Diagnosis of New Onset Vocal Cord Paralysis Using Acoustic Analysis

Yu Wei
College of Communication Engineering
Chongqing University, Chongqing 400044, China
wayne@yn133165.com

Hamid Gholam Hosseini
Institute of Biomedical Technologies (IBTec), Auckland University of Technology
Private page 92006, Auckland 1142 - New Zealand
hgholamh@aut.ac.nz

Andrew Cameron
Middlemore Hospital, Private Bag 93311, Otahuhu,
Auckland 1640, New Zealand
andrewj.cameron@middlemore.co.nz

Michael J. Harrison
University of Auckland,
Auckland, New Zealand
michael.harrison@auckland.ac.nz

Ahmed Al-Jumaily
Institute of Biomedical Technologies (IBTec), Auckland University of Technology
Private page 92006, Auckland 1142 - New Zealand
Ahmed Al-Jumailyn@ aut.ac.nz

Abstract

Vocal cord paralysis can occur as a complication of surgery or anaesthesia, and if permanent is a significant clinical problem. Early detection is important to optimize the chance of repair, and to avoid complications associated with an impaired swallow. We have developed an algorithm to detect altered vocal cord function by comparing pre and post-operative signals. It is based on Wavelet Packet Analysis (WPA) and Support Vector Machines (SVM). The coefficients of WPA are combined with other parameters, such as Spectral Centroid, Spectral roll-off point etc., as the feature vector. This method is compared to the Hoarseness Diagram method (HDm), which was reported as an objective voice quality evaluation approach and can be used for pathological voice discrimination. Although previous authors have analysed pathological voices, we believe that our algorithm is unique in that samples were obtained from the same patients before and after their surgery. Our experiments using voice signals recorded from subjects before and after the procedure show a high classification accuracy, whereas HDm fails in the detection of a hoarse voice. Based on our findings, we are working to develop a screening tool to detect damage to vocal structures that arises during surgery.

1. Introduction

Vocal cord paralysis caused by permanent damage to the recurrent laryngeal nerve is a highly significant clinical problem which can arise during surgery of the head and neck, most notably during thyroid surgery. Clinically, this causes hoarseness (a breathy, raspy voice) and an inability to safely swallow thin fluids. Despite the importance of early detection, assessment of recurrent laryngeal nerve function postoperatively is poorly standardized, and sometimes completely neglected. Some doctors directly visualize the vocal cords during phonation on emergence from anaesthesia. This is achieved by passing an endoscope or mirror into the patient's pharynx – an invasive technique which can be uncomfortable and requires sterile equipment, clinical expertise and a high level of patient compliance. Some clinicians listen for vocal change in an attempt to detect recurrent laryngeal nerve pathology. Whilst non-invasive, this is a highly subjective approach which relies on the observer's recollection of voice quality preoperatively.

Acoustic analysis-based techniques have been shown to be an effective screening tool for laryngeal dysfunction. Despite this, we are unaware of any previous use of acoustic analysis techniques in this clinical context. Our objective is to develop a non-invasive diagnostic tool to detect new onset vocal cord paralysis.

Brachial plexus blockade is a medical technique whereby local anaesthetic agents are injected around the nerves that supply the shoulder and arm. This eliminates the ability to feel or move the upper limb for several hours, and is commonly used to provide pain relief for patients undergoing surgery. A common side effect is hoarseness, which has been ascribed to a transient, unilateral loss of vocal cord function. This occurs because the recurrent laryngeal nerve lies close to the nerves supplying the arm, and can also be affected by the local anaesthetic [1]. Although the incidence of hoarseness appears to vary depending on the volume of local anaesthetic injected, rates of up to 50% have been reported in some groups [2]. Unlike nerve damage that arises as a consequence of surgery, this paralysis is not permanent, and therefore is not a clinically significant problem. However, the high frequency of vocal cord dysfunction among patients who have undergone brachial plexus blockade means that they represent an excellent study population.

Acoustic measures extracted from voice signals include the short-term perturbations of fundamental frequency and intensity (called jitter and shimmer, respectively), and glottal noise measures. Research to date has failed to yield a consensus on the utility of the various perturbation measures either for discriminating between normal and pathological voice samples (include hoarseness). Eskenazi [3] found that the percent jitter parameter exhibited a correlation of 0.55 with perceptual ratings of breathiness. Shrivastav [4, 5] reported a much better correlation of 0.86 for the percent jitter parameter with ratings of breathiness. However, Martin [6, 7] found very poor correlation between the jitter parameter and breathiness ratings. In addition, Parsa [8, 9] found that the jitter parameter resulted in a classification accuracy of only 68% on a database of 53 normal and 175 pathological talkers. A plausible reason for the mixed reports on the efficacy of perturbation measures is that they are dependent on the accuracy of fundamental frequency extraction, which is difficult on severely disordered voices. Several authors have investigated more global measures of abnormal voice quality. Yumoto [10, 11] introduced the Harmonics-to-Noise Ratio (HNR) parameter which quantifies the amount of glottal noise in the vowel waveform, and showed that it can be an effective indicator of pathological voice quality.

The best acoustic features for screening the voice signals would be those with the lowest correlation against the others and with the best discrimination capabilities [12, 13]. Among these lines, the GNE (Glottal-to-Noise Excitation Ratio), as reported, provides a low correlation with respect to the amplitude perturbation and noise features. Moreover, an advantage of this parameter is that its calculation is not based on a previous estimation of the fundamental frequency, a difficult task in the presence of pathology. Based on GNE (a noise measure) and other three aperiodicity measures: Period Perturbation Quotient (PPQ, a jitter measure), Energy Perturbation Quotient (EPQ, a shimmer measure) and Mean Waveform Matching Coefficient (CORR, a period correlation measure), Hoarseness Diagram method (HDm) was proposed by Michaelis [14-16] as a new approach to describe different acoustic properties of voices, and later studies showed that different pathological voices were localized in different areas in the diagram. To date, this method (HDm) has only been used in voice quality studies to represent the ‘‘hoarseness diagram,’’ but there are no existing extensive studies about the validity of this method for hoarseness screening purposes in related literature.

2. Data Collection

The sound data for this work were recorded in the Hospital for Special Surgery, New York. Data were collected before and after the operation by anesthesiologists. Following informed consent the patients (13 females and 20 males, aged 20 to 70 years old) were asked to speak 8 different words and each word was uttered 2 times. Patients with a hoarse voice preoperatively were excluded from the study. Patients then underwent insertion of a brachial plexus block, followed by shoulder surgery. No breathing tubes or other foreign bodies were placed in the patient’s airways during surgery. Postoperatively, patients were clinically assessed to detect the new onset of hoarseness. Patients were classified as hoarse if both they and an independent observer assessed their voice as being raspy or altered. The recorded signals were in WAV format and were categorized into two classes: pre-operative and post operative. A total of 66 data sets (33 pre and 33 post) were recorded using a digital multitrack recorder in stereo, 8-bit sampling resolution, 44,100 Hz sampling rate and 705kbps bit rate.

Typically, acoustic analysis is performed on the sustained vowel sound (1 to 3 seconds long). We choose the “yee” frame of each recorded file. The duration of every “yee” frame is from 0.8 to 2.8 seconds, which meets the requirements of sustained vowel phonation.

3. Classification Methods

3.1. Hoarseness Diagram method (HDm)

The hoarseness diagram method allows a quantitative two-dimensional description and graphical representation of voice characteristics based on four acoustic measures [14]. These four measures were found to yield a low-dimensional description of an originally 21-dimensional acoustic data space with the least loss of information [16-18]. Principal components analysis revealed that in this four-dimensional space the data were distributed in approximately a two-dimensional plane. The two principal components describing this plane provided the basis for the two axes of the HDm.

3.1.1. Four acoustic features of HDm

In the Hoarseness Diagram, three out of four measures (jitter, shimmer, mean period correlation) contribute to the horizontal axis labelled the irregularity component (IC) and the forth one contributes to the vertical coordinate. These four measures are:

1. Energy Perturbation Quotient (EPQ, similar to APQ, a shimmer measure).

$$PQ = \frac{100\%}{N-K} \sum_{v=\frac{K-1}{2}}^{N-\frac{K-1}{2}-1} \left| \frac{u(v) - \frac{1}{K}\sum_{k=-\frac{K-1}{2}}^{\frac{K-1}{2}} u(v+k)}{\frac{1}{K}\sum_{k=-\frac{K-1}{2}}^{\frac{K-1}{2}} u(v+k)} \right| \qquad (1)$$

   In this equation, U(n) is the energy sequence defined by the sequence of glottal cycle lengths (i.e., the summed squares of the samples of the i-th glottal cycle define the i-th energy) averaging over k=15 periods [14, 16].

2. Pitch Perturbation Quotient (PPQ, a jitter measure).

   PPQ uses the sequence of glottal cycle lengths for u(n), averaging over k=3 periods in above equation (1) [14, 16].

3. Mean Correlation Coefficient between successive cycles (CORR).

   CORR is defined as the mean of all correlation coefficients for every pair of consecutive periods. It indicates the overall similarity between the cycles of the time signal. Its upper limit of 1 is reached for signals with identical period shapes (i.e. strictly periodic signals) and it decreases with increasing differences in length or shape between consecutive periods.

   The above three parameters all measure the aperiodicity of voice signal and they enter the x-coordinate (IC) of the hoarseness diagram. The forth parameter is GNE which

4. Glottal-to-Noise Excitation Ratio (GNE, a noise parameter).

   The algorithm to calculate the GNE was first proposed by Michaelis [14-16]. The GNE represents an approach to quantify the amount of excitation due to vocal fold oscillations versus the excitation given by turbulent noise. Thus, it is closely related to breathiness, and it is considered a reliable measure for the relative noise level even in the presence of strong amplitude and frequency perturbations. It has an advantage that its calculation is not based on a previous estimation of the fundamental frequency [8-16].

   GNE indicates to what extent the voice excitation is due to a pulse train or due to noise. In tests using synthetic signals the GNE was shown to be sensitive to additive noise but unlike the normalized noise energy NNE [17, 18] or cepstral harmonics to noise ratio CHNR [19, 20] that are often used to assess additive noise in speech signals to be independent of jitter and shimmer [16]. The GNE enters the hoarseness diagram by the vertical coordinate labelled the noise component (NC) [14].

3.1.2. The coordinates definition of HDm

As mentioned in last section, EPQ, PPQ and CORR enter the x-coordinate (IC) of HDm, and GNE enters y-coordinate (NC). The coordinates are calculated by the following equations:

$$x-coordinate = 5 + \frac{1}{\sqrt{3}}\left(\frac{CORR+1.614}{0.574} + \frac{PPQ+0.374}{0.645} + \frac{EPQ-0.757}{0.368}\right) \qquad (2)$$

$$y-coordinate = 1.5 + \frac{0.695-(1-10^{GNE})}{0.242} \qquad (3)$$

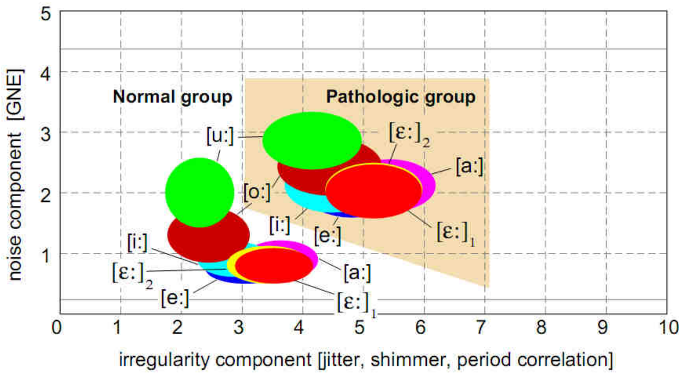

Figure 1 - Illustration of Hoarseness Diagram method using different vowels to discriminate pathologic group

Figure 1 illustrates using HDm with different vowel phonations to discriminate pathological voices from normal voices [14], but keep in mind that the pathologic group did not include hoarseness voices caused by anaesthesia.

It is our goal to screen hoarseness between post operative voice signals. We will explore the hoarseness discrimination capability of HDm and our proposed algorithm in the next section.

3.2. The proposed method

Our proposed method is based on Wavelet Packet Analysis (WPA) and Support Vector Machine (SVM). The goal of the present paper is to discriminate between hoarseness and nonhoarseness voices by extracting efficient features. WPA has been chosen for this task because of its superior time-frequency resolution, adaptivity and suitability for pattern recognition applications. After we got signal features through WPA, these features were input to a SVM classifier to attain the discrimination goal.

3.2.1. Wavelet Packet Analysis (WPA)

Wavelet transform, including discrete wavelet packet transform (DWPT), are generally acknowledged to be useful for studying nonstationary signals. It was developed as an alternative to the short time Fourier Transform (STFT) to overcome problems related to its frequency and time resolution properties. More specifically, unlike the STFT that provides uniform time resolution for all frequencies the WT (wavelet transform) provides high time resolution and low frequency resolution for high frequencies and high frequency resolution and low time resolution for low frequencies. In that respect it is similar to the human ear which exhibits similar time-frequency resolution characteristics. They have shown to be of particular value in the perceptually tuned analysis and parameterization of speech signals [21-23].

In the wavelet packet framework, the transform idea is exactly the same as those developed in the wavelet framework. The only difference is that wavelet packets offer a more complex and flexible analysis, because in wavelet packet analysis, the details as well as the approximations are split.

The extracted wavelet packet coefficients provide a compact representation that shows the energy distribution of the signal in time and frequency. In order to further reduce the dimensionality of the extracted feature vectors, statistics over the set of the wavelet coefficients are used.

The following wavelet-domain features are used in our algorithm:

1. WPT coefficients mean (WPCM)

   WPT coefficients mean is the median value of the coefficients in each subband. These features provide information about the frequency distribution of the voice signal.

2. WPT coefficients SD (WPCSD)

   WPT coefficients SD compute the standard deviation of the coefficients in each subband. These features provide information about the amount of change of the frequency distribution.

3.2.2. Support Vector Machine (SVM)

A Support Vector Machine (SVM) performs classification by constructing an *N*-dimensional hyperplane that optimally separates data into two categories. SVM models are closely related to neural networks. In fact, a SVM model using a sigmoid kernel function is equivalent to a two-layer, perceptron neural network. Using a kernel function, SVM's are an alternative training method for polynomial, radial basis function and multi-layer perceptron classifiers in which the weights of the network are found by solving a quadratic programming problem with linear constraints, rather than by solving a non-convex, unconstrained minimization problem as in standard neural network training [24, 25].

The goal of SVM modelling is to find the optimal hyperplane that separates clusters of vector in such a way that cases with one category of the target variable are on one side of the plane and cases with the other category are on the other side of the plane. All vectors lying on one side of the hyperplane are labelled as −1, and all vectors lying on the other side are labelled as 1. The training instances that lie closest to the hyperplane are called support vectors. More generally, SVM projects the original training data in space X to a higher dimensional feature space F via a Mercer kernel operator.

For median-sized problems, cross validation might be the most reliable way for parameter selection [24, 25]. First, the training data is separated to several folds. Sequentially a fold is considered as the validation set and the rest are for training. The average of accuracy on predicting the validation sets is the cross validation accuracy. Our SVM implementation, using k-level (k=5) cross validation method, can be summarized as follows:

Divide the sample sets (feature vector sets) into k disjoint equal-sized subsets or folds.

1. Construct a regressive model through these k-1 fold data and initialized parameter. In fact, this process is the training process. Use the last fold to validate the performance of this model based on MSE criterion.
2. Repeat (1) (2) k times, so that every fold will have the chance to be the validation fold, and get the optimized parameter. Use the optimized parameter to train the whole training set and generate the final model and k-level average classification rate.

In our experiments, C-SVM [26, 27] method was used with the upper bound of C=200.

4. Experimental Results

The results are presented in two sections. The first section deals with the discriminative capabilities of the HDm for the detection of hoarseness. The second section deals with the discrimination accuracy of our proposed algorithm.

4.1. HDm results

The voice recordings were analyzed in 500ms frames applying a shift of 300ms. For each complete frame (i.e., generally excluding the last frame of a vowel utterance) the hoarseness diagram coordinates were calculated. Table 1 shows some calculated results of the four acoustic features and then the IC (x-coordinate of HDm) and NC (y-coordinate of HDm) according to equations (1) (2) (3). These results are based on post-operative signals, not including pre-operative signals.

Table 1 - Acoustic features of some post-operative signals (subject no.1~10)

Signal (post-op)	EPQ	PPQ	GNE	CORR	IC (x-axis)	NC (y-axis)
Post01	11.924	0.147	0.957	0.95	5.409	0.419
Post02	7.285	4.989	0.539	0.992	5.62	2.144
Post03	4.715	0.603	0.48	0.78	5.97	2.387
Post04	19.799	5.068	0.863	0.85	7.608	0.806
Post05	12.876	3.106	0.678	0.819	7.205	1.571
Post06	19.217	12.198	0.666	0.879	7.835	1.621
Post07	4.334	0.776	0.913	0.969	5.16	0.599
Post08	11.373	0.98	0.256	0.95	6.107	3.314
Post09	26.281	6.378	0.411	0.805	8.003	2.674
Post10	7.039	0.149	0.724	0.886	5.413	1.382

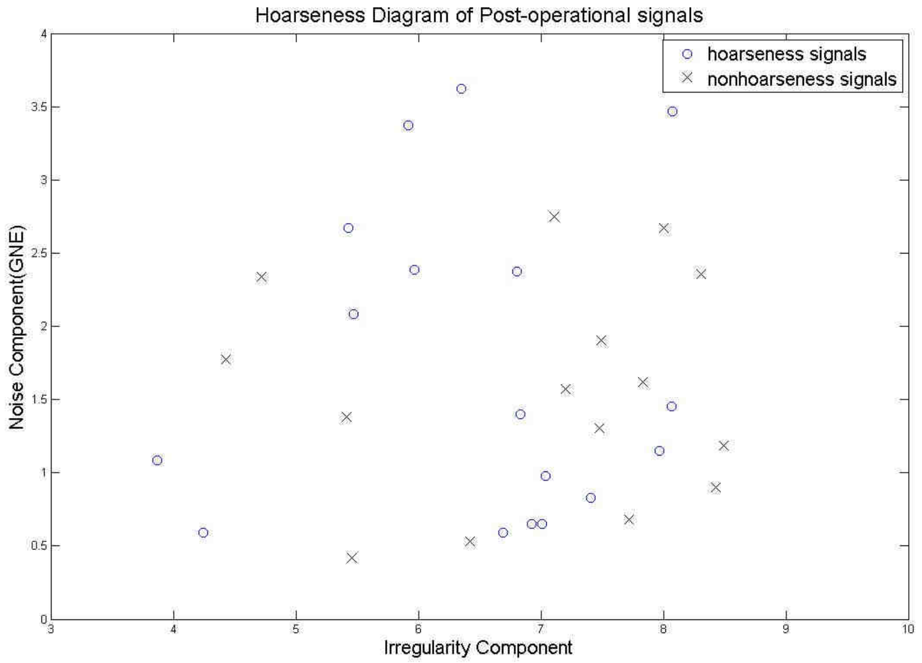

Figure 2 - Hoarseness Diagram based on post-operative signals

Then we can get the Hoarseness Diagram based on all 33 post-operative signals. Figure 2 shows that the two classes of signals (hoarseness and nonhoarseness) are not localized in separate areas therefore, it possesses poor classification result. We also calculated IC and NC based on the hoarseness-patient group (17 patients) whose voices changed to hoarseness after anaesthetic procedure. Table 2 shows the results.

Table 2 - IC and NC of hoarseness-patient group

Signal (pre-op)	**IC (x-axis)**	**NC (y-axis)**	**Signal (post-op)**	**IC (x-axis)**	**NC (y-axis)**
pre02	4.312	2.213	post02	5.47	2.086
pre03	8.788	2.588	post03	5.97	2.387
pre04	6.692	0.715	post04	7.407	0.832
pre07	5.935	1.281	post07	4.244	0.593
pre08	5.844	4.069	post08	6.353	3.625
pre11	5.666	1.127	post11	5.921	3.372
pre17	8.589	1.593	post17	7.037	0.978
pre18	6.483	1.028	post18	7.968	1.151
pre19	5.743	0.869	post19	5.432	2.672
pre20	3.471	2.213	post20	3.867	1.084
pre22	6.701	3.108	post22	8.076	3.468
pre23	6.223	0.773	post23	8.068	1.451
pre26	5.782	0.989	post26	6.808	2.374
pre27	4.199	2.081	post27	6.927	0.653
pre28	7.38	2.044	post28	6.836	1.4
pre29	6.03	3.138	post29	6.689	0.592
pre30	7.206	3.505	post30	7.011	0.648

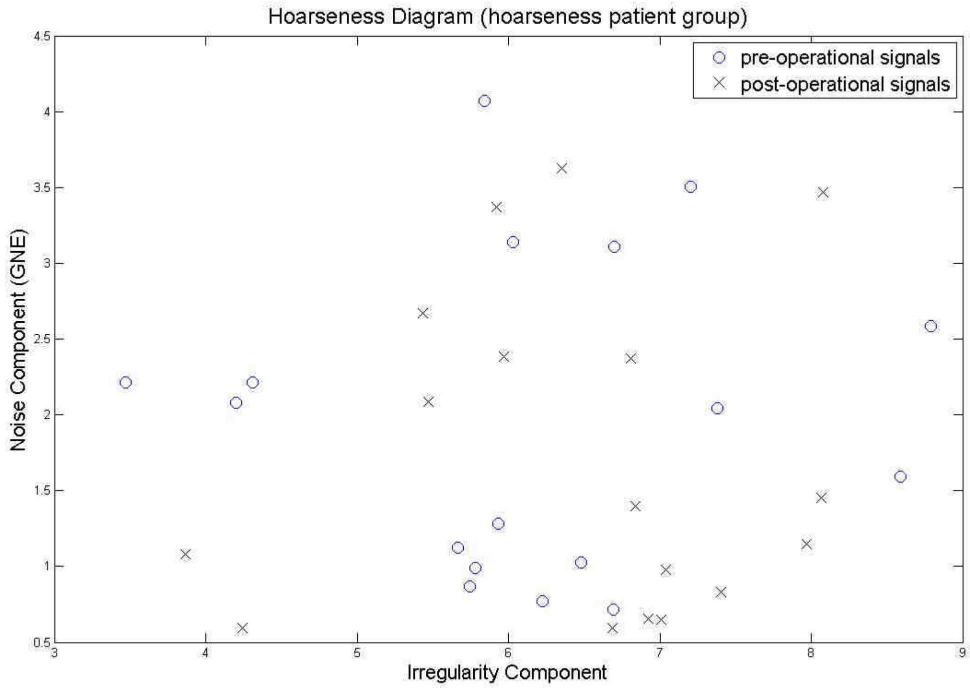

Figure 3 - Hoarseness Diagram based on hoarseness-patient group which includes pre and post-operative signals

Figure 3 shows the Hoarseness Diagram for patient group (17 patients) which includes pre and post-operative signals. Just as the last Hoarseness Diagram showed, this Hoarseness Diagram also can not classify signals into two categories: hoarseness and nonhoarseness.

4.2. The proposed algorithm results

We use a 4-level Haar-wavelet to decompose the signal and divide the voice band (0~8000Hz) into 16 subbands, and get 16 subbands of coefficients. The generated feature vectors using WPA then were input into in the SVM classifier.

Classification accuracy of the proposed algorithm in detecting new onset hoarseness is summarised below:

- Based on post-operative signals from 33 patients: 82.5%
- Based on pre and post-operative signals from 17 patients who developed hoarseness: 93.5%

5. Discussion and Conclusions

In this paper we have presented our study on voice signal classification in the context of a medically induced vocal cord pathology. To date, there are no studies on this topic. However, discrimination of pre and post anaesthesia voice signals has a significant potential to enhance patient care by allowing doctors to non-invasively detect damage to vocal structures that can occur during surgery.

The hoarseness diagram method was evaluated using our study population and was ineffective in classifying our recorded voice signals into hoarse and non-hoarse groups. By contrast our algorithm, based on WVA and SVM, was highly concordant with clinical assessments.

This finding appears to contradict reports by Michaelis [14], where the Hoarseness Diagram was explored to analyze the voice quality and different pathological voice signals caused by different laryngeal dysfunctions. However, both results can be considered complementary, and the choice of these two methods depends on the targeted pathology. The HDm has been shown to be appropriate for evaluation of voice quality and to detect other pathologies, such as polyps, nodules, paralysis etc, [28, 29]. Our proposed algorithm targets a specific pathology; namely hoarseness arising as a consequence of recurrent laryngeal nerve palsy, and is preferred in this limited but important clinical context. Experimental evaluation has shown that the proposed classification algorithm based on SVM is very effective and has an accuracy of over 80% when compared to clinical assessments.

A possible reason for the inability of HDm to screen for hoarseness in this context is the signals we used. All previous tests of HDm were based on stable, sustained vowel phonation either recorded from 'wide-awake' subjects in a controlled environment, or simply drawn from a commercial database [30]. Such data is poorly generalisable to our clinical context. For reasons related to patient safety, post-surgical patient cannot be left in a sound-proofed, isolated environment, and accordingly our raw data was recorded in a busy hospital setting. The pitch and loudness of the post-surgical patient's phonation varied substantially in comparison to the signals used in HDm. It is not surprising that the proposed method is superior to HDm under these circumstances because WPA is a nonstationary analysis technique.

In the future we will introduce more data into the proposed algorithm to evaluate its validity. We will also focus on detection of hoarseness using continuous speech voice signals because sustained vowel signals do not incorporate such important vocal function attributes as rapid voice onset and termination, variations in voice fundamental frequency and amplitude, and voice breaks. These attributes may be highly relevant to the perception of vocal quality in everyday life as they make a vital contribution to the perceptual judgment of the person talking.

6. References

[1] Boezaart AP, de Beer JF, du Toit C, van Rooyen K. A new technique of continuous interscalene nerve block. Can J Anesth. 1999; 46(3): 275-81.

[2] Singelyn FJ, Seguy S, Gouverneur JM. Interscalene brachial plexus analgesia after open shoulder surgery: continuous vs patient controlled infusion. Anesthesia and Analgesia. 1999; 89: 1216-20.

[3] Eskenazi L, Childers DG, and Hicks DM. Acoustics correlates of vocal quality. J. Speech Hear. Res. 1990; 33: 298–306.

[4] Shrivastav R. The use of an auditory model in predicting perceptual ratings of breathy voice quality. J. Voice. 2003; 17(4): 502–512.

[5] Yamaguchia H, Shrivastavc R, Andrewsc ML, Niimib S. A Comparison of Voice Quality Ratings Made by Japanese and American Listeners Using the GRBAS Scale. Folia Phoniatr Logop. 2003; 55: 147-157.

[6] Martin D, Fitch J, and Wolfe V. Pathologic voice type and the acoustic prediction of severity. J. Speech Hear. Res. 1995; 38: 765–771.

[7] Beranek LL, Martin D. Acoustical Measurements. J. Acoust. Soc. Am. 1989; 86(4): 1625-1626.

[8] Parsa V, and Jamieson DG. Acoustic discrimination of pathological voice: Sustained vowels versus continuous speech. J. Speech, Language Hear. Res. 2001; 44: 327–339.

[9] Parsa V, Jamieson DG. Identification of pathological voices using glottal noise measures. Speech Lang Hear Res. 2000; 43: 469–485.

[10] Yumoto E, Gould W, and Baer T. Harmonics-to-noise ratio as an index of the degree of hoarseness. J. Acoust. Soc. Am., 1982; 71: 1544–1549.

[11] Yumoto E, Sasaki Y, Okamura H. Harmonics-to-Noise Ratio and Psychophysical Measurement of the Degree of Hoarseness. J. Speech Hear. Res.1984; 27: 2-6.

[12] Scheirer DE, Slaney M. Construction and evaluation of a robust multifeature speech/music discriminator. IEEE Transactions on Acoustics, Speech and Signal Processing. 1997; 1331-1334.

[13] Preciado JA, Ferna S. Digital analysis of the acoustic signal in vocal pathology diagnosis: Sensitivity and specificity of shimmer and jitter measurements. Acta Otorrinolaringol Esp. 1998; 49: 475–481.

[14] Michaelis D, Strube HW. Selection and combination of acoustic features for the description of pathologic voices. J. Acoust. Soc. Am. 1998; 103(3): 1628–1639.

[15] Michaelis D, Strube HW. Acoustic breathiness measures in the description of pathologic voices. In: Proceedings ICASSP 98, Seattle, WA;1998. 937–940.

[16] Michaelis D, Gramss T, Strube HW. Glottal-to-Noise Excitation Ratio–a new measure for describing pathological voices. Acustica/Acta acustica. 1997; 83: 700–706.

[17] Kasuya H, Ogawa S, Kikuchi Y, Ebihara S. An acoustic analysis of pathological voice and its application to the evaluation of laryngeal pathology. Speech Commun. 1986; 5:171–181.

[18] Kasuya H, Ogawa S, Kikuchi Y Ebihara S. Normalized noise energy as an acoustic measure to evaluate pathologic voice. J. Acoust. Soc. Am. 1986; 80: 1329–1334.

[19] de Krom G. A cepstrum-based technique for determining a harmonics-to-noise ratio in speech signals, J. Speech Hear. 1993; 36: 254–266.

[20] de Krom G. Some spectral correlates of pathological breathy and rough voice quality for different types of vowel fragments. J. Speech Hear. 1995; 38: 794–811.

[21] Strang G, Nguyen T. Wavelets and filter banks, Wellesley: Wellesley-Cambridge Academic Press, 1997; 94–107.

[22] Mallat SG. A wavelet tour of signal processing, San Diego: Academic Press, 1999.

[23] Williams JR, Amaratunga K. Introduction to wavelets in engineering. Int. Journal for Numerical Methods in Engineering. 1994; 37: 2365–2388.

[24] Suykens JAK, J. Vandewalle J. Least squares support vector machine classifiers. Neural Process. Lett. 1999; 9 (3): 293–300.

[25] Suykens JAK, Van Gestel T, de Brabanter J, De Moor B, Vandewalle J. Least squares support vector machine classifiers. World Scientific 2002; Singapore.

[26] Cortes C. Vapnik V. Support-vector network. Machine Learning. 1995; 20: 273–297.

[27] Vapnik V. Statistical Learning Theory, New York, NY: Wiley, 1998.

[28] Alvarez Gonzalez T, Cervera T, Miralles JL. Acoustic analysis of speech: Reliability of a set of acoustic multidimensional parameters. Multidimensionales Acta Otorrinolaringol Esp. 2002; 53: 256–268.

[29] Oppenheim AV, Schafer RW, Buck JR. Discrete-Time Signal Processing, 2 edn. Saddle River, NJ: Prentice Hall, 1999. 221–229.

[30] Kay Elemetrics Corp. Disordered Voice Database. Version 1.0, 3-CD-ROM. Lincoln Park, NJ: Kay Elemetrics Corp, 1994.