View metadata, citation and similar papers at core.ac.uk

brought to you by CORE

provided by HAL-CEA

Detection of recurrent nerve paralysis: Development of a Computer Aided Diagnosis system

Mike-Ely Cohen, Muriel Lefort, Héloïse Bergeret-Cassagne, Siham Hachi, Ang Li, Gilles Russ, Diane Lazard, Fabrice Menegaux, Laurence Leenhardt, Christophe Trésallet, et al.

To cite this version:

Mike-Ely Cohen, Muriel Lefort, Héloïse Bergeret-Cassagne, Siham Hachi, Ang Li, et al.. Detection of recurrent nerve paralysis: Development of a Computer Aided Diagnosis system. IRBM, Elsevier Masson, 2015, 36 (6), pp.367-374. <10.1016/j.irbm.2015.09.008>. <inserm-01227885>

HAL Id: inserm-01227885

http://www.hal.inserm.fr/inserm-01227885

Submitted on 12 Nov 2015

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

ABSTRACT

Recurrent nerve paralysis (RP) is one of the most frequent complications of thyroid surgery. It reduces vocal fold mobility. Nasal endoscopy, a mini-invasive procedure, is the reference procedure to detect RP, and is based on the examination of vocal fold mobility. A new approach based on laryngeal ultrasound acquisition and using a dedicated data analysis was designed to help with the automated detection of RP.

One hundred and fifty subjects were enrolled for this feasibility study: 50 controls, 50 patients with RP and 50 patients without RP, according to nasal endoscopy. The ultrasound protocol was based on a ten seconds B-mode acquisition in an axial plane during normal breathing. Image processing included three steps: 1) the detection of two consecutive closing and opening images corresponding to extreme positions of vocal folds in the sequence of B-mode images; 2) the positioning of three landmarks and the robust tracking of these points using multi-pyramidal refined optical flow approaches; 3) the estimation of quantitative parameters indicating left and right fractions of mobility and motion symmetry.

Results provided by the first two image processing steps were compared to those obtained by an expert. Motion symmetry and fraction of mobility indices were systematically computed using the automated procedures. Associated sensitivity and specificity values for detecting RP were then calculated. To optimize the performances of the system, a mixed CAD system, which integrates the automatic steps of image processing and a possible correction of its results by a trained operator was developed.

Laryngeal ultrasound combined with appropriate image processing helped in the diagnosis of recurrent nerve paralysis and could be proposed as a first–line method.

Keywords: laryngeal ultrasound, recurrent nerve paralysis, optical flow, motion analysis

1. INTRODUCTION

Thyroid surgery concerns tens of thousands of people every year in France (http://stats.atih.sante.fr). It is performed in similar proportions in developed countries. One of the major complications of this surgery is recurrent nerve paralysis (RP), which can occur in up to 10% of patients (1). Clinical markers are voice dysphonia or swallowing disorders. In most cases, the troubles are temporary and disappear within the days following the surgery. However this paralysis, especially in case of bilateral paralysis can be serious and generate higher morbidity (2). Therefore, an early detection of these troubles is crucial to propose a customized and fast care to the patients. Indeed, an early reeducation including speech therapy favors the recovery.

Nasal endoscopy, a minimally invasive procedure, is presently the gold standard method to establish RP. This technique can require some local anesthesia and may be hard to bear for patients just after surgery. This procedure is thus not systematically used, but it is proposed to patients presenting some symptoms of RP after surgery. Furthermore, the nasal endoscopy requires some dedicated material and the presence of a trained physician and thus all patients cannot beneficiate from such an exam shortly after the surgery.

Ultrasound has already been proposed to study laryngeal anatomy. Larynx motion was also studied using cine loops (3,4). Therefore, to detect recurrent nerve paralysis, we decided to use dynamic B-mode ultrasound imaging, which is a fast, widely available, cheap and non-invasive tool (5). As currently done in nasal endoscopy and as proposed by previous studies in ultrasound, the analysis of dynamic acquisitions was first based on the visual assessment of the vocal folds mobility. The anatomical description of the observable structures from ultrasound acquisitions, namely the true vocal folds, the vestibular folds, the thyroid cartilage and the arytenoids was reported in a first paper (5). True vocal folds appear as isoechoic or

hypoechoic structures, while the thyroid cartilage and arytenoids, on which vocal folds are attached, are hyperechoic well defined structures on ultrasound images. Furthermore the dynamic analysis during free breathing conditions reveals that the true vocal folds cannot be fully tracked in a 2D plane. It was thus proposed to choose the thyroid cartilage and the arytenoids as surrogate markers to study the vocal fold motion. Dedicated software was subsequently developed to study these landmarks on the dynamic acquisition. Briefly, the operator first selects two images in the sequence corresponding to abduction (opening) and adduction (closing) of vocal folds, thus showing vocal folds extreme positions during motion. Then the operator positions three landmarks: thyroid cartilage (T), and left (L) and right (R) arytenoids on each image from which several geometrical indices, including length and area measurements are finally computed. On fifty control subjects, these indices were significantly larger for abduction than for adduction. These indices were similar when computed for the left and right sides, the mid-sagittal plane being the symmetry plane.

Using these first results, this study proposes to develop and test some automated image processing methods to assist the operator in the selection of 1) the abduction and adduction images and 2) the landmarks. For that purpose, an automated detection of images corresponding to abduction and adduction of vocal folds, and thus to extreme positions of vocal folds in the sequence of B-mode images, was proposed. Furthermore a tracking of the landmarks using multi-pyramidal refined optical flow approaches was defined. These steps were evaluated by comparing results obtained by the automated methods with the selection performed by an expert on a large database including fifty control subjects, fifty patients (after thyroid surgery) with recurrent paralysis and fifty patients (after thyroid surgery) without recurrent paralysis. Finally two quantitative parameters indicating fractions of mobility, and motion symmetry were proposed and evaluated.

2. METHODS

2.1 Database and associated data

2.1.1 Population

Fifty control subjects and a hundred patients were included in this database. All selected patients had post-surgery voice disorders and thus were examined by nasal endoscopy. A paralysis of the recurrent nerve was established for fifty patients. For further classification, control subjects and patients without recurrent nerve paralysis were pooled in a first group (n=100), constituting the group without paralysis of the recurrent nerve. Patients with paralysis of recurrent nerve (n=50) defined the second group.

2.2.2 Ultrasound data acquisition

Each subject of the database underwent an ultrasound acquisition, using portable SonixTouch® from UltraSonix™ (Richmond, Canada) with a 7-14 MHz linear probe. The probe was placed over the middle portion of the thyroid cartilage. The operator was instructed to position the probe in order to provide an axial view, the horizontal axis being perpendicular to the sagittal plane, and showing thyroid and arytenoids cartilages (Figure 1). A B-mode cine loop was then recorded during 10 seconds at a frame rate of 30 images per second, while the subject breathed normally. During breathing, vocal folds and arytenoids showed successive phases of opening and closing.

2.2.3 Motion analysis

To quantify motion, a three-steps approach was proposed (5) including: 1) the detection of two images corresponding to extreme locations of the vocal folds; 2) the selection and

tracking of three landmarks; 3) the quantification of motion features from the location of these landmarks.

The expert first selected a pair of images from the cine loop using a dedicated video analysis tool; these two images I_C (closing phase) and I_O (opening phase) corresponded to two extreme positions of the vocal folds and arytenoids. Figure 1 shows an example of the selected images. Then the expert defined on each image three landmarks (Figure 1): the top of the thyroid cartilage (T), since it was a motionless point during the examination which enabled to define the two halves of the larynx, and the centers of the left and right arytenoid cartilages (L and R) since their apparent motion in the plane could be surrogate markers for studying motion of vocal folds. This procedure was applied to the 150 studies. The expert results were thereafter considered as the gold standard results for further image processing steps.

From the thyroid cartilage, the intersection of the mid-sagittal plane with the axial plane was first defined (dashed line in Figure 2), then the intersection M_L (M_R) of this axis and its perpendicular passing through the left (right) arytenoid provides one left (right) triangle TLM_L (TRM_R), the areas of which are computed. Being estimated on both I_C and I_O images, four areas were thus defined: LA_C, RA_C, LA_O and RA_O. In the example shown in Figure 2, areas defined on I_C image are obviously smaller than areas defined on I_O image.

2.2 Automated tracking of landmarks

Our first objective was to assist the operator in the definition of the landmarks on images I_C and I_O, in order to make this task less tedious. The automatic definition of the landmarks on each image appeared to be a very challenging task, not easily reachable due to the variability of patterns observed in ultrasound images. Thus we proposed to track the three landmarks (for

instance defined on the I_C Image by the operator) and to automatically report their position on the I_O image. A first approach consisted in applying different optical flow approaches to estimate motion between I_O and I_C. However our tests showed quite deceiving results (6). Thus we opted for an alternate solution consisting in estimating motion between each pair of consecutive images and to follow step by step the three landmarks. This solution was not initially retained because it was more time consuming and could generate some additional errors in the tracking, due to the iterative process. However, the motion estimation between two consecutive images can be more successful than between I_O and I_C, since two consecutive images are relatively similar to each other, and displacement amplitudes are reduced and thus more easily estimated by optical flow approaches.

In this work, an iterative optical flow approach based on a robust pyramidal implementation (7) of the Lucas-Kanade algorithm (8) was defined for each pair of consecutive images: going from I_C to I_C+1, till I_O-1 to I_O (assuming that I_C was prior to I_O). This defined the forward approach. A backward method was also tested in the opposite direction of the previous forward method: it was based on the same algorithm but it started from I_O towards I_C. In addition, a third strategy (forward, increment 2) was evaluated: the motion was estimated in the forward direction, considering every other image from I_C to I_C+2 and so on till I_O.

To evaluate these different methods of tracking, the Euclidian distances between the positions of the three landmarks estimated after automatic tracking and those defined by the expert were computed.

2.3 Definition of quantitative features to test paralysis of recurrent nerve

Assuming the paralysis of one vocal fold (the contralateral vocal fold moving normally), the difference of areas between the left larynx and the right larynx was then expected to be

different from 0. In order to get a parameter independent of the size of the larynx, this difference was scaled by the global area of the larynx, approximated by the sum of the areas of the left and right triangles defined in Figure 2. This parameter was called surface ratio (SR). As paralysis can occur either on closing or opening positions, it was computed on the images I_O (SR_O) and I_C (SR_C) and defined by equations (1) and (2):

$$SR_O = \frac{|LA_O - RA_O|}{LA_O + RA_O} \quad (1),$$

$$SR_C = \frac{|LA_C - RA_C|}{LA_C + RA_C} \quad (2).$$

SR_O and SR_C values were thus expected to be close to 0 (positive or negative) in case of symmetrical motion. One of these values, especially SR_O, was expected to be higher in case of dissymmetry (suspicion of RP).

Furthermore to quantify the degree of mobility of the left and right vocal folds, the left and right fractions of mobility LFM and RFM were defined between the two extreme images I_O and I_C by equations (3) and (4):

$$LFM = \frac{|LA_O - LA_C|}{LA_O} \quad (3),$$

$$RFM = \frac{|RA_O - RA_C|}{RA_O} \quad (4).$$

These fractions were expected to be high for normal cases and to be reduced in case of recurrent nerve paralysis. Considering these two measures, the absolute difference between left fraction of mobility and right fraction of mobility was defined according to equation (5):

$$DFM = \left| LFM - RFM \right| \quad (5).$$

All these geometrical parameters and more than 30 others which were based on the comparison of angles and line segment lengths were systematically tested considering the expert measurements. According to areas under Receiver Operating Characteristics (ROC)

curves based on expert results, which were computed for each parameter in order to select the most discriminant ones, the two parameters SR_O and DFM were shown to be the most powerful to estimate patients with recurrent nerve paralysis (9). From these learning ROC curves, two cut-off values, one minimizing the number of misclassified cases and one defined such as sensitivity and specificity have similar values, were estimated for each parameter. Compared to the first cut-off providing a very high specificity, the second cut-off improves the sensitivity. These two cut-off values were used here to estimate the sensitivity and the specificity of each automated procedure.

2.4 Automatic detection of abduction and adduction images

The detection of abduction and adduction images being a tedious task for the operator, requiring the careful observation of the cine loop, a second tool for detecting them automatically was proposed. Two different approaches for extreme images selection were defined and tested.

The first approach (PCA based method) was inspired from previous works done in our team: a method used to detect wall motion abnormalities in echocardiography (10) and a method to *a posteriori* gate ultrasound dynamic acquisitions according to breathing phases (11,12). The estimation was based on a principal component analysis (PCA) of the image sequence, which was applied to the entire set of time-intensity curves It(p), these curves being the intensity variations observed in a pixel, p, during the 10 seconds acquisition. Using appropriate linear combinations of the first two principal components, two time-intensity curves were then defined: the most constant curve $C_1(t)$ and a second curve $C_2(t)$ related to motion, and particularly to breathing motion. The time-intensity curve $C_2(t)$ was then filtered and stretched in order to find its local extreme values. Finally all pairs of extreme images, associating one minimal value and the next following maximal value (or one maximal value and the next

following minimal value) were defined (Figure 3). Using this automatic selection, different pairs could thus be tested and several cycles could be analyzed. However resulting parameters could vary from one cycle to another for physiological reasons (the motion amplitude could vary from one cycle to another) and for image quality reasons. To get rid of these sources of variability in this first feasibility study, the image pair the closest to the images I_C and I_O chosen by the expert was selected.

The second alternate approach (triangle area based method) that we tested was directly inspired from the previous one and from the tracking of the three landmarks T, L, and R by the optical flow approach described in section 2.2. Indeed, the three landmarks defined on the image I_C were then tracked by the optical flow algorithm image by image, both in the forward and backward directions, until reaching the extremities of the image sequence. From this tracking, the temporal curve showing the area of the triangle defined by the three landmarks could be computed. Then following the procedure described for the $C_2(t)$ curve in the previous paragraph, all couples of images corresponding to a local minimum and to a local maximum of the area were defined. Once again, the pair of images, which was the closest to the images I_C and I_O selected by the expert, was chosen.

3. RESULTS

3.1 Automatic tracking methods

For each tracking approach, the mean Euclidian distances and standard deviations (in pixels) between the three estimated landmarks and reference landmarks (defined by the expert) were calculated on the database. Figure 4 presents the resulting diagrams landmark by landmark.

These results show that the three tracking approaches have quite similar performances, showing higher distance values for the arytenoids (mean between 10.5 and 12.5 pixels) than for the top of the thyroid cartilage (mean between 7 and 8.5 pixels).

Defining a limit distance of ten pixels to validate the tracking procedure, the percentage of 'non acceptable' landmarks was obtained (Figure 5).

The backward method provided the lowest number of acceptable points: 77% for point T, 50% and 60% for points L and R. The most performant approach was the forward method with the lowest number of points out of the 10 pixels threshold. The forward and 'forward, increment 2' methods showed close results.

For each tracking method, sensitivity and specificity of recurrent nerve paralysis associated with parameters SR_O (Table 1) and DFM (Table 2) were calculated according to the two cut-off values presented in section 2.3.

Best results were obtained for the RS_O parameter, using the backward approach. Indeed, the sensitivity and specificity respectively reached 72% and 92% for the first cut-off value, and 84% and 73% for the second cut-off value. The forward method provided also satisfying results for this RS_O parameter. For the DFM parameter, the performances of the method showed higher sensitivity and lower specificity values with the backward method than with the forward method.

3.2 Automated selection of abduction and adduction images

It was difficult to compare the two methods of selection alone and to find an automatic way of quantifying it. Thus it was decided to study its impact on the final estimation of sensitivities and specificities for the detection of the recurrent nerve paralysis. To avoid asking the expert

to select the landmarks on each pair of selected images, we decided to use the automated tracking procedure, using either the forward or the backward approaches defined in section 2.2 for defining the landmarks location on the selected images. Tables 3 and 4 showed the results obtained for the SR_O and DFM parameters.

These results first proved that for both parameters, the second method of image selection (the one based on the temporal variations of triangle TLR area) was more relevant than the first method of image selection (the one based on the PCA approach). Indeed, the second method always presented both higher sensitivity values and higher specificity values. Moreover, comparing these results to those obtained for the follow-up of landmarks by the forward optical flow method, sensitivity and specificity values of the second method of image selection were close to that of the landmarks tracking method. The PCA based image selection method showed lower sensitivity and specificity values.

4. DISCUSSION

Ultrasound acquisitions of laryngeal tract have recently been proposed as a new non-invasive procedure to detect recurrent nerve paralysis (3–5,9). From the dynamic cine-loop, including about 300 images, a three-steps procedure was defined including: 1) an appropriate selection of two images corresponding to abduction and adduction phases; 2) a selection of three landmarks, the top of thyroid cartilage and the left and right arytenoids, on each selected image; 3) the computation of some geometrical indices (5). The first two steps require the intervention of an expert. To reduce the tediousness of these steps, the present study proposes to define an automated computer-aided diagnosis system. Three strategies for an automatic

tracking were tested and two methods for selecting the pair of abduction and adduction images were proposed.

The three tracking methods based on a multi-pyramidal optical flow algorithm showed that the arytenoids were more difficult to track than the top of the thyroid cartilage (Figures 4 and 5). These results could be explained by the image contrast variation in the arytenoids during the image sequence. Furthermore the motion amplitude is larger for the arytenoids than for the thyroid cartilage. It is known that the Lukas-Kanade algorithm used for tracking is less robust for large amplitude motion. Thus an alternative solution could be to track landmarks locally by a strategy based on a block-matching algorithm, and not globally on the whole field of view, as it was done in this study. Furthermore the backward method provides results that differ from those obtained by the forward method, which confirms the importance of testing the two directions to evaluate a tracking approach based on optical flow. The iterative procedure of tracking leads inevitably to an accumulation of errors in estimations, but they can be different in the forward and backward directions. Thus future work should investigate the automated combination of both approaches. Despite the reduced processing time brought by the two by two estimation strategy (about twenty seconds for fifty images on a standard workstation), this forward, increment 2 approach, appears less interesting, because its results are always worse than results obtained by the one by one forward method. Errors can be very important in case of motion with large amplitude due to swallowing or large respiratory motion. To avoid such a phenomenon, it should be helpful to exclude images with high amplitude motion. Another source of errors in the tracking of landmarks could come from the low quality of some B-mode images sequences, which are unavoidable when dealing with a large number of data, for which no *a priori* selection was done. However, despite these imperfections, the sensitivity and specificity values that were obtained for the SR_O parameter

were quite high. For these reasons, the best compromise would be to ask the operator to systematically draw the three landmarks on I_O images.

Results for the two automated methods for the detection of abduction and adduction images indicate that the method based on the variations of the larynx area (represented by the area of the triangle defined by the three landmarks) is more relevant than the method based on PCA algorithm. The possible reason for explaining such differences is that PCA synthesizes not only the breathing motion but also all sources of variation of intensities during the image sequence, thus increasing its sensitivity to noise. However, the second method requires an initial positioning of the landmarks. Furthermore its robustness was studied only in the neighborhood of the initially selected images. Its robustness when selecting other pairs of images in the sequence remains to be tested. Thus the decision process concerning the choice of either the selection based on PCA or the selection based on the triangle area remains to be established.

Being aware that a fully automatic procedure for landmarks positioning, tracking and extreme images selection cannot be 100% effective, a dedicated software component for making the diagnosis easier was developed. Briefly it proposes an automatic image selection while keeping for the operator the possibility to modify the selection of images. Then the operator is invited to position the three landmarks on the I_O image. The automatic tracking is then executed, offering the operator the possibility to modify the estimated position of the landmarks after tracking in case of large errors. Finally motion-based parameters including SR_O, SR_C, LFM and RFM are calculated to suggest the diagnosis: recurrent nerve paralysis or absence of paralysis. The resulting interface (Figure 6) has already been tested by final users. It now needs to be included in ultrasound devices in order to facilitate the diagnosis task.

5. CONCLUSION

An innovative procedure based on laryngeal ultrasound imaging and dedicated software defining specific image processing steps was proposed in order to detect recurrent nerve paralysis. Automatic methods for extreme images selection and landmarks tracking were defined and evaluated on a large database including 150 subjects. The performance of the automatic procedure was directly expressed in terms of sensitivity and specificity for the detection of recurrent nerve paralysis. A combined approach between operator and image processing algorithms provided a new platform dedicated to the diagnosis of recurrent nerve paralysis. This hybrid system will be evaluated in the next future by end users in terms of precision in the diagnosis, inter/intra operator variability, time of analysis, and ergonomics. Further developments could allow using this platform as software implemented in ultrasound imaging devices for clinical applications. Moreover some additional image processing procedures could be developed and tested to improve the automatic part of the system.

ACKNOWLEDGMENTS

The authors are grateful to Claire Barakat for her careful proofreading. They would like to thank the SATT Lutech for its financial support of the VOCALE project. H. Bergeret-Cassagne thanks the AFCE for its financial support during her Master Thesis.

6. REFERENCES

1. Duclos A, Lifante J-C, Ducarroz S, Soardo P, Colin C, Peix J-L. Influence of Intraoperative Neuromonitoring on Surgeons' Technique During Thyroidectomy. World J Surg. 2011;35(4):773–8.

2. Gardner GM, Smith MM, Yaremchuk KL, Peterson EL. The cost of vocal fold paralysis after thyroidectomy. The Laryngoscope. 2013;123(6):1455–63.

3. Wang C-P, Chen T-C, Yang T-L, Chen C-N, Lin C-F, Lou P-J, et al. Transcutaneous ultrasound for evaluation of vocal fold movement in patients with thyroid disease. Eur J Radiol. 2012 Mar;81(3):e288–91.

4. Dedecjus M, Adamczewski Z, Brzeziński J, Lewiński A. Real-time, high-resolution ultrasonography of the vocal folds--a prospective pilot study in patients before and after thyroidectomy. Langenbecks Arch Surg Dtsch Ges Für Chir. 2010;395(7):859–64.

5. Cohen ME, Lefort M, Bergeret-Cassagne H, Hachi S, Li A, Russ G et al. Quantification of vocal fold motion using echography: application to recurrent nerve paralysis detection Proc. SPIE 9414, Medical Imaging 2015: Computer-Aided Diagnosis, 94142M,1-6.

6. Hachi S. Détection et quantification du mouvement des cordes vocales sur des échographies cervicales. Master 2 Réseaux Image Parole. Université Paris Descartes; 2013.

7. Bouguet J. Pyramidal implementation of the Lucas Kanade feature tracker. Intel Corp Microprocess Res Labs. 2000;

8. Lucas BD, Kanade T. An Iterative Image Registration Technique with an Application to Stereo Vision. Proceedings of the 7th International Joint Conference on Artificial Intelligence - San Francisco, CA, USA: Morgan Kaufmann Publishers Inc.; 1981 p. 674–9.

9. Bergeret-Cassagne H. Modélisation échographique du larynx : étude de la mobilité des cordes vocales après chirurgie thyroïdienne. Evaluation de la faisabilité et de la sensibilité d'un nouvel outil diagnostic de paralysie récurrentielle. Master 2 Sciences Chirurgicales. Université Paris XII; 2013.

10. Diebold B, Delouche A, Abergel E, Raffoul H, Diebold H, Frouin F. Optimization of factor analysis of the left ventricle in echocardiography for detecting wall motion abnormalities. Ultrasound Med Biol. 2005;31(12):1597–606.

11. Renault G, Tranquart F, Perlbarg V, Bleuzen A, Herment A, Frouin F. A posteriori respiratory gating in contrast ultrasound for assessment of hepatic perfusion. Phys Med Biol. 2005;50(19):4465-80.

12. Mulé S, Kachenoura N, Lucidarme O, Oliveira AD, Pellot-Barakat C, Herment A, et al. An automatic respiratory gating method for the improvement of microcirculation evaluation: application to contrast-enhanced ultrasound studies of focal liver lesions. Phys Med Biol. 2011;56(16):5153-65.

Captions of figures

Figure 1: Ultrasound images of the laryngeal tract: the two images being the selected images I_C (image 66/300) and I_O (image 123/300) corresponding respectively to adduction and abduction phases. The three landmarks - at the top: thyroid cartilage (T), at the bottom: left (L) and right (R) arytenoid cartilages defined by the operator on each image are shown as red circles. The green circles are the results of the forward tracking of the three points defined on I_C, using the iterative optical flow approach from I_C till I_O.

Figure 2: Definition of the geometrical parameters from the three landmarks and the median line (shown in white dashes). These elements define the left and the right triangles of interest and their areas on images I_C and I_O.

Figure 3: Example of two temporal curves corresponding to: 1) the breathing motion component $C_2(t)$ issued from PCA-based procedure (in blue color); 2) the variation of the triangle area delimited by the three landmarks (in yellow color). The vertical lines correspond to the choice of images I_C and I_O performed by the expert.

Figure 4: Diagrams showing the mean Euclidian distances, expressed in pixels, (and their standard deviations) between the three estimated landmarks T, L and R by forward, backward, 'forward, increment 2' tracking methods and the selected landmarks by the expert.

Figure 5: Diagrams showing the percentage of estimated landmarks positioned at a Euclidian distance greater than 10 pixels from the expert positioning.

Figure 6: Interface of the CAD software developed for the project showing the selected pair of images: I_C (on the left) and I_O (on the right). The three landmarks are shown in red color before tracking and estimated landmarks after optical flow tracking are shown in yellow color.

Titles of Tables

Table 1: Sensitivity and specificity for detecting recurrent nerve paralysis obtained according to the different automatic tracking approaches and using the surface ratio on opening images (SR_O) parameter.

Table 2: Sensitivity and specificity for detecting recurrent nerve paralysis obtained according to the different automatic tracking approaches and using the difference of fraction of mobility (DFM) parameter.

Table 3: Sensitivity and specificity for detecting recurrent nerve paralysis obtained for the two methods of image selection that were tested (the one based on the PCA, and the one based on the temporal curve showing the area of the triangle constituted by the three landmarks), using the surface ratio on opening images (SR_O) parameter.

Table 4: Sensitivity and specificity for detecting recurrent nerve paralysis obtained for the two methods of image selection that were tested (the one based on the PCA, and the one based on the temporal curve showing the area of the triangle constituted by the three landmarks), using the difference of fraction of mobility (DFM) parameter.

Table 1

Tracking method	Cut-off value	Sensitivity (%)	Specificity (%)
forward	0.1257	66	77
backward		72	92
forward, increment 2		66	73
forward	0.077	80	59
backward		84	73
forward, increment 2		76	60

Table 2

Tracking method	Cut-off value	Sensitivity (%)	Specificity (%)
forward	0.2059	56	82
backward		61	74
forward, increment 2		54	80
forward	0.112	74	66
backward		80	55
forward, increment 2		72	49

Table 3

Images selection method	Cut-off value	Sensitivity (%)	Specificity (%)
PCA based	0.1257	58	68
Triangle area based		62	76
PCA based	0.077	70	42
Triangle area based		76	60

Table 4

Images selection method	Cut-off value	Sensitivity (%)	Specificity (%)
PCA based	0.2059	50	75
Triangle area based		58	81
PCA based	0.112	76	52
Triangle area based		74	59

Figure 1

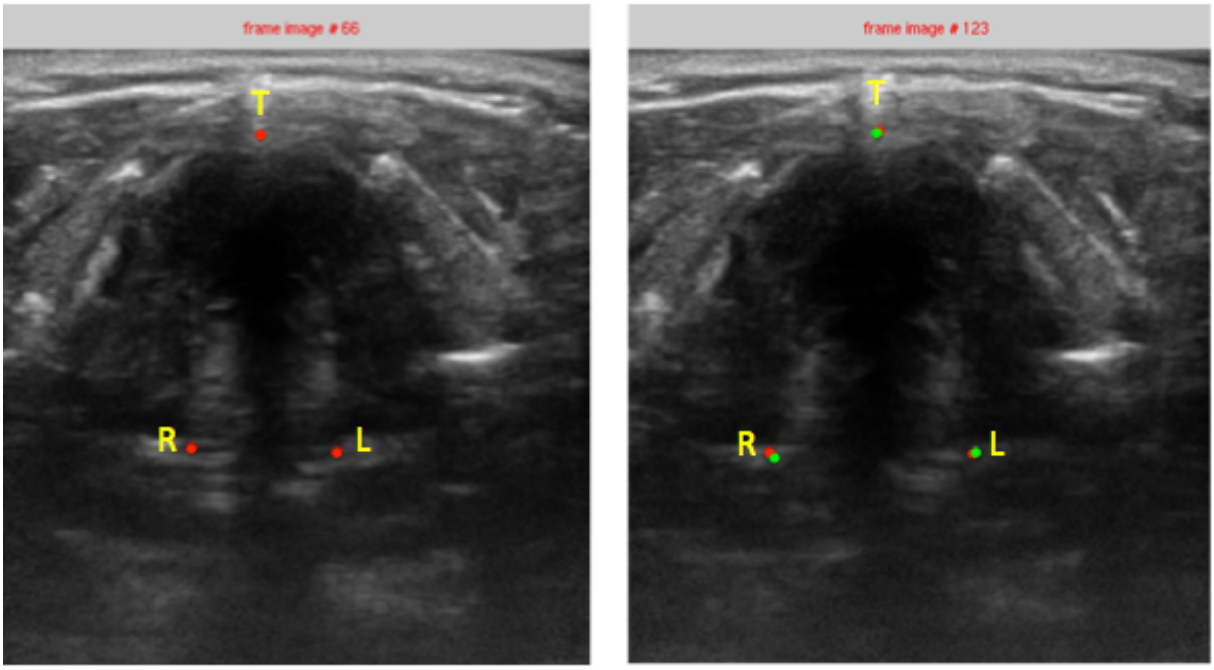


Figure 2

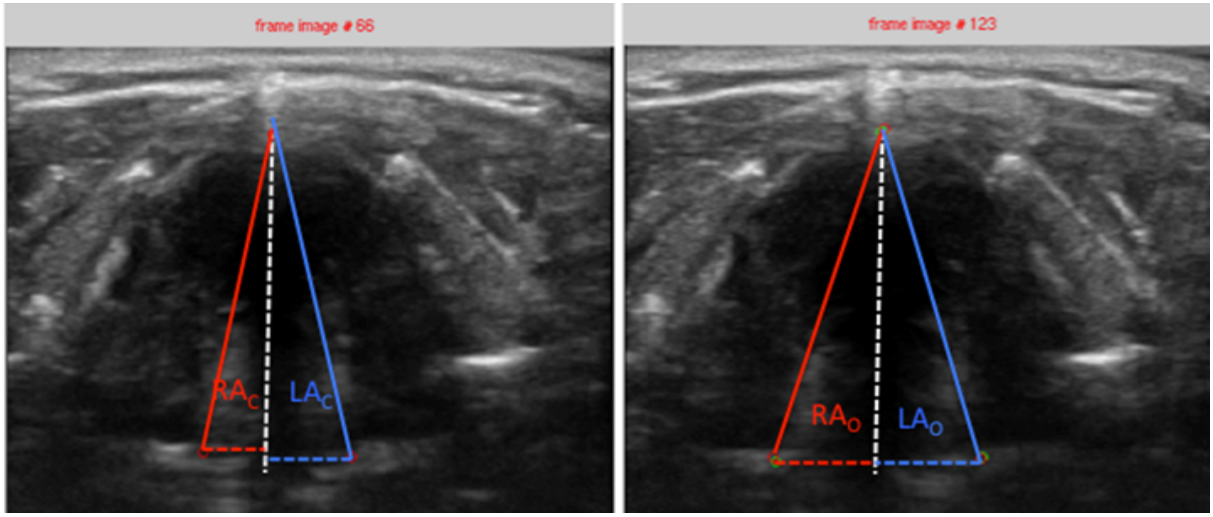


Figure 3

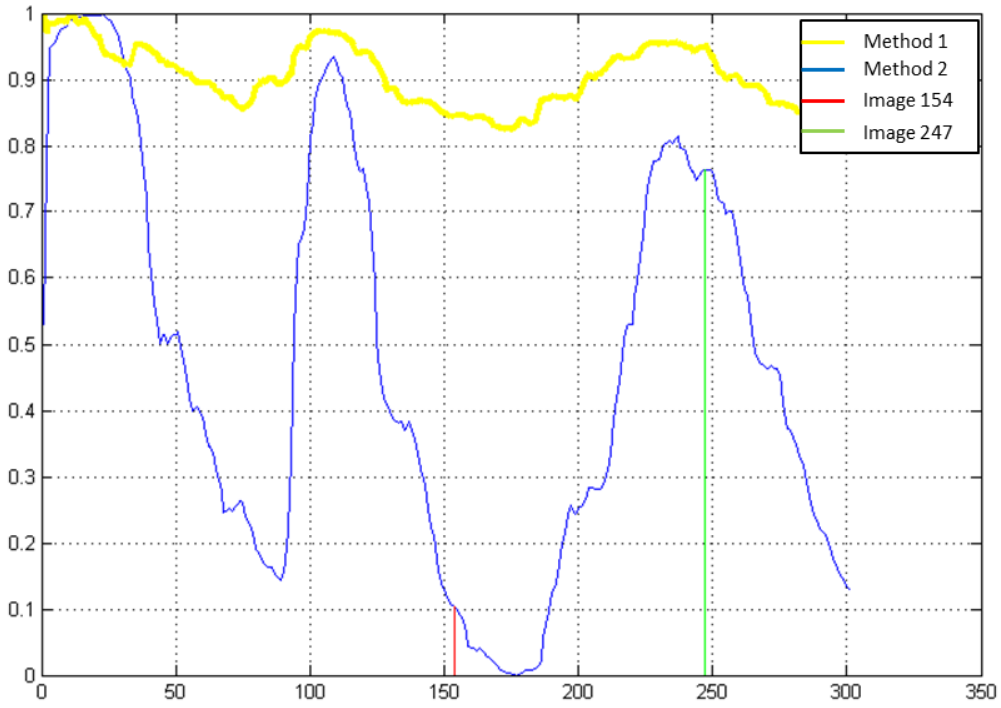

Method 1
Method 2
Image 154
Image 247
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1
0
50
100
150
200
250
300
350

Figure 4

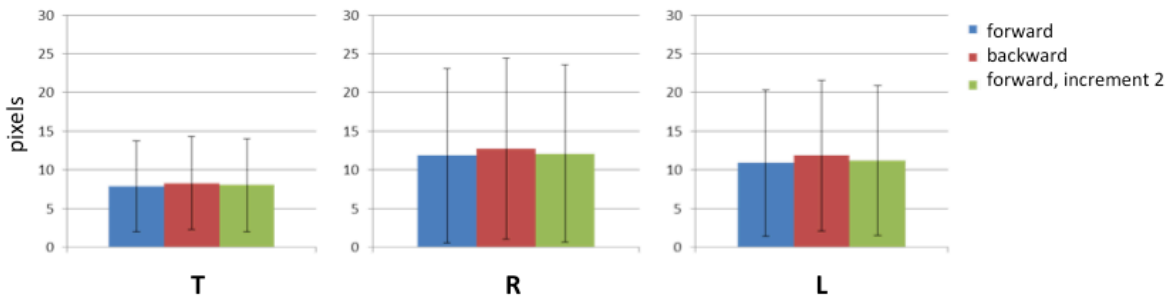

pixels
30
25
20
15
10
5
0
T
R
L
forward
backward
forward, increment 2

Figure 5

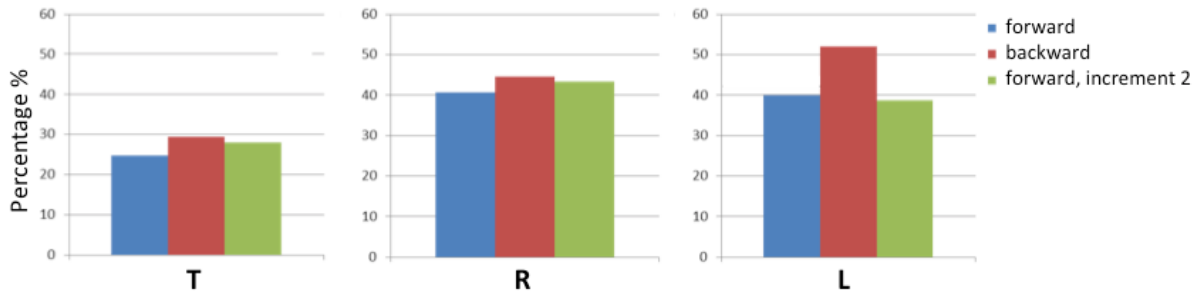
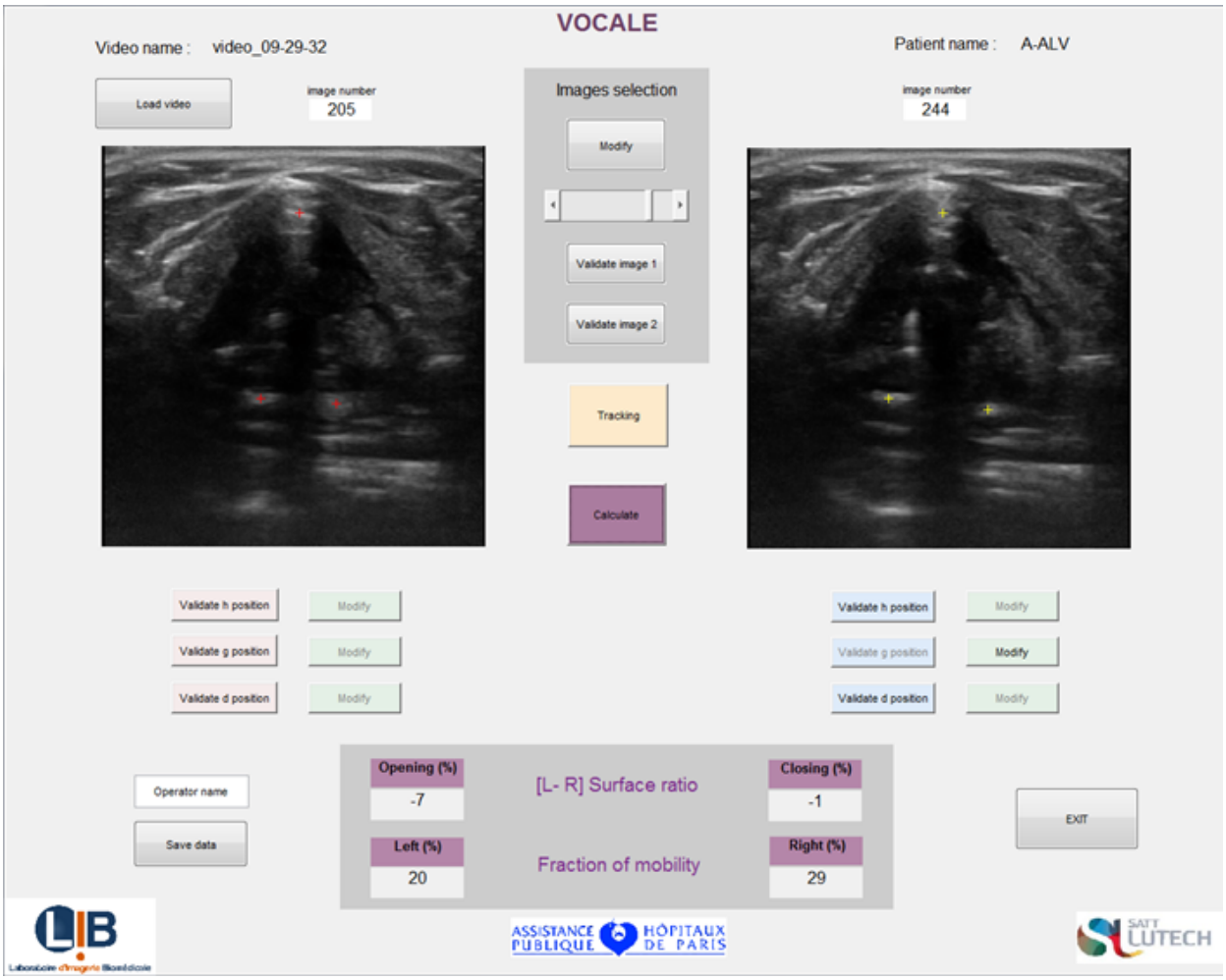

Percentage %
60
50
40
30
20
10
0
T
R
L
forward
backward
forward, increment 2

Figure 6


VOCALE
Video name : video_09-29-32
Patient name : A-ALV
Load video
image number
205
Images selection
Modify
Validate image 1
Validate image 2
Tracking
Calculate
image number
244
Validate h position
Modify
Validate g position
Modify
Validate d position
Modify
Validate h position
Modify
Validate g position
Modify
Validate d position
Modify
Operator name
Save data
Opening (%)
-7
[L- R] Surface ratio
Closing (%)
-1
Left (%)
20
Fraction of mobility
Right (%)
29
EXIT
ASSISTANCE PUBLIQUE HÔPITAUX DE PARIS
SATT LUTECH