

Quantification of vocal fold motion using echography: application to recurrent nerve paralysis detection

Mike-Ely Cohen^{a,b}, Muriel Lefort^a, Héloïse Bergeret-Cassagne^{a,c}, Siham Hachi^{a,d}, Ang Li^a,
Gilles Russ^e, Diane Lazard^c, Fabrice Menegaux^c, Laurence Leenhardt^{a,e},
Christophe Trésallet^c, Frédérique Frouin^{a,f}

^a Sorbonne Université, Univ Paris 06, Inserm, CNRS, Laboratoire d'Imagerie Biomédicale, Paris, France ^b Satt Lutech, Paris, France ^c Service de Chirurgie Générale, Viscérale, Endocrinienne, CHU Pitié-Salpêtrière, AP-HP, Univ Paris 06, Paris, France ^d Université du Luxembourg, Luxembourg Centre for Systems Biomedicine, Esch/Alzette ^e Service de Médecine Nucléaire, CHU Pitié-Salpêtrière, AP-HP, Paris, France, ^f Inserm/CEA/Université Paris Sud/CNRS, CEA/I2BM/SHFJ, Laboratoire IMIV, Orsay, France

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 conventional way to detect RP. We suggest a new approach based on laryngeal ultrasound and a specific data analysis was designed to help with the automated detection of RP.

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

Results provided by automated image processing were compared to those obtained by an expert. Detection of extreme images was accurate; tracking of landmarks was reliable in 80% of cases. Motion symmetry indices showed similar values for controls and patients without RP. Fraction of mobility was reduced in cases of RP. Thus, our CAD system helped in the detection of RP.

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: ultrasound, vocal fold motion, recurrent nerve paralysis, optical flow

1. INTRODUCTION

Thyroid surgery concerns several tens of thousands of people every year in France (<http://stats.atih.sante.fr>), and is performed in similar proportion 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 procedure. 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.

The 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. Thus this procedure is not systematically used but proposed to patients presenting some symptoms of RP after surgery.

*frederique.frouin@inserm.fr; phone 33 1 69 86 77 39; fax 33 1 53 82 84 48; lib.upmc.fr

Ultrasound has already been proposed to study laryngeal anatomy. Furthermore, the dynamics was 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. As it is currently done for the nasal endoscopy and as it was proposed by previous studies in ultrasound, the analysis of dynamic acquisitions was first based on the visual assessment of the vocal folds mobility. On ultrasound acquisitions, the vocal folds appear as isoechoic or hypoechoic structures, while the thyroid cartilage and arytenoids, on which vocal folds are attached, are hyperechoic structures, better defined on ultrasound images. Thus it was proposed to choose the thyroid cartilage and the arytenoids as surrogate markers to study the vocal fold motion. Dedicated software was first developed to study the above-mentioned landmarks on the dynamic acquisition. In its first version, the expert selected two images in the sequence corresponding to abduction and adduction of vocal folds, thus showing their extreme positions during motion. Then he/she positioned three landmarks (thyroid cartilage and left and right arytenoids) on each image and several geometrical indices were finally computed and tested to detect RP.

The present study proposes to develop a second CAD system to assist the operator in the selection of the abduction and adduction images and of the landmarks. An automated detection of images corresponding to abduction and adduction of vocal folds was proposed. Furthermore a frame-by-frame 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 preliminary database including four control subjects, three patients (after thyroid surgery) with RP and three patients (after thyroid surgery) without RP. Different geometrical indices indicating fractions of mobility, and motion symmetry were compared.

2. METHODS

2.1 Patients and Data acquisition

Four control subjects (without any symptoms) and six patients (after thyroid surgery) were considered for this preliminary task. All selected patients had post-surgery voice disorders and were examined by nasal endoscopy. A paralysis of the recurrent nerve was established for three of these six patients.

All subjects underwent an ultrasound acquisition in our institution, 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, providing an axial view. The acquisition protocol was based on the acquisition of a 10 seconds B-mode image sequence (with a frame rate of 30 images per second), asking the subject to breath freely. The operator was asked to keep the probe still, in order to neglect this possible source of motion. During breathing, vocal folds and arytenoids showed successive phases of opening and closing.

2.2 Detection of abduction I_O and adduction I_C images

The first selection was visual: an expert was asked to select a couple of extreme images, by reviewing the ten seconds image sequence, using a dedicated cine loop viewer. This was done off-line and produced two images per acquisition, I_C and I_O .

A second method, more automated, was proposed in order to assist the operator in this tedious task. This second approach was based on a Singular Value Decomposition (SVD), which was applied to the set of P time-intensity curves $I_i(p)$; these time-intensity curves corresponded to the grey-level variations associated to each pixel p ($1 \leq p \leq P$) according to the 10 seconds acquisition (t being the number of images of the ultrasound sequence, $1 \leq t \leq T$). If \mathbf{I} is the (P,T) matrix associated with these P curves, the SVD factorization of \mathbf{I} was given by $\mathbf{U}\mathbf{\Sigma}\mathbf{V}^*$, \mathbf{U} being an orthogonal (P,P) matrix, \mathbf{V} an orthogonal (T,T) matrix and $\mathbf{\Sigma}$ a diagonal (P,T) matrix of non negative values, in descending order, called singular values. Considering the two orthogonal vectors $\mathbf{V}_1(t)$ and $\mathbf{V}_2(t)$ associated with the two largest singular values, two normalized curves $K_1(t)$ and $K_2(t)$ with non negative values were then estimated as linear combinations of $\mathbf{V}_1(t)$ and $\mathbf{V}_2(t)$ such as $K_1(t)$ was the most constant curve and its area under the curve was equal to 1, and $K_2(t)$ was a curve with 0 as minimal value and its area under the curve was equal to 1. Such conditions could be fulfilled without ubiquity and provided a unique solution. This procedure was inspired from previous approaches developed to detect wall motion abnormalities in echocardiography [5, 6] and breathing motion in contrast ultrasound [7, 8]. The second curve $K_2(t)$ was associated with main variations of intensity observed in the series. Assuming that global motion was negligible and the mean grey levels in the different structures did not vary with time, the curve could be associated with breathing motion,

as it had been associated with cardiac motion in echocardiography [5,6] or breathing motion in contrast ultrasound studies [7,8]. Finally the time-intensity curve related to breathing motion was stretched [7] in order to find its local extreme values and possible couples of extreme images (Figure 1) were then defined. For this feasibility study, the couple of images, which was the closest to the selection made by the operator, were selected.

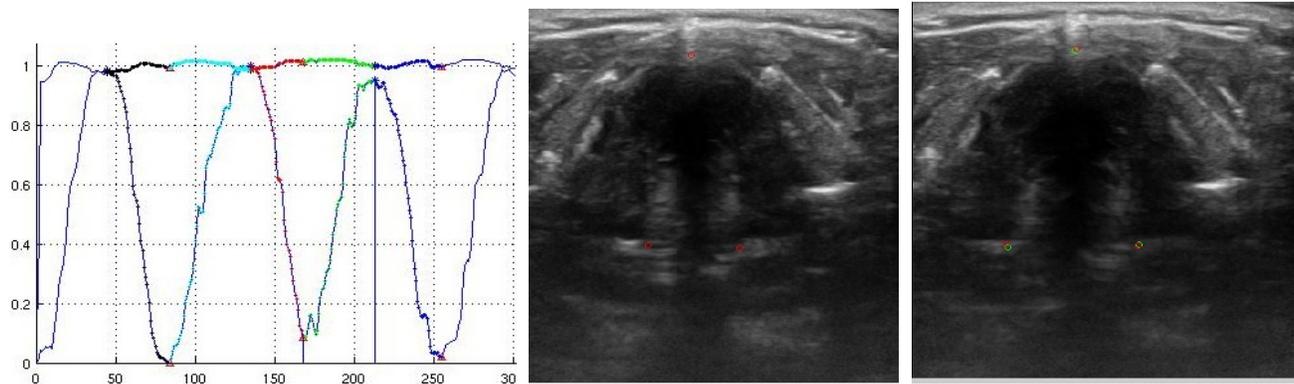


Figure 1 Automated selection of adduction and abduction images. From left to right: time signal-intensity curves $K_1(t)$ and $K_2(t)$ estimating one nearly constant curve and one curve related to motion, colored parts of the curves show the different possible intervals between two successive extreme positions, vertical lines show the two extreme images that were selected by the expert; associated adduction (I_C) and abduction (I_O) images.

2.3 Definition and tracking of landmarks

As mentioned in the introduction, we assumed that arytenoid cartilages and their apparent motion in the plane could be surrogate markers for studying vocal folds motion. Furthermore, the top of the thyroid cartilage was also considered since it was a motionless point during the examination and it enabled to define the symmetric plane and to consider the two halves of the larynx. To evaluate motion quantitatively, we had to define the amplitude of motion. Furthermore we consider indices related to the symmetry of motion. Indeed they could be good predictors of abnormalities since the paralysis affects one side generally (in rare cases both sides are affected).

In a first study, an expert defined three landmarks corresponding to the top of the thyroid cartilage (T), and to the centers of the left (L) and right (R) arytenoid cartilages on the two images I_C and I_O , as shown by Figure 2.

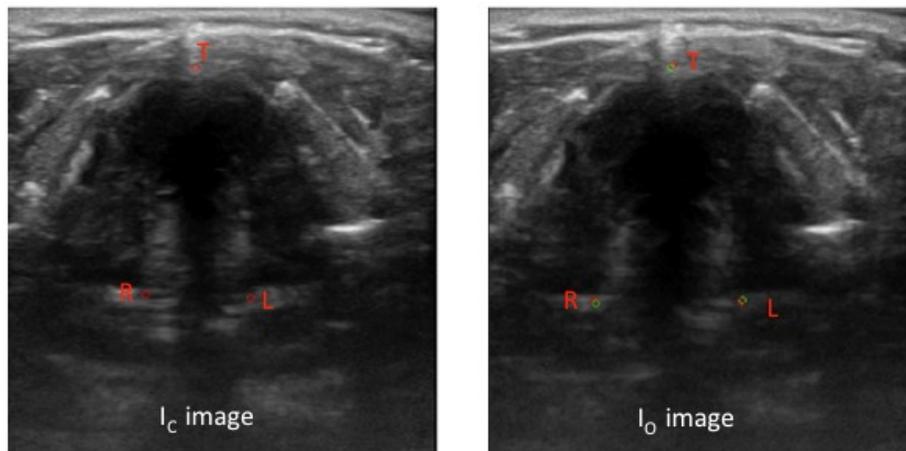


Figure 2: the three landmarks (T: thyroid cartilage, R: right arytenoid cartilage, L: left arytenoid cartilage) defined by the expert are shown in red color on both I_C and I_O images. The points in green color are the results of the tracking of these three points first defined on I_C image, using the optical flow approach from I_C till I_O .

To make the procedure more robust, we proposed to define landmarks on images I_C and an iterative optical flow approach based on a robust pyramidal implementation [9] of the Lucas-Kanade-Tomasi algorithm was defined for each pair of consecutive images going from I_C to I_{C+1} , I_{C+1} to I_{C+2} , ... till I_{O-1} to I_O . An additional constraint was to define a priori the parameters of the algorithm in order to avoid the user to set them. We chose to have the same values for all the images and for all the acquisitions: two pyramid levels, a neighborhood of 20x20 pixels (assuming that inside the neighborhood, the pixels have a constant displacement) and a number of iterations equal to 15.

2.4 Definition of motion based features

Using the three landmarks, the areas of the left and right triangles were defined on both I_C and I_O images yielding LA_C , RA_C , LA_O , RA_O (Figure 3).

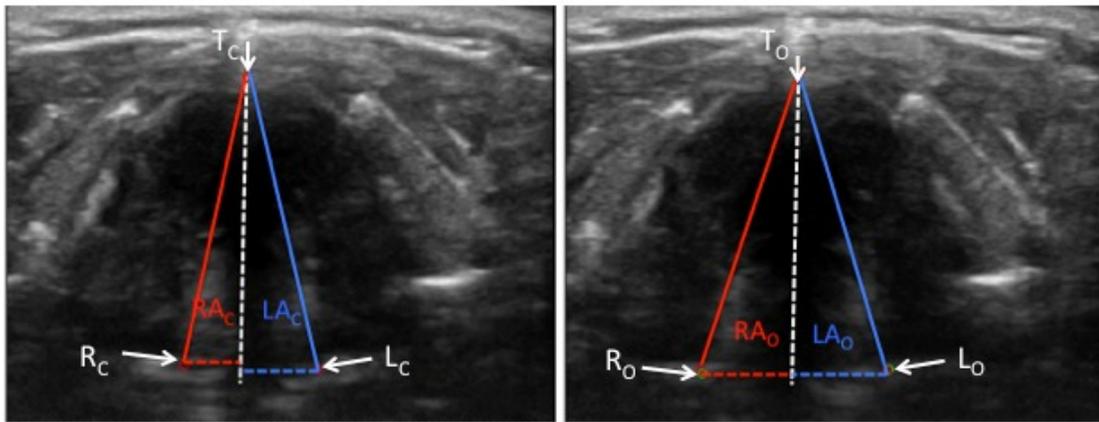


Figure 3: Definition of triangles of interest on I_C and I_O images and estimation of associated areas: LA_C , RA_C , LA_O , and RA_O .

A first parameter was defined by considering the laryngeal symmetry (SYM). It was computed on images I_O (SYM_O) or I_C (SYM_C), SYM_O being equal to $|LA_O - RA_O| / (LA_O + RA_O)$, and SYM_C being equal to $|LA_C - RA_C| / (LA_C + RA_C)$. SYM_O was expected to be close to 0 (positive or negative) in case of symmetrical motion. Its absolute value was expected to be higher in case of dissymmetry, indicating a suspicion of recurrent nerve paralysis.

Then to quantify the motion amplitude, left and right fractions of mobility LFM and RFM were defined between images I_O and I_C as following: $LFM = (LA_O - LA_C) / LA_O$ and $RFM = (RA_O - RA_C) / RA_O$. These fractions were expected to be high for normal cases and to be reduced (low positive values and possibly negative values) in case of recurrent nerve paralysis.

3. RESULTS

Table 1 shows the parameters SYM_O , SYM_C , LFM , and RFM for three settings S1, S2, and S3: S1) the expert selected the images I_C and I_O and the three points of interest on both images I_C and I_O ; S2) the expert selected the images I_C and I_O and the three points of interest on images I_C (respectively I_O , depending on the position of these two images in the sequence), the three points on I_O (respectively I_C) were defined by the tracking approach based on the optical flow algorithm; S3) the images I_C and I_O were defined according to the procedure defined in section 2.2, and the three points of interest were defined as described in the previous setting S2.

From these results, a first analysis was done using the values that were obtained by the expert (first setting). It appears that SYM_O could isolate two out of three PR cases which had much larger values (>0.30) than values obtained for control and NPR subjects (<0.15). The SYM_C did not appear to be discriminant even if smaller values were generally obtained for control than for patients. For all cases, the LFM was high (values between 0.20 and 0.60,) while RFM showed low values (less than 0.1) for the three patients having PR. The diagnosis of right recurrent nerve paralysis was confirmed for these three subjects. However one NPR subject showed also a low value of RFM .

Considering the expert as the reference, results obtained by the optical flow tracking task (second setting) were quite encouraging showing similar results (similar being defined as providing the same classification) in 9 out of 10 cases. However a visual inspection of T, L, and R points revealed that the tracking could be improved. Results obtained by the third setting combining automated selection of adduction and abduction images and the automated tracking task were less concordant with the expert (two different cases for SYM₀ and three different cases for LFM and RFM)

Subjects	SYM ₀ -S1	SYM ₀ -S2	SYM ₀ -S3	SYM _c -S1	SYM _c -S2	SYM _c -S3	LFM-S1	LFM-S2	LFM-S3	RFM-S1	RFM-S2	RFM-S3
C#1	0.01	0.04	0.06	0.05	0.05	0.01	0.41	0.36	0.37	0.47	0.46	0.45
C#2	0.08	0.03	0.07	0.06	0.06	0.03	0.21	0.24	0.07	0.40	0.37	0.15
C#3	0.03	0.01	0.0	0.07	0.07	0.11	0.48	0.38	0.20	0.37	0.27	0.36
C#4	0.06	0.06	0.44	0.10	0.07	0.07	0.44	0.28	0.01	0.39	0.44	0.55
NPR#1	0.0	0.41	0.25	0.38	0.38	0.43	0.57	0.21	0.62	0.03	0.26	0.42
NPR#2	0.11	0.14	0.09	0.22	0.22	0.06	0.38	0.40	0.61	0.50	0.49	0.58
NPR#3	0.04	0.01	0.06	0.13	0.13	0.07	0.29	0.37	0.02	0.13	0.15	0.04
PR#1	0.34	0.37	0.26	0.47	0.47	0.38	0.31	0.31	0.29	0.06	0.10	0.07
PR#2	0.0	0.04	0.03	0.09	0.09	0.04	0.21	0.15	0.17	0.03	0.07	0.06
PR#3	0.37	0.13	0.24	0.15	0.15	0.06	0.54	0.37	0.29	-0.34	-0.12	-0.32

Table 1: Quantitative parameters according to the three configurations S1, S2 and S3. Considering S1 as the reference, numbers in red color indicate large deviations of the results obtained by automated methods compared with the reference, suggesting a different diagnosis.

4. DISCUSSION

Dynamic ultrasound acquisitions of laryngeal tract have been proposed to detect recurrent nerve paralysis. 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 geometrical indices in order to suggest the diagnosis of recurrent nerve paralysis or to discard it. To reduce the tediousness of these steps, we proposed to test some automatic methods in order to define a computer-aided diagnosis system.

Three settings were tested on a preliminary database including 10 subjects: a first setting including a fully manual processing, a second setting using an automatic tracking of the three landmarks, a third setting using an automatic tracking of the three landmarks and an automated selection of adduction and abduction images.

The visual analysis of the tracking done with the multi-pyramidal optical flow algorithm showed better results for the top of the thyroid cartilage than for the arytenoids. These results seem to be related with the amplitude of vocal fold motion, which is larger for the arytenoids than for the thyroid cartilage. Indeed, despite the pyramidal approach, the Lukas-Kanade algorithm is known to be less robust for large amplitude motion. As only forward tracking was tested, it could be useful to test the backward direction too. Furthermore the iterative procedure of tracking led inevitably to an accumulation of errors in estimations, but they could be different in the forward and backward directions. Thus future work should investigate the automated combination of both approaches. 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 clinical data, for which no a priori selection was done. However, despite these imperfections, the values that were obtained for the SYM₀, LFM and RFM parameters were quite similar to the values that were obtained by the expert.

The last setting combining both the automated selection of adduction and abduction images and the tracking of the landmarks shows to be less efficient. However in 7 cases out of 10 cases, it provides quite similar results. Being aware that a fully automatic procedure for landmarks positioning, tracking and image selection cannot be fully effective, a

dedicated software component for making the diagnosis easier was developed. 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_0 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 quantitative parameters are calculated in order to suggest the diagnosis: recurrent nerve paralysis or absence of paralysis.

The small number of cases enrolled in this study makes the identification of relevant parameters complex. Our next goal is to increase the number of studies on our clinical database in order to better define criteria to detect recurrent nerve paralysis. Thereafter a multi-parametric approach combining both fractions of mobility and symmetry indices could be tested to improve the overall classification rate.

5. CONCLUSIONS

A new procedure based on laryngeal ultrasound imaging and specific image processing steps is presented to detect recurrent nerve paralysis. Image processing steps provide encouraging results when compared to those obtained by an expert. The results obtained on a preliminary database of ten subjects show the potential of symmetry and fraction of mobility indices to help in the automated diagnosis of recurrent nerve paralysis. An approach combining expert decision and image processing algorithms could provide a powerful procedure and help in the automated diagnosis of this pathology.

REFERENCES

- [1] Duclos, A., Lifante, J.C., Ducarroz, S., Soardo, P., Colin, C., and Peix, J.L. "Influence of Intraoperative Neuromonitoring on Surgeons' Technique During Thyroidectomy," *World J Surg.* 35(4):773–778 (2011).
- [2] Gardner, G.M., Smith, M.M., Yaremchuk, K.L., and Peterson E.L. "The cost of vocal fold paralysis after thyroidectomy," *The Laryngoscope* 123(6):1455–63 (2013).
- [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.* 81(3):e288–e291 (2012).
- [4] Dedejusz, M., Adamczewski, Z., Brzeziński, J., and 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* 395(7):859–864 (2010).
- [5] Frouin, F., Delouche, A., Raffoul, H., Diebold, H., Abergel, E., and Diebold, B. "Factor analysis of the left ventricle by echocardiography (FALVE): a new tool for detecting regional wall motion abnormalities," *Eur J Echocardiogr.* 5(5):335-46 (2004).
- [6] Diebold, B., Delouche, A., Abergel, E., Raffoul, H., Diebold, H., and Frouin, F. "Optimization of factor analysis of the left ventricle in echocardiography for detecting wall motion abnormalities," *Ultrasound Med Biol.* 31(12):1597–1606 (2005).
- [7] Renault, G., Tranquart, F., Perlberg, V., Bleuzen, A., Herment, A., and Frouin, F. "A posteriori respiratory gating in contrast ultrasound for assessment of hepatic perfusion," *Phys Med Biol.* 50(19):4465-80 (2005).
- [8] Mulé, S., Kachenoura, N., Lucidarme, O., Oliveira, A.D., Pellot-Barakat, C., Herment, A., and Frouin, F. "An automatic respiratory gating method for the improvement of microcirculation evaluation: application to contrast-enhanced ultrasound studies of focal liver lesions," *Phys Med Biol.* 56(16):5153-5165 (2011).
- [9] Bouguet, J. "Pyramidal implementation of the Lucas Kanade feature tracker," Intel Corporation, Microprocessor Research Labs, Technical reports, OpenCV Documents (1999).