ISCA Archive
http://www.isca-speech.org/archive

First International Conference on
Spoken Language Processing (ICSLP 90)
Kobe, Japan
November 18-22, 1990

A NOVEL MODEL OF PATHOLOGICAL VOCAL CORDS AND ITS APPLICATION TO THE DIAGNOSIS OF VOCAL CORD POLYP

Takuya Koizumi and Shuji Taniguchi

Faculty of Engineering, Fukui University, 3-9-1 Bunkyo, Fukui 910, Japan

ABSTRACT

The present paper deals with a new noninvasive method of diagnosing vocal cord polyp through hoarse voice analysis. A noteworthy feature of this method is that it enables us not only to discriminate hoarse voices caused by pathological vocal cords with a single polyp from hoarse voices due to other laryngeal diseases but also to estimate polyp features such as the mass and dimension of polyp through the use of a novel model of pathological vocal cords which has been devised to simulate the subtle movement of the vocal cords with a single polyp. A synthetic hoarse voice produced with a hoarse voice synthesizer is compared with a natural hoarse voice caused by the vocal cord polyp in terms of a distance measure and the polyp features are estimated by minimizing the distance measure. Some estimates of polyp dimension that have been obtained by applying this procedure to hoarse voices are found to compare favorably with actual polyp dimensions, demonstrating that the procedure is effective for the diagnosis of vocal cord polyp.

I. INTRODUCTION

A number of research works on the acoustic analysis of hoarse voices which have been reported in the last decade may be divided into two groups by their objectives. Some of those works tried to establish techniques of discriminating voices caused by pathological vocal cords from normal voices, while others aimed at developing a procedure for the diagnosis, that is, the classification, of laryngeal disease only through the analysis of hoarse voice. To the best of our knowledge, however, none of them have attempted to develop a noninvasive procedure for estimating the pathological condition of the larynx.

An attempt toward this end has recently been made, resulting in a new noninvasive procedure developed for the diagnosis of vocal cord polyp through hoarse voice analysis. The present paper is concerned with this new procedure. A noteworthy feature of this procedure is that it enables us not only to discriminate hoarse voices caused by the vocal cord polyp from hoarse voices due to other laryngeal diseases but to estimate polyp features such as the mass and dimension of polyp through the use of a novel model of pathological vocal cords which has been devised to simulate the subtle movement of the vocal cords with a single polyp.

In what follows, first a hoarse voice synthesizer which plays an important role in estimating the polyp features is described in detail. Then a newly developed method for estimating the polyp features is discussed. This method, which might be called an analysis by synthesis approach, is characterized by a fact that the hoarse voice synthesizer and a distance measure for hoarse voices formed of several specific physical variables that are closely related to the pathological condition of the vocal cords are used to estimate the polyp features. Finally the effectiveness of this method is examined by comparing estimates of the dimensions of polyps that were derived from over ten samples of hoarse voice caused by the vocal cord polyp with the actual dimensions of polyps.

II. THE HOARSE VOICE SYNTHESIZER

2.1 A five-mass model of pathological vocal cords

The hoarse voice synthesizer consists of a vocal tract model and a model of pathological vocal cords. The pathological vocal cords which are usually accompanied by anomalous change in physiological and physical properties, such as the mass, volume, and tension, of the tissue and also by the lack of symmetry between the opposing cords can be modeled as a mechanical vibration system made up of four independent masses coupled by springs and dampers. To properly model the vocal cords with a single polyp, however, it is necessary to add one more mass representing the polyp to the four-mass model, because the motion of the polyp relative to a vocal cord to which it is attached is known to be like that of a pendulum. This results in a five-mass model of the vocal cords with a single polyp as shown in Fig. 1, which is able to demonstrate some subtle movements of the vocal cords with a single polyp such as the pendulum motion of the polyp. A brief description of this model will be given below. The upper and lower edges of the vocal cords are represented by the masses m_1, m_2 and M_1, M_2, respectively. The lower masses M_1 and M_2 are allowed to move only in the horizontal direction, whereas the upper masses m_1 and m_2 have both horizontal and vertical degrees of freedom. The upper and lower masses m_i and M_i (i=1, 2) are coupled by the springs S_{i2x} and S_{i2y}, and the dampers r_{i2x} and r_{i2y}. Assuming that the thickness of the upper masses is variable leads to a behavior of the upper masses somewhat like the mucosal surface wave [1]. Since the volumes of the upper masses vary with time, both M_i and m_i will also vary with time, if it is assumed that the total mass remains constant. Obviously the change in mass is proportional to the change in the thickness of the upper masses.

The mass m_p representing the polyp is connected with the upper mass m_2 by the spring S_p and the damper r_p and is allowed to move on m_2 in the horizontal direction. For simplicity of

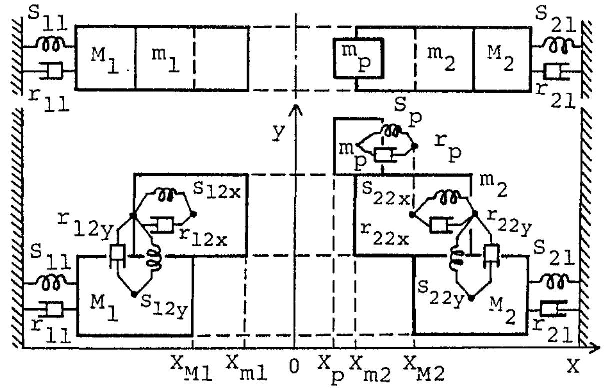

Fig. 1 The structure of a five-mass model of vocal cords with a single polyp.

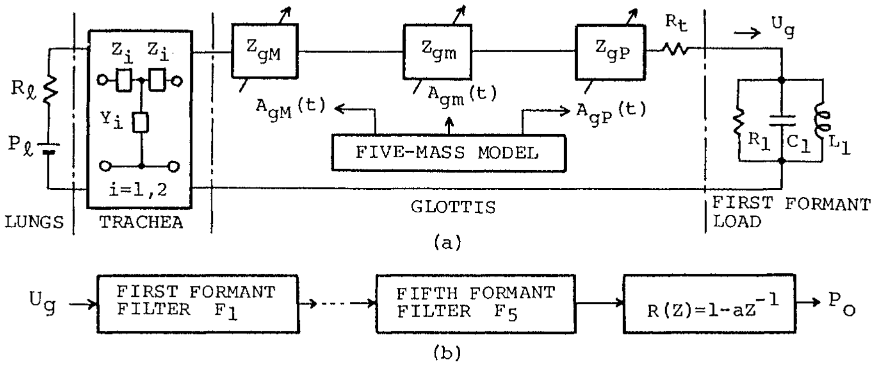

Fig. 2 The hoarse voice synthesizer: (a) the glottal volume flow source and (b) the vocal tract model.

analysis it is assumed that the polyp is in the form of a cube and that its form and mass remain fixed at all times. Thus the thickness δ_p of the polyp is given by $\delta_p=(m_p/\rho_p)^{1/3}$, where ρ_p is the density of the polyp and is assumed to be equal to the average density of the vocal cords, a typical value of which is 1.1 g/cm³.

2.2 The hoarse voice synthesizer

The hoarse voice synthesizer [4], as shown in Fig. 2, consists of a glottal volume flow source and a vocal tract model. The glottal volume flow source comprises a subglottal system, a glottal impedance controlled by the five-mass model of the vocal cords, and a first formant load. The vocal tract model that is made up of five formant filters and a simple radiation load produces the output P_o which is a synthetic hoarse voice, when excited by the output U_g of the glottal volume flow source.

Using the hoarse voice synthesizer, a number of hoarse voices have been synthesized. The effect of varying the values of m_p and K_p on the waveforms of output sound pressure and glottal volume flow, glottal areas, and displacements of masses has been examined with the following results:

a. The double-pendulum motion of the polyp such that the motion of the polyp is always behind that of the upper mass carrying the polyp has been observed.

b. This delay of the polyp in motion increases with its mass m_p, while its displacement decreases with m_p.

c. The polyp comes to move across the plane x=0, as m_p is decreased.

d. When the value of m_p exceeds 100 mg, the movement of the polyp becomes irregular, and waveforms of the glottal volume flow and output sound pressure turn out to be different from period to period.

III. THE METHOD OF ESTIMATING THE POLYP FEATURES

The procedure of the diagnosis of vocal cord polyp is shown in Fig. 3. First, an input voice sample of, say, vowel /a/ undergoes a preliminary classification [2], [3], and is classified into three classes of normal voice, hoarse voice caused by the vocal cord polyp, and hoarse voice due to other laryngeal disease. Second, a hoarse voice is subjected to a test based on a couple of specific acoustic parameters called PPQ and APQ to discriminate a hoarse voice caused by the vocal cord polyp from a normal voice. Last, for a hoarse voice whose cause has been decided to be the vocal cord polyp, the polyp features, i.e., the mass and dimension of the polyp are estimated through the use of an optimization technique described below.

3.1 PPQ and APQ

The acoustic parameters of voice called PPQ (Pitch Perturbation Quotient) and APQ (Amplitude Perturbation Quotient) are known to be useful for the classification of laryngeal disease

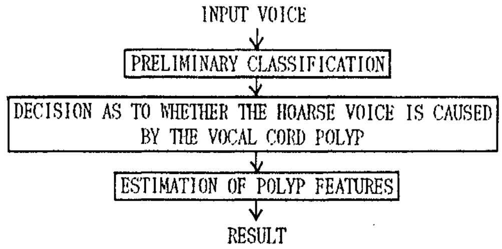

Fig. 3 The procedure of the diagnosis of vocal cord polyp.

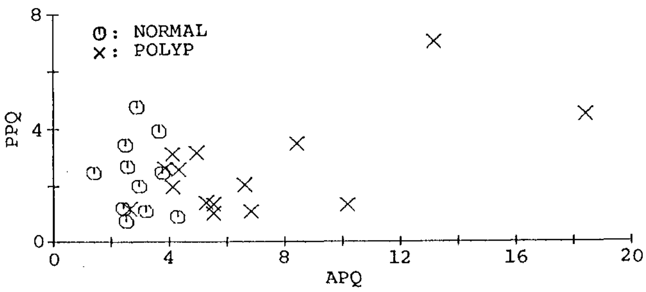

Fig. 4 The distribution of normal voices and hoarse voices caused by the vocal cord polyp on the PPQ-APQ plane.

through hoarse voice analysis. The PPQ and APQ which represent the fluctuation of the fundamental period and amplitude of hoarse voice, respectively, are defined as

$$PPQ = (1/N-1)\sum_{i=1}^{N-1} |T_{i+1} - T_i| / \{(1/N)\sum_{i=1}^{N-1} T_i\}, \quad (1)$$

$$APQ = (1/N-1)\sum_{i=1}^{N-1} |A_{i+1} - A_i| / \{(1/N)\sum_{i=1}^{N-1} A_i\}, \quad (2)$$

where N is the number of fundamental periods in a given voice sample and T_i and A_i (i=1,2,...,N-1) are sequences of the fundamental period and amplitude, respectively. The usefulness of these parameters to the present study is twofold, for they serve as means of discriminating hoarse voices caused by the vocal cord polyp from normal voices and are used for the distance measure to estimate the polyp features.

3.2 Relations between the polyp features and the PPQ and APQ

The PPQ and APQ were calculated for 11 normal voices and 15 hoarse voices caused by the vocal cord polyp with a result shown in Fig. 4 and the following findings:

a. The average value of the PPQs for the hoarse voices is nearly the same as that of the PPQs for the normal voices. Thus the PPQ does not seem to be strongly correlated with the polyp features.

b. The average value of the APQs for the hoarse voices is approximately two and half times as great as that of the APQs for the normal voices and nearly three times as great as that of the PPQs. In general values of the APQ seem to be greater than those of the PPQ for hoarse voices caused by the vocal cord polyp. This is particularly the case with hoarse voices due to a single, large polyp.

3.3 The discrimination of hoarse voices due to the vocal cord polyp

Hoarse voices due to the vocal cord polyp were discriminated from nomal voices by using a PPQ-APQ plane as shown in Fig. 4. Voices whose PPQs and APQs fell in a rectangular region on the PPQ-APQ plane corresponding to PPQ<5.0 and APQ<4.0 were categorized as normal voices and voices whose PPQs and APQs fell outside the region were identified as hoarse voices caused by the vocal cord polyp. With this discrimination scheme a rate of discrimination as high as 87 % has been achieved.

3.4 Estimation of the polyp features

It is possible to estimate the polyp features from a given hoarse voice caused by the vocal cord polyp by maximizing the the similarity between the hoarse voice and a synthetic hoarse voice produced with the hoarse voice synthesizer in terms of the distance measure. The method of estimating the polyp features in this work is based on this analysis by synthesis approach.

It is illustrated in Fig. 5. In order to produce a synthetic hoarse voice which is similar to the input voice in the sense described above, several important acoustic parameters including the PPQ and APQ have to be extracted from the input voice. First the input voice is subjected to the linear predictive analysis and frequencies and bandwidths of the first five formants are estimated. Then an average fundamental period is calculated from the sequence of the fundamental period to determine a pitch control parameter Q. Using these parameters and a pair of m_p and K_p specified somehow in the hoarse voice synthesizer will result in a synthetic hoarse voice which may or may not be similar to the input voice.

If the distance measure which will be defined in the next section is considered as a function of m_p and K_p, then the problem of estimating the polyp features will be equivalent to that of finding a minimum of a surface representing the distance measure over the m_p-K_p plane. To find the minimum requires to evaluate the distance measure at sufficiently many points which are formed into a lattice in the m_p-K_p plane. Finding the position of the minimum through the use of this array of points provides us with the polyp features.

3.5 The distance measure

The distance measure to be used for the purpose described above comprises four classes of acoustic features, i.e., the PPQ and APQ that are known to be effective in the discrimination of laryngeal diseases, the log magnitude of frequency spectrum at discrete frequencies, and the average fundamental period. It is given by

$$C = W_1 \sum_{i=1}^{N} \{DB_1(i) - DB_2(i)\}^2 + W_2 \mid PPQ_1 - PPQ_2 \mid + W_3 \mid APQ_1 - APQ_2 \mid + W_4 \mid T_{f1} - T_{f2} \mid , \quad (3)$$

where $DB_1(i)$ and $DB_2(i)$ denote log magnitudes (dB) of the ith frequency components of the spectra of input voice and synthetic

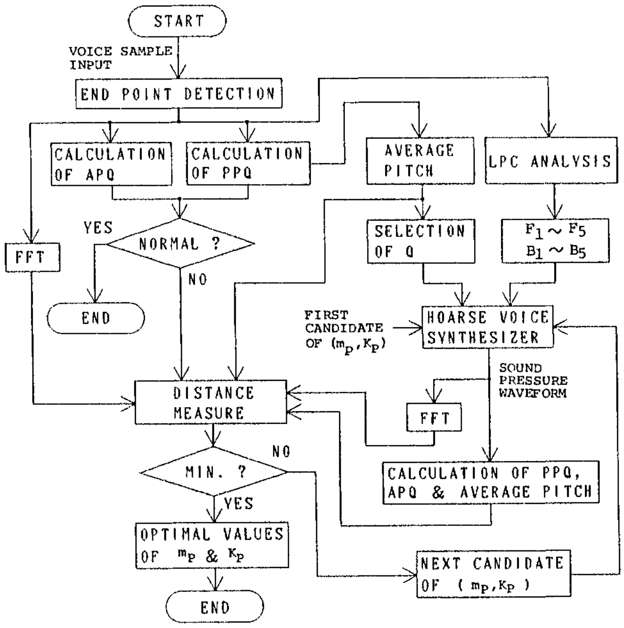

Fig. 5 The method of estimating the polyp features.

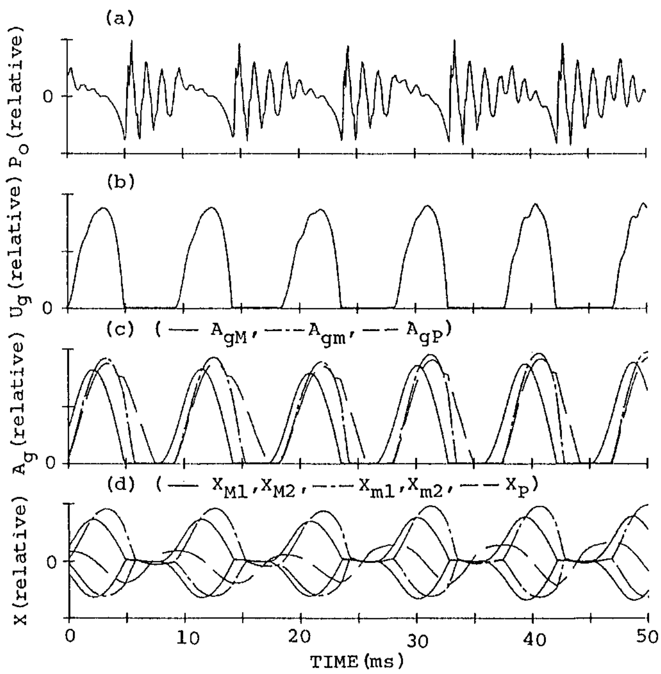

Fig. 6 Waveforms derived from the estimation scheme:
(a) synthetic hoarse voice (sound pressure),
(b) glottal volume flow, (c) glottal areas, and
(d) displacements of masses for the voice sample 14.

Table I Measured diameters of polyps and corresponding estimates of the diameters of polyps obtained by the estimation scheme.

VOICE SAMPLE	m_p (mg)	K_p (kdyn/cm)	MEASURED DIAMETER(mm)	ESTIMATE OF DIAMETER(mm)
1	38	16	4.4	4.04
2	62	9	4.9	4.76
3	46	52	4.6	4.31
4	21	5	2.7	3.32
5	24	12	3.0	3.47
6	27	32	3.5	3.61
7	11	18	1.6	2.67
8	19	39	1.4	3.21
9	16	24	1.8	3.03
10	32	23	3.6	3.82
11	30	32	3.0	3.74
12	121	13	5.7	5.95
13	20	32	1.8	3.26
14	54	18	4.5	4.54
15	19	22	1.4	3.21

voice, respectively, and T_{f1} and T_{f2} represent the average fundamental periods of the input and synthetic voices, respectively. The suffix 1 denotes the input voice and the suffix 2 the synthetic voice. The spectra are obtained through the use of FFT and a 102.4 ms window. The weights W_1, W_2, W_3, and W_4 have been chosen as follows, considering the magnitude and importance of each term: $W_1=0.1$, $W_2=W_3=W_4=1$.

IV. RESULTS OF EXPERIMENT

The method of estimation described above was applied to hoarse voices /a/ uttered by fifteen patients with vocal cord polyp. Since video pictures of the vocal cords of those patients were available, it was possible to rather precisely measure dimensions of the polyps from the pictures taking into account a fact that the average length of the vocal cords is 14 mm for adult men and 10.5 mm for adult women. If the mass of polyp m_p is estimated, then on the simplifying assumptions that the polyp is shaped like a sphere and its density is known its diameter can be estimated from the mass m_p. Measured diameters of the polyps and corresponding estimates of the diameters obtained by the method of estimation are shown in Table I along with estimates of m_p and K_p.

For the case of typical polyp corresponding to the voice sample 14, waveforms of synthetic hoarse voice (sound pressure), glottal volume flow, glottal areas, and displacements of the constituent masses of the model of pathological vocal cords are demonstrated in Fig. 6. Similar waveforms were obtained for other voice samples too, though they are not shown here. Those waveforms were derived from the synthesizer by supplying it with values of all the necessary parameters including the estimates of m_p and K_p.

Note that for the voice sample 14 that was produced by vocal cords with a single, spherical polyp, the estimate of m_p is reasonably good in view of the actual diameter of the polyp. From waveforms derived from the voice samples 1, 3, and 14, each of which was produced by vocal cords with a single, spherical polyp, the following observations have been made: For all the voice samples the movement of the mass representing polyp is behind that of the upper mass coupled to the polyp and the displacement of the polyp is rather small compared with those of the upper and lower masses. For the voice samples 1 and 14 K_p is rather small, thus the upper mass is loosely coupled to the polyp, causing the polyp to vibrate irregularly beyond the plane x=0. In addition, the output sound pressure waveforms are found to be rather irregular for these cases, demonstrating that a single, spherical polyp has a great effect on the voice produced.

V. CONCLUSIONS

In this work a new noninvasive procedure for the diagnosis of vocal cord polyp through the use of a novel model of pathological vocal cords with a single polyp has been developed, which might well be called an analysis by synthesis approach. By this procedure it is possible not only to discriminate a hoarse voice caused by the vocal cord polyp from a normal voice but to estimate the state of the vibration of vocal cords through the use of the acoustic parameters, PPQ and APQ.

This procedure has been applied to several hoarse voice samples with results showing that it is capable of estimating the state of the vibration of vocal cords and the polyp features including the mass of polyp with a satisfactory accuracy for vocal cords with a single, large polyp. For small polyps, however, the estimates of the mass of polyp were found to be in error because of mismatch between the model and actual vocal cords. If the density of polyp were known, one could estimate the dimension, e.g., diameter, of polyp. Even though the dimension of polyp is not available, the estimate of the mass of polyp would be useful for the diagnosis of vocal cord polyp through hoarse voice analysis.

ACKNOWLEDGMENTS

The authors would like to thank Dr. Gohta Tsuda of Fukui Medical College, Fukui, Japan, for all his assistance and many helpful discussions. Thanks are also due to Mr. Hiroyasu Katoh of Toyota Motors Co. Ltd., Toyota, Japan and graduate student Kazutaka Hattori of Fukui University, Fukui, Japan, for their help with the computer simulation.

REFERENCES

[1] T. Koizumi, S. Taniguchi, and S. Hiromitsu, "Two-mass models of the vocal cords for natural sounding voice synthesis," J. Acoust. Soc. Am., Vo. 82, pp. 1179-1192, 1987.

[2] T. Koizumi and S. Taniguchi, "Analysis of hoarse voice applied to the diagnosis of laryngeal disease," Technical Report, the Institute of Electronics, Information and Communication Engineers, SP88-62, pp. 1-4, 1988.

[3] T. Koizumi and S. Taniguchi, "Analysis of hoarse voice applied to the diagnosis of laryngeal disease," J. Acoust. Soc. Am., Vo. 84, S84 (BB15), 1988.

[4] T. Koizumi and S. Taniguchi, "A novel model of pathological vocal cords and its application to the diagnosis of vocal cord polyp," Technical Report, the Institute of Electronics, Information and Communication Engineers, SP89-42, pp. 23-30, 1989.