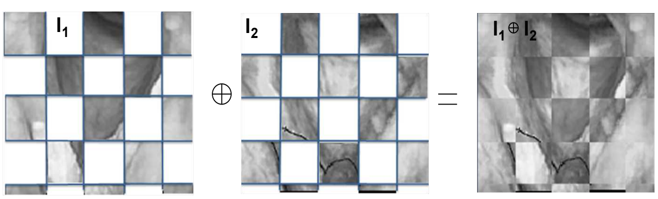

Fig. 5. Illustration of the construction of the result image ($\oplus$ operator to combine the two images I_1 and I_2).

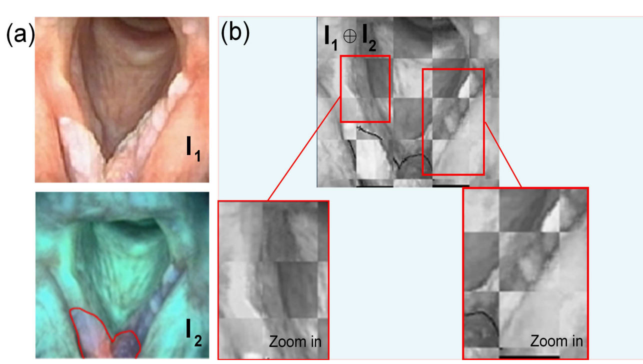

Fig. 6. Illustration of the registration results between I_1 and I_2. Column (a) represents the test images (*ie.* fluorescence image *vs.* white light image), and column (b) represents the obtained combined image with a zoom in to show the details.

By analyzing the Fig. 6, it can be noticed the continuity of the combination of I_1 in I_2 as shown in Fig. 6(b).

Otherwise, the obtained values of the different parameters of $(\widehat{T}, \widehat{\rho})$ for the test shown in Fig. 6 are: t_x = -13.92 pixels, t_y = -1.31 pixels, θ_z = -0.0141°, and ρ = 0.158 pixels. This result is obtained in 110 seconds for a 170×150 image size using a 2.5 GHz PC.

V. Conclusion

This paper presented the initial results using the proposed mutual information registration method. The considered images are multimodal (*ie.* fluorescence image *vs.* white light image), and such image pairs will be used to detect the cancerous tumor in the vocal chords. The mutual information based registration method gives promising results and shows interesting robustness with respect to different kind and quality of the images.

The next step is to develop a new metric for visual servoing using the proposed technique. A new control law, which does not require any matching nor tracking step (direct visual servoing), based on mutual information will be designed. This is to ensure that the surgeon can plan a trajectory (*ie.* incision) on the fluorescence image and the endoluminal laser phonosurgery microrobot system will be able to follow this reference in the high speed white light images.

Acknowledgment

This work was supported by μRALP, the EC FP7 ICT Collaborative Project no. 288663 (`http://www.microralp.eu`), and by ACTION, the French ANR Labex no. ANR-11-LABX-01-01 (`http://www.labex-action.fr`).

References

[1] J. A. Burns, G. Har-El, S. Shapshay et al., *Endoscopic Laser Resection of Laryngeal Cancer: Is It Oncologically Safe? Position Statement From the American Broncho-Esophagological Association*, Annals of Otology Rhinology and Laryngology, vol. 118(6), pp. 399-404, 2009.

[2] M. C. Jackel, A. Martin, and W. Steiner, *Twenty-five Years Experience with Laser Surgery for Head and Neck Tumors*, European Archives of Oto-Rhino-Laryngology, vol. 264(6), pp. 577-585, 2007.

[3] Chaumette and S. Hutchinson. *Visual Servo Control, Part I: Basic Approaches*, IEEE Robotics and Automation Magazine, vol. 13(4), pp.8290, December, 2006.

[4] Barbara Zitova, Jan Flusser, *Image Registration Methods: a Survey*, Image and Vision Computing. vol. 977-1000, June 24, 2003.

[5] R. V. L. Hartley, *Transmission of Information,* Bell System Technical Journal, vol. 7, pp. 535563, 1928.

[6] Josien P. W. Pluim, J. B. Antoine Maintz and Max A. Viergever, *Mutual Information based Registration of Medical Images: a Survey*, IEEE Transaction on Medical Imaging. vol. 22(9), pp. 986-1004, August, 2003.

[7] Amaury Dame, Eric Marchand, *Second Order Optimization of Mutual Information for Real-time image Registration*, IEEE Transaction on image Processing, vol. 21(9), pp. 4190-4203, September, 2012.

[8] J. A. Nelder and R. Mead. *A Simplex Method for Function Minimization,* Computer Journal, vol. 7, pp. 308-313, 1965.

[9] Yang-Ming Zhu, PhD, and Steven M. Cochoff, MS, *Influence of Implementation Parameters on Registration of MR and SPECT Brain Images by Maximization of Mutual Information*, IEEE transactions on Medical Imaging. December, 2001.

[10] R. John Koshel, *Enhancement of the Downhill Simplex Method of Optimization*, vol. 4832, pp. 270-282, December, 2002.