91

Inverse Filtering of Speech Signal for Detection of Vocal Fold Paralysis After Thyroidectomy

Andrius RYBAKOVAS[1], Virgilijus BEIŠA[1], Kęstutis STRUPAS[1],
Jonas KAUKĖNAS[2], Gintautas TAMULEVIČIUS[2]*

[1]*Vilnius University Faculty of Medicine Institute of Clinical Medicine Clinic of Gastroenterology,
Nephrourology and Surgery, Santariškių 2, LT-08661, Vilnius, Lithuania*
[2]*Vilnius University Institute of Data Science and Digital Technologies
Akademijos 4, LT-08663, Vilnius, Lithuania*
e-mail: gintautas.tamulevicius@mii.vu.lt

Abstract. The Autoregressive model-based digital inverse filtering technique is applied in non-invasive detection of vocal fold paralysis. The vocal tract filter is modelled using variable order (up to 20) AR model which is adequate to individual characteristics of human vocal properties. This postulates the more accurate estimation of the glottal flow, disturbances of which are direct evidence of the vocal fold paralysis.

Key words: inverse filtering, autoregressive model, speech analysis, vocal fold paralysis.

1. Introduction

Clinically, vocal fold paralysis (immobility) is detected using invasive techniques like laryngoscopy, kymography, and others. These techniques mean unpleasant procedure with the possible traumatic output, the need for expensive clinical equipment.

As an alternative to invasive techniques, acoustic signal analysis-based non-invasive techniques are explored extensively during the last two decades. Various parametric and non-parametric analysis techniques were proposed for assessment of vocal fold immobility type and degree.

In this paper, we present the Autoregressive (AR) model-based digital inverse filtering approach for estimation of the glottal flow. The quality of estimated flow is evaluated using prediction error which is used as an objective indicator of the vocal fold functionality. Experimental analysis of the proposed technique was performed using recordings of healthy and pathological voices. The results obtained show the ability of the inverse filtering technique to characterize the quality of the glottal flow and make it possible to detect the paralysis of the vocal folds.

---

*Corresponding author.

2. The Background

2.1. *Vocal Fold Paralysis*

Voice and speech have very important roles in human social life and professional performance. The negative impact of laryngeal nerve injury on voice is well known in thyroid surgery, but unfortunately, the correlation between them is little studied. The literature shows that altered voice is a common problem after thyroid surgery. The voice changes were reported in 25% to almost 90% of patients within the first few weeks after thyroidectomy (Henry *et al.*, 2010). Other studies represent similar numbers (30–87%) (de Pedro Netto *et al.*, 2006; Musholt *et al.*, 2006; Stojadinovic *et al.*, 2002; Page *et al.*, 2007; Sinagra *et al.*, 2004; Elsheikh *et al.*, 2016). Voice changes can be classified as neural and non-neural related. The true incidence of recurrent laryngeal nerve injury following thyroid surgery is probably underrated, as it strongly depends on postoperative laryngeal examination. According to a systematic review (Jeannon *et al.*, 2009), which included 27 articles and 25,000 patients, the average of temporary incidence of recurrent laryngeal nerve after thyroid operation was 9.8% and the incidence of permanent injury of the same nerve was 2.3%. The rate varied from 26% to 2.3%. The data of 3,605 patients from 5 high-volume centres in France (Lifante *et al.*, 2017) shows similar results: immediate injury rate was 9.3% (range 3.8–21.8%), permanent rate was 3.1% (0–9.1%). The Scandinavian multicentre audit of 3,660 patients reports postoperative unilateral paresis of the recurrent laryngeal nerve in 3.9% of cases (Bergenfelz *et al.*, 2008). It is very important to realize that vocal cord paralysis may occur without any voice changes. Voice could be normal in case of vocal cord paralysis in up to 28% of cases (Mihai and Randolph, 2009) or even in more than 50% (Ortega *et al.*, 2009). Majority of endocrine and general surgeons agree that pre- and postoperative laryngoscopy should be mandatory in all patients undergoing thyroid surgery, as it is the most trustworthy method in determining vocal cord paralysis. Despite reliability of this method, it could be uncomfortable and unpleasant for the patient, adds extra costs, needs special instruments and trained personal, causes logistic problems (Ortega *et al.*, 2009). Probably computerized acoustic voice analysis could be used as a screening method to select patients for laryngoscopic examination.

2.2. *Acoustic Speech Analysis for Voice Disorders*

The idea to apply acoustical analysis of speech for voice disorder detection and evaluation is not new. Similar ideas were proposed 50–60 years ago (Lieberman, 1963; Koike, 1967), and has been studied since now. Various acoustic parameters were proposed and employed for this purpose. These include but are not limited to perturbations of fundamental tone (Kasuya *et al.*, 1983), various noise estimation techniques (Yumoto *et al.*, 1982; Kasuya *et al.*, 1986; Fukazawa *et al.*, 1988), cepstral features (Dejonckere and Wieneke, 1994; Hillenbrand *et al.*, 1994), nonlinear operators and techniques (Cairns *et al.*, 1994; Giovanni *et al.*, 1999), MFCC features (Dibazar *at al.*, 2002), fractal dimensions (Baljekar and Patil, 2012; Ali *et al.*, 2016). During the last decade the task of acoustic analysis-based

detection and evaluation of pathological voices was studied intensively. Vast majority of studies focus on combining various features without any physiological reasoning. Extensive and summarized reviews on acoustic analysis of pathological voice can be found in Arroyave *et al.* (2012), Vaičiukynas *et al.* (2015), Panek *et al.* (2015).

The speech signal is generated in two stages. Firstly, the so-called source signal is induced. The air flow generated by the lung causes the vibration of the vocal folds. This vibration is called phonation process, and its intensity is described by the fundamental frequency value. In the next step, the glottal flow is modulated by the voice tract. The result of this modulation is the speech signal, transmitting information on both the vocal fold and the voice tract resonant properties. Disorder of vocal folds (paralysis among them) affects the speech inevitably. The effect depends on dysfunction degree of the folds and can vary from inaudible changes up to severe changes of voice, for example, it becomes breathy, harsh, and weak.

Acoustical analysis of the speech signal is considered as an objective evaluation of the vocal tract functionality rather than perceptual analysis of the speech. Acoustic parameters represent generative and articulatory properties of the voice and thus could be applied for pathology detection and evaluation. Different acoustic parameters describe different stages of the speech signal production, thus should be chosen reasonably. To estimate the functionality (or immobility) of vocal folds, we have to analyse the glottal flow.

2.3. *Inverse Filtering Technique*

The most common technique to estimate the glottal flow is to employ *source-filter* production model. This model describes the speech signal as the convolution of the source signal (glottal flow) and a filter (vocal tract). Both source signal and vocal tract can be modelled using various joint estimation models or separately, ignoring or considering close phase of the glottal cycle (Walker and Murphy, 2007; Alku, 2011).

If we consider the glottal flow and the vocal tract as independent, the glottal flow can be extracted by inverse filtering of the speech signal (Alku, 2011). The inverse filter eliminates the effect of the vocal tract thus giving the estimate of the glottal flow. The process of inverse filtering can be simplified using linear modelling of the vocal tract.

Linear modelling has played a very important role in speech analysis domain because of its mathematical tractability and applicability, spectral estimation properties. For speech analysis purposes, the linear all-pole filter was applied mostly. Various linear prediction techniques were employed for glottal flow extraction: constrained linear prediction with reduced distortion of filter frequency response (Alku and Magi, 2009), weighted linear prediction with temporal weighting of the residual (Airaksinen *et al.*, 2014) and its stabilized modification (Kafentzis *et al.*, 2011).

All voice pathology detection and inverse filtering studies can be summarized as follows:

- The prediction model order varies from 8 up to 12 in different studies. The order number is related with the number of modelled vocal tract formant frequencies, p-th order model describes $p/2$ formants. Typically, a fixed order value is used.

- Vast majority of studies employ complex feature sets for vocal fold paralysis detection. So far, only small part of them are physiologically motivated, i.e. reflect glottal flow directly. Most of employed features (like MFCC, PLP) contain redundant information like linguistic content, emotional status of the speaker, etc.
- Despite numerous studies, acoustical analysis of vocal pathologies (including paralysis of vocal folds) still remains a challenging task.

In this paper, we present the AR model-based inverse filtering approach for estimation of glottal flow and detection of vocal fold paralysis. A variable order AR model was employed to model the vocal tract and the glottal flow.

3. The Proposed Method

3.1. *The Autoregressive Model-Based Inverse Filtering Model*

The glottal flow can be obtained from the voiced speech segments. Considering the *source-filter* approach, the speech signal $s(t)$ can be expressed as the convolution of the glottal flow $g(t)$ and the vocal tract filter $h(t)$

$$s(t) = g(t) * h(t). \tag{1}$$

Here the lip radiation effect (modelled as a first-order differentiating filter) is included in the vocal tract processing and is not considered separately. Traditionally, the vocal tract is modelled using an all-pole filter for speech analysis purposes.

If we obtain an estimate of the inverse vocal tract filter $\hat{h}^{-1}(t)$ and apply it to the analysed speech signal $s(t)$, we will eliminate the effect of the vocal tract thus obtaining the estimate of the glottal flow

$$\hat{g}(t) = s(t) * \hat{h}^{-1}(t). \tag{2}$$

In this study, we applied AR model for the modelling of the vocal tract. The choice was due to the following reasons:

- The AR model (also known as Linear Predictive Coding model) is an all-pole filter and had great success in speech applications. The adequacy of the AR model parameter estimation technique (Kaukėnas, 1983) to the speech signal was shown in Kaukėnas and Tamulevičius (2016) and Tamulevičius and Kaukėnas (2016).
- The linearity of filter enables us to obtain an inverse version of the filter very easy.
- The chosen parameter estimation technique enables us to obtain a variable model order which is adequate to individual characteristics of human vocal properties. Therefore, we can expect a more accurate estimation of the glottal flow.

The AR model parameter estimation technique is presented in the next subsection.

3.2. *Estimation of the AR Model*

Let us explore the speech signal as the process $\{S_t\}$ with zero mean and describe it using the AR model

$$\sum_{i=0}^{M} a_i S_{t-i} = bV_t, \quad a_0 = 1, \quad t = 1, 2, \ldots, N, \tag{3}$$

where N is the length of the signal S_t, $\{V_t, \ t = 1, 2, \ldots\}$ is the process of mutually independent and normally distributed random variables.

Our task is to estimate the model order M, the parameters $\{a_1, a_2, \ldots, a_M\}$ and b of the AR model.

From (3) we can obtain

$$
\begin{aligned}
-s_{M+1} &= a_1 s_M + a_2 s_{M-1} + \cdots + a_M s_1 + b v_{M+1}, \\
-s_{M+2} &= a_1 s_{M+1} + a_2 s_M + \cdots + a_M s_2 + b v_{M+2}, \\
&\ \ \vdots \\
-s_N &= a_1 s_{N-1} + a_2 s_{N-2} + \cdots + a_M s_{N-M} + b v_N,
\end{aligned}
\tag{4}
$$

If we denote

$$
\begin{aligned}
Y' &= (-s_{M+1}, -s_{M+2}, \ldots, -s_N), \\
X_1' &= (s_M, s_{M+1}, \ldots, s_{N-1}), \\
X_2' &= (s_{M-1}, s_M, \ldots, s_{N-2}), \\
&\ \ \vdots \\
X_M' &= (s_1, s_2, \ldots, s_{N-M}), \\
X &= (X_1, X_2, \ldots, X_M), \\
A' &= (a_1, a_2, \ldots, a_M), \\
V' &= (V_{M+1}, V_{M+2}, \ldots, V_N),
\end{aligned}
$$

we get the following expression of the AR model

$$Y = X \cdot A + bV. \tag{5}$$

The equation is solved using the recurrent evaluation approach (Kaukėnas, 1983). The Efroymson matrix is composed

$$E = \begin{bmatrix} R(M \times M) & T'(M \times 1) & I(M \times M) \\ T(1 \times M) & I(1 \times 1) & O(1 \times M) \\ -I(M \times M) & O(M \times 1) & O(M \times M) \end{bmatrix}, \tag{6}$$

where R is the cross-correlation matrix of X_i and X_j, $i, j = 1, \ldots, M$; T is the cross-correlation vector of Y and X_i, $i = 1, \ldots, M$; O denotes zero vectors and matrices, and I is a unit matrix.

Each new sequence X_i is included during the recurrent modification of the Efroymson matrix

$$E'(i, j) = E(i, j)/E(i, i), \quad j = 1, 2, \ldots, 2M + 1, \tag{7}$$

$$E'(k, l) = E(k, l) - \frac{E(k, i)E(i, l)}{E(i, i)}, \quad k, l = 1, 2, \ldots, 2M + 1, \ l \neq i. \tag{8}$$

$E(i, j)$ denotes the Efroymson matrix before including X_i, $E(i, j)'$ is an updated version of the Efroymson matrix with included X_i.

Finally, the model parameter a_i is estimated

$$\hat{a}_i = E(i, M + 1)\sqrt{(Y'Y)/(X_i'X_i)}, \quad i = 1, 2, \ldots, M. \tag{9}$$

The estimate of b^2 is obtained as follows

$$\hat{b}^2 = E(M + 1, M + 1)(Y'Y)/(N_0 - M). \tag{10}$$

The model order estimation task is solved by comparing M_1 and M_2 order models. Usually, the estimated variances of prediction error $\hat{b}_{M1}^2$ and $\hat{b}_{M2}^2$ are compared. The following estimator for the model order was formulated in Kaukėnas (1983)

$$\hat{M}_i = \begin{cases} i, & \text{if } \left(\frac{\hat{b}_i^2 - \hat{b}_{i-1}^2}{\hat{b}_i^2}\right)(N_0 - i) > F_{cr}(1, N - i), \\ 0, & \text{otherwise}, \end{cases} \tag{11}$$

$$\hat{M} = \max_i\{\hat{M}_i\}, \quad i = 1, \ldots, M_{\max},$$

where $F_{cr}(1, N - i)$ is the quantile of Fisher distribution with 1 and $(N - i)$ degrees of freedom; $\hat{b}_0^2$ is the estimate of variance, i.e. $\hat{b}_0^2 = \hat{D}$, $\hat{D}$ is the estimated variance of the process. $M_{\max}$ is the maximum model order value, it is based on empirical knowledge of the signal.

In this study, we have chosen $M_{\max} = 20$ for the vocal tract model. The filter with order up to 20th will model up to 10 resonant frequencies (formants), which is completely sufficient for description of speaker's individual articulation (vocal tract) properties.

For the modelling of the glottal flow we have chosen $M_{\max} = 200$. The decision is based on the results obtained in Tamulevičius and Kaukėnas (2017), where description of individual speakers qualities demanded AR model order up to 170.

The quality of the estimated glottal flow was assessed on the basis of the ratio of estimated squared prediction error and estimated signal variance $\hat{b}^2/\hat{D}$. The value of $\hat{b}^2/\hat{D}$ indicates the relative part of the unmodelled signal: the higher ratio value we obtain,

the higher signal prediction error is. Therefore, we can expect a low $\hat{b}^2/\hat{D}$ value for normal glottal flow and high values for pathological voices (with paralysis).

In this study, we will express this ratio in percentage and call it the estimated error of glottal flow. We think that for healthy and normal voices this ratio will approach towards zero level, and for pathological voices, it will converge to 100% (in case of full paralysis or dysfunction of vocal folds).

3.3. *Inverse Filtering of the Speech Signal*

In this subsection we will present the algorithm of the inverse filtering of the speech signal and estimation of the glottal flow quality.

Step 1. The vocal tract filter $h(t)$ is modelled using AR model: the estimates of model order M (11) and parameters $\{a_1, a_2, \ldots, a_M, b\}$ (9)–(10) are obtained with $M_{max} = 20$.

Step 2. The estimate of inverse filter $\hat{h}^{-1}(t)$ is constructed and the estimate of the glottal flow is obtained

$$\hat{g}_t = \sum_{i=0}^{M} a_i S_{t-i}, \quad t = 1, 2, \ldots, N.$$

Step 3. The AR model order M' and parameter estimates $\{a'_1, a'_2, \ldots, a'_{M'}, b'\}$ are obtained for the glotal flow $\hat{g}(t)$ with $M'_{max} = 200$. The quality of the $\hat{g}(t)$ is assessed by value $\hat{b}'^2/\hat{D}'$.

4. Experimental Analysis

4.1. *Experimental Data*

For experimental analysis of the proposed method, records of two voice types were collected.

Starting in 2016, patients scheduled for thyroidectomy and included in the study (study launched in Vilnius University Faculty of Medicine Institute of Clinical Medicine Clinics of Gastroenterology, Nephrourology and Surgery in cooperation with the Institute of Data Science and Digital Technologies) were selected for voice recording and vocal folds movement evaluation before and after the operation. Vilnius regional Biomedical research committee permission No. 158200-15-819-331 has been given in 2015.12.08. The interval comparison of sequential voice recording was matched against change in vocal folds movement. The vocal folds function was assessed by a laryngoscopy in each case before and after thyroidectomy procedure.

A prospective trial was launched in March 2016 and finished in May 2017. 112 patients with known thyroid pathology were prospectively enrolled in this study. All 112 patients were operated on in Vilnius University Hospital Santaros Klinikos. The study protocol

included voice recording and laryngeal exam in all patients preoperatively and postoperatively by a qualified ENT specialist. 6 cases of temporary vocal cord palsy were diagnosed on postoperative examination (5.4% injury rate per patient and 3% per nerve at risk). No cases of permanent or bilateral vocal cord palsy were recognized postoperatively.

All the patient voices were recorded using headset microphones in a clinician's room environment. There were 4 recording sessions organized: one day before surgery, one day, 2 weeks, and 4 months after surgery.

The control group consisted of healthy people with no complaints or throat/mouth surgery procedures in last 3 months. The voices of 10 female and 10 male speakers were recorded using voice recorder with an external microphone in a silent room environment.

All the recorded persons were asked to pronounce vowel [a] in a sustained manner for 3–4 seconds. This vowel is characterized by a minimal lip restriction during radiation phase and a fully expressed phonation level. Besides, vowel [a] is common for most languages, what makes it universal for comparison purposes.

4.2. *Case Analysis*

For analysis of pathological and healthy voices, we have selected two voices for inverse filtering procedure and estimation of glottal flow. The estimated signal of glottal flow and its spectral density function were analysed to estimate the qualities of the pathological and healthy voices.

Figure 1 presents the results obtained for the healthy female's voice. The estimated order of the vocal tract filter was 11 (i.e. the vocal tract had 6 resonant frequency values). The estimated glottal flow can be evaluated as periodic and normal (Fig. 1(b)). Spectral density function (Fig. 1(c)) is also periodic, the harmonic components are vivid through the entire frequency range of the signal.

The results of pathological male voice analysis are given in Fig. 2. Here we can see the distorted waveform of the utterance (Fig. 2(a)). The vocal tract was modelled by 20-th order model which means ten resonant frequencies of the tract. The estimated glottal (Fig. 2(b)) flow is noisy with no sign of periodicity (what is characteristic for the vocalized vowel). The spectral density (Fig. 2(c)) of the flow is noise-like, here we can see only 4–5 harmonic components. This is the evidence of vocal fold immobility which can be the result of the vocal fold paralysis.

Similar results were obtained for all pathological voices: non-periodicity of the estimated glottal flow, noise-like spectral density function. The degree of non-periodicity was different for the individual voices. This difference may be with individual characteristics of the voices and require a more detailed study with larger datasets.

4.3. *Experimental Results*

First of all, we evaluated the error level of glottal flow for healthy and pathological voices. The averaged results are given in Fig. 3.

We can see the clear difference between healthy and pathological voices. The patients' voices (before thyroidectomy surgery) have at 50% higher error level than healthy ones.
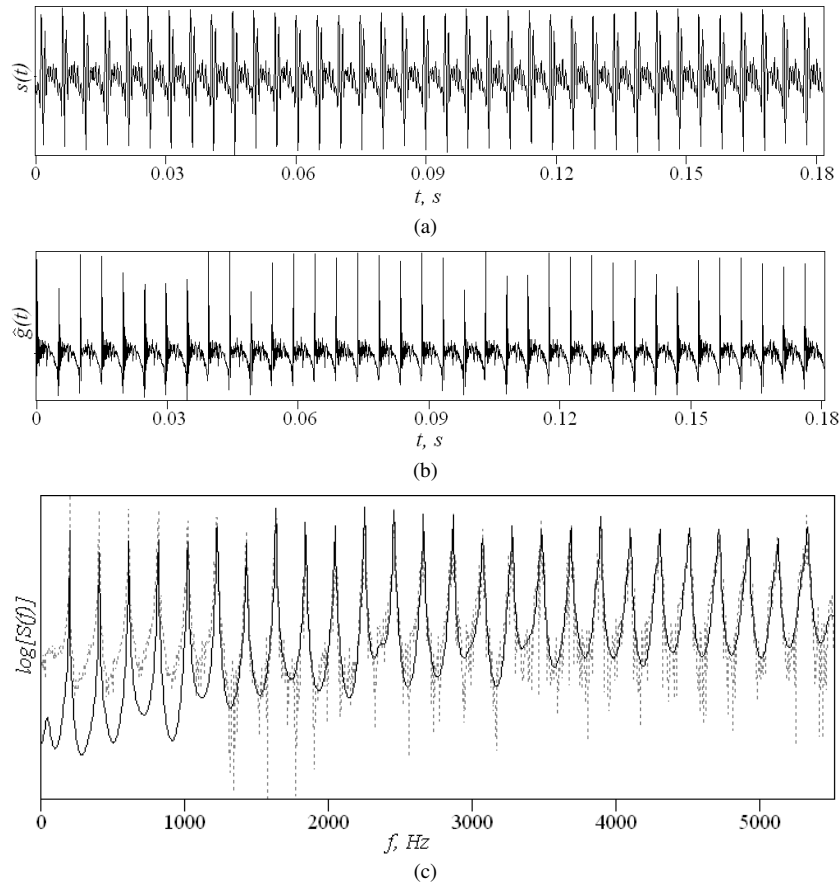
Fig. 1. The healthy voice: (a) the waveform of the vowel [a]; (b) the estimated glottal flow; (c) the spectral density of the glottal flow (AR model-based spectral density is given in solid line, Fourier transform-based spectral density is given in dotted line).

The thyroidectomy procedure with the output of the immobility of the vocal folds increased the error level by 15–50% (by 2–3 times in comparison with healthy voices). Therefore, the prediction error level of the glottal flow enables us to identify the case of vocal fold paralysis.

Nevertheless, the amount of analysed data is not sufficient to make statistically reasoned conclusions and to propose some global criteria for detection of vocal fold paralysis. The main reason is the scattering of the results because of individual properties of the persons' voice. Every person is characterized by his own inherent qualities of glottal flow, so the output of the surgery (which is also very characteristic to person) should be estimated individually, taking into account these qualities. To illustrate this statement the data about the status of 3 patient's vocal folds is given in Fig. 4.

Comments in Fig. 4:

Female #1. This patient has been diagnosed with paralysis of the vocal folds after thyroidectomy surgery. Only a partial recovery of folds mobility has been stated after
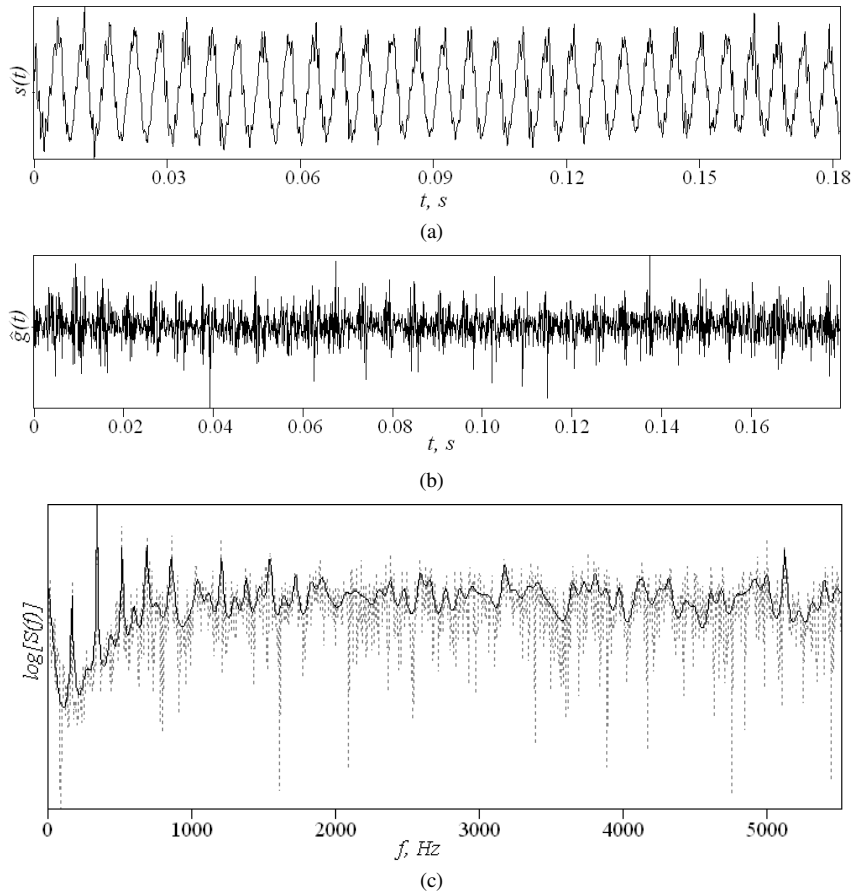
Fig. 2. The pathological voice: (a) the waveform of the vowel [a]; (b) the estimated glottal flow; (c) the spectral density of the glottal flow (AR model-based spectral density is given in solid line, Fourier transform-based spectral density is given in dotted line).
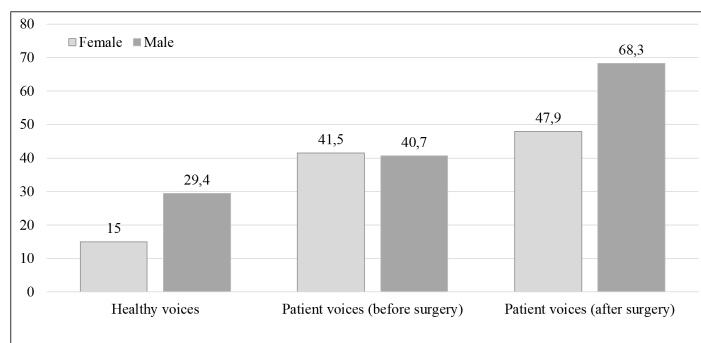

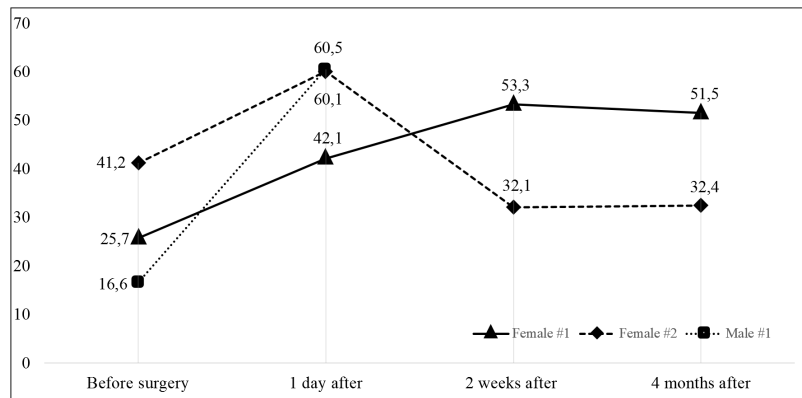Fig. 3. The estimated error level of glottal flow for different voices.

Fig. 4. The change of estimated vocal fold status for 3 patients.

4 months. In Fig. 4 we can see only slightly improving status of vocal folds (solid line).

Female #2. In this case, we also can see the change of folds mobility after surgery (paralysis was diagnosed). However, after two weeks the status of the folds had improved significantly and became much better than before the surgery and remained unchanged after 4 months (dashed line). The dynamics of the glottal flow quality is given in Fig. 5. There we can see the obvious improvement of the glottal flow quality. The glottal flow after thyroidectomy has become noisy and non-periodic (Fig. 5 (b)). After two weeks the flow was more stable and periodic (Fig. 5 (c)) even compared with preoperative status (Fig. 5 (a)).

Male #1. This patient's data show the drastic change of vocal fold status (dotted line). The estimated error level of the glottal flow had increased almost 4 times. So far, the monitoring of this patient has not yet been completed, so there is no data on the current state of this patient's vocal folds.

It is obvious that glottal flow prediction error-based estimation of the vocal fold functionality should be performed individually. As we can see in Fig. 4, the preoperative and postoperative status of vocal folds were different for patients, and the recovery process is also individual. Therefore, this assessment can be implemented as monitoring the dynamics of vocal fold functionality for screening examination method to select patients for laryngoscopic procedure. Relative change of the glottal flow prediction error reflects changes in glottal flow. For application purposes, the change should be parametrized.

5. Conclusion

The formulated vocal fold mobility assessment technique and the experimental results obtained can be summarized as follows:

- The Autoregressive model-based digital inverse filtering technique is presented for estimation of the glottal flow. The novelty of the proposed method is the objec-
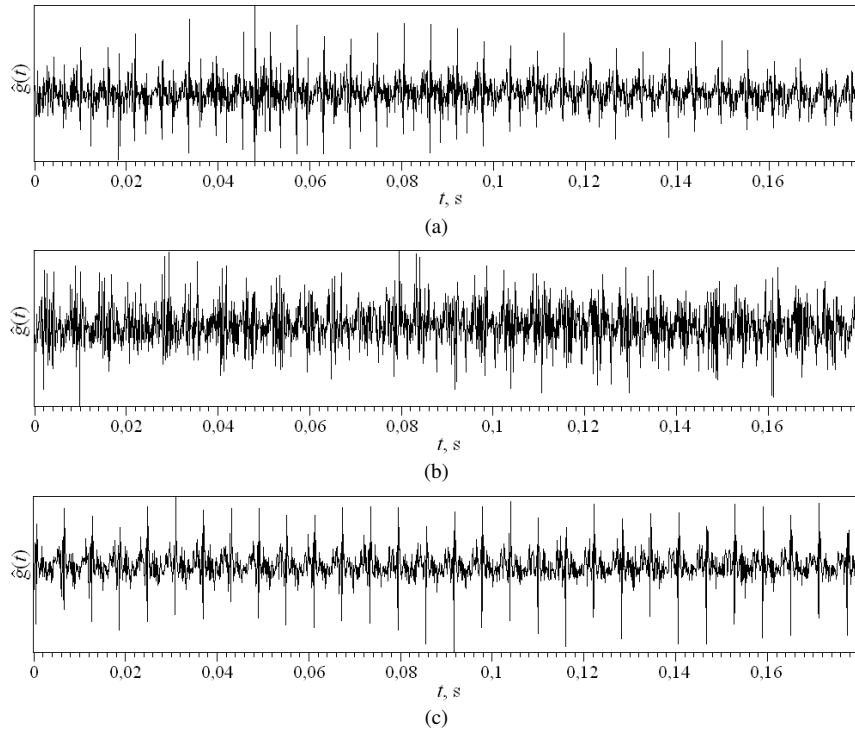
Fig. 5. Dynamics of the glottal flow for patient *Female #2*.

tive and adequate selection of a variable model order, which enables us to obtain a more accurate evaluation of individual articulation properties than a fixed-order modelling. This postulates the more accurate estimation of the glottal flow, disturbances of which are direct evidence of the vocal fold paralysis.

- The glottal flow differs for healthy and pathological voices. AR modelling of the glottal flow gives at least 50% higher prediction error level for pathological voices (before the thyroidectomy procedure). The surgery procedure increases this difference 2–3 times. Nevertheless, the results were obtained for 20 healthy and 6 pathological voices. Therefore, statistical significance of the results is not high.

- Prediction error-based global and universal glottal flow assessment criteria for paralysis detection cannot be formulated so far. The voice production system is very specific to each speaker, the impact of the surgery is also very specific. Thus mobility of the vocal folds should be estimated individually, taking into account individual qualities, comparing preoperative and postoperative voice qualities. The employed AR model parameter estimation technique is capable of describing these individual properties and using of a prediction error to monitor the dynamics of vocal fold functionality before and after thyroidectomy procedure.

References

Airaksinen, M., Raitio, T., Story, D., and Alku, P. (2014). Quasi closed phase glottal inverse filtering analysis with weighted linear prediction. *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, 22(3), 596–607.

Ali, Z., Elamvazuthi, I., Alsulaiman, M. (2016). Detection of voice pathology using fractal dimension in multiresolution analysis of normal and disordered speech signals. *Journal of Medical Systems*, 40(20).

Alku, P., Magi, C. (2009). Closed phase covariance analysis based on constrained linear prediction for glottal inverse filtering. *The Journal of the Acoustical Society of America*, 125(5), 3289–3305.

Alku, P. (2011). Glottal inverse filtering analysis of human voice production – a review of estimation and parametrization methods of the glottal excitation and their applications. *Sadhana*, 36(5), 623–650.

Arroyave, R.O., Bonilla, F.V., Trejos, D.T. (2012). Acoustic analysis and non linear dynamics applied to voice pathology detection: a review. *Recent Patents on Signal Processing*, 2(2), 96–107.

Baljekar, P.N., Patil, H.A. (2012). A comparison of waveform fractal dimension techniques for voice pathology classification. In: *Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 4461–4464.

Bergenfelz, A., Jansson, S., Kristoffersson, A., Mårtensson, H., Reihnér, E., Wallin, G., Lausen, I. (2008). Complications to thyroid surgery: results as reported in a database from a multicenter audit comprising 3660 patients. *Langenbeck's Archives of Surgery*, 393(5), 667–673.

Cairns, D.A., Hansen, J.H.L., Riski, J.E. (1994). Detection of hypernasal speech using a nonlinear operator. In: *Proceedings of the IEEE Conference on Engineering in Medicine and Biology Society*, pp. 253–254.

Dejonckere, P., Wieneke, G.H. (1994). Spectral, cepstral and aperiodicity characteristics of pathological voice before and after phonosurgical treatment. *Clinical Linguistics & Phonetics*, 8(2), 161–169.

Dibazar, A.A., Narayanan, S., Berger, T.W. (2002). Feature analysis for automatic detection of pathological speech. In: *Proceedings of the Second Joint EMBS/BMES Conference*, Houston, USA, October 23–26, 2002, pp. 182–183.

Elsheikh, E., Quriba, A.S, El-Anwar, M.W. (2016). Voice changes after late recurrent laryngeal nerve identification thyroidectomy. *Journal of Voice*, 30(6), 762.e1–762.e9.

Fukazawa, T., el-Assuooty, A., Honjo, I. (1988). A new index for evaluation of the turbulent noise in pathological voice. *Journal of Acoustical Society of America*, 83(3), 1189–1193.

Giovanni, A., Ouaknine, M., Triglia, J.M. (1999). Determination of largest Lyapunov exponents of vocal signal: application to unilateral laryngeal paralysis. *Journal of Voice*, 13(3), 341–354.

Henry, L., Helou, L., Solomon, N., Howard, R.S., Gurevich-Uvena, J., Coppit, G., Stojadinovic, A. (2010). Functional voice outcomes after thyroidectomy: an assessment of the Dysphonia Severity Index (DSI) after thyroidectomy. *Surgery*, 147(6), 861–870.

Hillenbrand, J., Cleveland, R.A., Erickson, R.L. (1994). Acoustic correlates of breathy vocal quality. *Journal of Speech and Hearing Research*, 37(4), 769–777.

Jeannon, J.P., Orabi, A.A, Bruch, G.A., Abdalsalam, H.A, Simo, R. (2009). Diagnosis of recurrent laryngeal nerve palsy after thyroidectomy: a systematic review. *International Journal of Clinical Practice*, 63(4), 624–629.

Kafentzis, G.P., Stylianou, Y., Alku, P. (2011). Glottal inverse filtering using stabilised weighted linear prediction. In: *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 5408–5411.

Kasuya, H., Kobayashi, Y., Kobayashi, T., Ebihara, S. (1983). Characteristics of pitch period and amplitude perturbations in pathologic voice. In: *Proceedings of International Conference on Acoustics, Speech, and Signal Processing ICASSP*, pp. 1372–1375.

Kasuya, H., Ogawa, S., Mashima, K., Ebihara, S. (1986). Normalized noise energy as an acoustic measure to evaluate pahologic voice. *Journal of Acoustical Society of America*, 80(5), 1329–1334.

Kaukėnas, J. (1983). On estimation of ar model order and parameters. *Statistical Problems of Control*, 61, 46–60 (in Russian).

Kaukėnas, J., Tamulevičius, G. (2016). Analysis of autoregressive model adequacy for Lithuanian vowels. *Proceedings of Lithuanian Mathematical Society (Series B)*, 57, 19–24 (in Lithuanian).

Koike, Y. (1967). Application of some acoustic measures for the evaluation of laryngeal dysfunction. *Journal of Acoustical Society of America*, 42(5), 1209.

Lieberman, P. (1963). Some acoustic measures of the fundamental periodicity of normal and pathologic larynges. *Journal of Acoustical Society of America*, 35(3), 344–353.

Lifante, J.C., Payet, C., Menegaux, F., Sebag, F., Kraimps, J.L., Peix, J.L., Pattou, F., Colin, C., Duclos, A. (2017). Can we consider immediate complications after thyroidectomy as a quality metric of operation? *Surgery*, 161(1), 156–165.

Mihai, R., Randolph, G.W. (2009). Thyroid surgery, voice and the laryngeal examination-time for increased awareness and accurate evaluation. *World Journal of Endocrine Surgery*, 1(1), 1–5.

Musholt .T.J., Musholt, P.B., Garm, J., Napiontek, U., Keilmann, A. (2006). Changes of the speaking and singing voice after thyroid or parathyroid surgery. *Surgery*, 140(6), 978–988.

Ortega, J., Cassinello, N., Dorcaratto, D., Leopaldi, E. (2009). Computerized acoustic voice analysis and subjective scaled evaluation of the voice can avoid the need for laryngoscopy after thyroid surgery. *Surgery*, 145(3), 265–271.

Page, C., Zaatar, R., Biet, A., Strunski, V. (2007). Subjective voice assessment after thyroid surgery: a prospective study of 395 patients. *Indian Journal of Medical Sciences*, 61(8), 448–454.

Panek, D., Skalski, A., Gajda, J., Tadeusiewicz, R. (2015). Acoustic analysis assessment in speech pathology detection. *International Journal of Applied Mathematics and Computer Science*, 25(3), 631–643.

de Pedro Netto, I., Fae, A., Vartanian, J.G., Barros, A.P., Correia, L.M., Toledo, R.N., Testa, J.R., Nishimoto, I.N., Kowalski, L.P., Carrara-de Angelis, E. (2006). Voice and vocal self-assessment after thyroidectomy. *Head Neck*, 28(12), 1106–1114.

Sinagra, D.L., Montesinos, M.R., Tacchi, V.A., Moreno, J.C., Falco, J.E., Mezzadri, N.A., Debonis, D.L., Curutchet, H.P. (2004). Voice changes after thyroidectomy without recurrent laryngeal nerve injury. *Journal of American College of Surgeons*, 199(4), 556–560.

Stojadinovic, A., Shaha, A.R., Orlikoff, R.F., Nissan, A., Kornak, M.-F.,Singh, B., Boyle, J.O., Shah, J.P., Brennan, M.F., Kraus, D.H. (2002). Prospective functional voice assessment in patients undergoing thyroid surgery. *Annals of Surgery*, 236(6), 823–832.

Tamulevičius, G., Kaukėnas, J. (2016). Adequacy analysis of autoregressive model for Lithuanian semivowels. In: *Proceedings of IEEE 4th Workshop on Advances in Information, Electronic and Electrical Engineering (AIEEE)*, pp. 1–4.

Tamulevičius, G., Kaukėnas, J. (2017). High-order autoregressive modeling of individual speaker's qualities. In: *Proceedings of IEEE 5th Workshop on Advances in Information, Electronic and Electrical Engineering (AIEEE)*. (Accepted for publishing).

Vaičiukynas, E., Verikas, A., Gelžinis, A., Bačauskienė, M., Minelga, J., Hålander, M., Padervinskis, E., Uloza, V. (2015). Fusing voice and query data for non-invasive detection of laryngeal disorders. *Expert Systems With Applications*, 42, 8445–8453.

Walker, J., Murphy, P. (2007). A review of glottal waveform analysis. In Stylianou, Y., Faundez-Zanuy, M., Esposito, A. (Eds.), *Progress in Nonlinear Speech Processing. Lecture Notes in Computer Science*, Vol. 4391, pp. 1–21.

Yumoto, E., Gould, W.J., Baer, T. (1982). Harmonics to Noise Ratio as hoarseness index of degree of hoarseness. *Journal of Acoustical Society of America*, 71(6), 1544–1550.

A. Rybakovas is a PhD student in Vilnius University Medical Faculty, an abdominal surgeon in the Centre of Abdominal Surgery, Vilnius University Hospital Santaros Klinikos. Scientific interests include endocrine surgery, upper GI surgery.

V. Beiša is a head of the Centre of Abdominal Surgery at Vilnius University Hospital Santaros Klinikos, professor. In 1989 he received a PhD degree in biomedical sciences, Vilnius University. In 2009 he received habilitated doctor degree in biomedical sciences. Scientific interests include endocrine surgery, surgical treatment of thyroid, parathyroid, adrenal gland, pancreatic endocrine tumors.

K. Strupas is a head of Clinic of Gastroenterology, Nephrourology, and Surgery, professor. In 1989 he received a PhD degree in biomedical sciences, Vilnius University. In 1997 he received habilitated doctor degree in biomedical sciences. Starting from 2002 he is a chairman of Clinic for Visceral Surgery and Gastroenterology Medical Faculty Vilnius University, member of the Vilnius University Senate, director of Clinic for General and Visceral Surgery. Starting from 2014 he is a full member of the Lithuanian Academy of Sciences. Scientific interests include minimally invasive surgery, strategies of treatment in HPB surgery, transplantation. He published 366 scientific articles.

J. Kaukėnas is a long-time employee of the Institute of Mathematics and Informatics (now Vilnius University Institute of Data Science and Digital Technologies). His main research areas were an analysis of random signals, analysis of heart rate, speech signal analysis, and modelling.

G. Tamulevičius is a researcher at the Vilnius University Institute of Data Science and Digital Technologies. Currently, he also is an associate professor at Vilnius Gediminas Technical University. His main research interests include speech signal modelling and its application for speech and speech emotion recognition, speech pathology detection.