Distinguishing Normal and Abnormal Tracheal Breathing Sounds by Principal Component Analysis

Michel John Mussell, Yoshimi Nakazono, Yoshimi Miyamoto, Shinichi Okabe,* and Tamotsu Takishima*

Department of Information, Faculty of Engineering, Yamagata University, Yonezawa, 992 Japan
*First Department of Internal Medicine, Tohoku University School of Medicine, Sendai, 980 Japan

Abstract Expired and inspired tracheal breathing sounds (BS) were recorded from 10 normal subjects and 8 patients with respiratory diseases, including bronchial asthma, sarcoidosis, fibrosing lung disease, chronic bronchitis, and radiation pneumonitis. Frequency spectra were generated using Fast Fourier Transform (FFT), and we observed considerable differences between BS spectra of normal subjects and patients. The frequency of peak amplitude and mean frequency of the BS spectra of patients were significantly higher than those of normal subjects. Spectral features were extracted by dividing each spectra into equal frequency bands—each feature being the mean amplitude of each FFT element within a frequency band. We used Principal Component Analysis to compare spectral feature sets and found a clear separation between normal and abnormal tracheal BS for 10, 20, and 40 features/spectra. We conclude that Principal Component Analysis of BS could become a new method of diagnosing respiratory disease in an automated fashion.

Key words: breathing sounds, lung disease, Principal Component Analysis.

Since Laennec invented the stethoscope in 1819, auscultation of breathing sounds (BS) has become a well-established method of diagnosing respiratory disease. However, it is subjective and limited by the sensitivity of human hearing. In this paper we report a more objective technique of recording and analyzing BS. BS recording goes back many decades but only in the past 10 years has advanced computing enabled more detailed analysis to be performed on BS.

The fundamental aims of BS research is to determine the origin of BS in relation to the underlying pathology of respiratory disease, and to develop BS analysis systems that can automate respiratory disease diagnosis. The key to

Received for publication June 27, 1990
automation of diagnosis is to be able to rapidly compare the relatively large amount of spectral data produced by BS recording. In this study, we report recording tracheal BS in normal subjects and patients with lung disease, and analyzing them using Principal Component Analysis (PCA). Our aim was to determine if the PCA technique is sensitive enough to distinguish between tracheal breathing sounds from patients and normal subjects.

SUBJECTS AND PATIENTS

For this study we recorded the BS of 10 normal subjects (students at Yamagata University who had no history of cardiopulmonary disease) and 8 patients who were undergoing treatment for a variety of lung diseases at Tohoku University Hospital, Sendai. All patients had abnormal spirometry and abnormal BS as defined by auscultation, and some had severe respiratory disease. Informed consent was obtained from all participants. The normal subjects were either 23 or 24 years of age and all but one were male. Table 1 gives the patient details.

RECORDING EQUIPMENT AND PROTOCOL

Figure 1 shows the BS recording equipment and tracheal microphone (Foster M268E00 electret condenser microphone). The microphone assembly was constructed from a piece of plastic tube (20 mm bore, 3 mm wall) with an aluminum end plate and foam supporting material. A hole (2 mm diameter) was drilled through the tube wall to equalize its inner pressure to that of atmospheric pressure when pressed against the trachea. The microphone was hand-held on the trachea of each subject or patient and produced a signal of around 5 mV peak-to-peak during normal quiet breathing, and around 25 mV during forced breathing. The signal was amplified with a gain of 100 using a low-noise battery-powered instrumentation amplifier.

Table 1. Patient details.

<table>
<thead>
<tr>
<th>Age</th>
<th>Sex</th>
<th>Disease</th>
<th>VC (l)</th>
<th>VC (%pred.)</th>
<th>FEV1 (l)</th>
<th>FEV1 (%pred.)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>21</td>
<td>M</td>
<td>Bronchial asthma</td>
<td>5.0</td>
<td>71</td>
<td>3.58</td>
</tr>
<tr>
<td>2</td>
<td>41</td>
<td>M</td>
<td>Bronchial asthma</td>
<td>4.48</td>
<td>115</td>
<td>3.48</td>
</tr>
<tr>
<td>3</td>
<td>59</td>
<td>F</td>
<td>Sarcoiosis</td>
<td>2.84</td>
<td>89</td>
<td>2.54</td>
</tr>
<tr>
<td>4</td>
<td>65</td>
<td>M</td>
<td>Fibrosing lung disease</td>
<td>2.11</td>
<td>86</td>
<td>1.4</td>
</tr>
<tr>
<td>5</td>
<td>80</td>
<td>M</td>
<td>Chronic bronchitis</td>
<td>2.82</td>
<td>100</td>
<td>1.59</td>
</tr>
<tr>
<td>6</td>
<td>72</td>
<td>F</td>
<td>Fibrosing lung disease</td>
<td>1.31</td>
<td>64</td>
<td>0.84</td>
</tr>
<tr>
<td>7</td>
<td>66</td>
<td>F</td>
<td>Fibrosing lung disease</td>
<td>1.39</td>
<td>89</td>
<td>1.24</td>
</tr>
<tr>
<td>8</td>
<td>61</td>
<td>M</td>
<td>Radiation pneumonitis</td>
<td>3.33</td>
<td>70</td>
<td>2.33</td>
</tr>
</tbody>
</table>

VC, vital capacity; FEV1, forced expired volume in 1 s; %pred., percentage of predicted value.

Japanese Journal of Physiology
During BS recording, the subjects breathed through a low resistance flow meter (Minato Medical Science Co., RF-2 respiratory flow meter); the flow signal drove a large center-zero analogue meter placed in front of the subject, so they could target their breathing maneuver at approximately 3 times normal resting inspiratory and expiratory flow levels. The object of the target flow meter was to have all subjects and patients breath at the same flow and with a similar respiratory maneuver. Some patients found it difficult to reach the target, and so did the best they could.

The flow signal and amplified microphone signal were recorded on 2 channels of an 8-track tape recorder (Sony, FE-3000), with a bandwidth of 5 kHz and signal-to-noise ratio of 50 dB per channel on its fastest tape speed of 19 cm/s. Several respiratory maneuvers were recorded from each patient and subject. Later, the recorded microphone signal was replayed through an anti-aliasing filter (2.2 kHz cut-off frequency) and instrumentation amplifier to increase the signal to ±5 V at the loudest point, and then digitized at 5 kHz (12-bit resolution). (Note: an anti-aliasing filter is a low-pass filter cutting off at half, or less than half, the sampling frequency to avoid erroneous sampling.) The flow signal was not filtered or amplified. Six-second lengths (containing at least one inspiration and expiration maneuver) of the two signals were stored on a floppy disk.

FREQUENCY SPECTRA

Figure 2 shows the mean expiratory spectra for all subjects and patients. Each spectra was computed using multiple 1,024-point Fast Fourier Transforms (FFT) with a cosine window function (i.e., time-shifted FFT averaging), as follows. The
first 1,024 points (i.e., 1,024/5,000 s) of the digitized expired microphone signal (identified from the expired flow signal) was FFT'ed, to produce a 512-point frequency spectra (0 Hz (d.c.) to 2.5 kHz). Then, after a shift of 200 data points, the next 1,024-point block (i.e., 200 to 1,224 points) was FFT'ed and the resultant spectra added to the first spectra. This process of shifting 200 points, FFT'ing a 1,024-point block, and accumulating the spectra was continued to the end of the expiratory portion of the BS. Each of the 512 accumulated elements from each spectra was divided by the number of FFT’s done to produce the mean expiratory frequency spectra. The relative-amplitude spectra in Fig. 2 were all normalized to have the same area under the curve as normal spectra 1.

This process was also repeated for the inspiratory portion of the digitized microphone signal. It was considered important to analyze expired and inspired breathing sounds separately because: (i) the resistance at the vocal cord changes with respiratory phase; and (ii) different sounds are produced by the differing respiratory phases by differing mechanisms (Hallgren et al., 1982; Charbonneau

*Japanese Journal of Physiology*
et al., 1983; Lessard and Wong 1986).

For an initial assessment of the frequency spectra, the frequency of maximum amplitude (FMA) and the mean frequency of the amplitude (MFA) spectra, were computed for the 18 expired and 18 inspired spectra (i.e., one inspired and one expired spectra for each of the subjects and patients). FMA is the frequency of peak amplitude, and MFA divides the spectra into two equal areas (i.e., the center of rotation of the spectra). They are the amplitude spectra equivalents of the power spectra parameters used by Schreiber et al. (1981) and Hallgren et al. (1982)—both are considered important descriptors of BS spectra. Non-paired t-tests show that both expired and inspired FMA for normal subjects (with means of 461 and 425 Hz respectively) are significantly lower than those for patients (with means of 751 and 645 Hz, respectively) ($p < 0.001$ for both expiration and inspiration). Similarly, both expired and inspired MFA for normal subjects (with means of 355 and 359 Hz, respectively) are significantly lower than those for patients (with means of 550 and 496 Hz, respectively) ($p < 0.001$ and $p < 0.05$ for both expiration and inspiration).

FEATURE EXTRACTION AND PRINCIPAL COMPONENT ANALYSIS

Ideally, to compare all the mean spectra in Fig. 2, the 512 elements of each spectra should be compared. However, this would require considerable computation time and power, so each spectra was simplified into a smaller number of “features” by “feature extraction.” The first 300 points of each spectra (i.e., d.c. to 1.46 kHz), in which most frequency components lie, was divided into equal frequency bands, and the amplitude of each band (a feature) was calculated as the mean amplitude of the frequency elements within the frequency band. Figure 3 illustrates this in one spectra. It shows a typical normal spectra (subject 3) divided into 8, 18, and 38 frequency bands (features). FMA and MFA were also used as features of each spectra, bringing the respective total number of features/spectra (denoted $M$ features/spectra) to $M = 10, 20, 40$.

The feature sets (frequency bands and FMA and MFA) for each spectra describe each spectra in a reasonably detailed, if simplified, form. To formally compare the 18 expiratory spectra (denoted $N$ spectra) by comparing their feature sets for $M = 10$ features/spectra, the features were grouped together as $N$ sets of $M$ features and input into a standard multi-variate PCA program (Tanaka et al., 1986). The same was repeated for $M = 20$ and 40, and also for inspiration for $M = 10, 20, 40$.

A detailed mathematical description of the standard PCA program is beyond the scope of this paper. Essentially, the $N$-by-$M$ array of spectral features are input into the PCA program, which outputs a reordered $N$-by-$M$ array—via computing the eigen numbers and vectors of the data and sorting them in the descending order of eigen numbers. In so doing, the PCA program determines the vector of most variance (with the highest eigen number), the vector of second-most variance, and so on, to the vector of least variance (with the lowest eigen number) in the data.
Fig. 3. The frequency banding method of extracting features from a typical mean expired spectra (normal subject 3)—the amplitude of each band is a feature. Labels F and M on the frequency scale denote the frequency of maximum amplitude (FMA) and mean frequency of the amplitude (MFA) spectra. The top trace shows an 8-band approximation to the spectral shape and, combined with FMA and MFA, has $M = 10$ features/spectra. The lower two traces have 18 and 38 frequency bands and combined with FMA and MFA have $M = 20$ and 40 features/spectra.

Most of the important (principal) variance in the data (i.e., 85% or more, for example) is contained within the vectors of most and second-most variance, and the other vectors can be discarded without significant loss of information; in practice this means reploting the original data using the vectors of most and second-most variance as 2-dimensional axes to produce a scatter diagram. This generates a good visual representation of the variance in all the data in a form humans can easily understand. Each point on the scatter diagram represents a feature set of each spectra (i.e., represents a spectra), and the relative positions of the points on the scatter diagram illustrates the relative differences or similarities between the original spectra.

*Japanese Journal of Physiology*
Fig. 4. Expiratory and inspiratory scatter plots for 10, 20, and 40 features/spectra produced by Principal Component Analysis. Dimensionless axes A and B are, respectively, the vectors of most and second-most variance in the data. Each point represents a single spectra, with normal and abnormal spectra denoted by × and ○, respectively, and the relative position of the points determine the degree of difference between spectra. Numbers against each point signify subject and patient numbers.
Thus, for example, if all the original spectra were identical, all the points on the scatter diagram would be in the same position. However, as the spectra become dissimilar, their respective points spread out on the scatter diagram.

The resulting expiratory and inspiratory scatter plots for $M=10$, 20, and 40 (features/spectra) for the 18 spectra are shown in Fig. 4; axis A is the vector of most variance in the data and axis B is the vector of second-most variance in the data—both are dimensionless. It is clear that for all values of $M$ used, the group of points representing normal spectra ($\times$) is well separated (with little overlap) from the group of points representing spectra of patients ($\bigcirc$). Also, as $M$ increases, the within-group scattering of points, and the separation between the normal and patient groups, becomes larger.

**DISCUSSION**

By visually inspecting Fig. 2, it is apparent that the spectra of patients have higher frequency components and a wider bandwidth than those of normal subjects. This is statistically illustrated by the fact that FMA and MFA of the patients' spectra are both significantly higher than FMA and MFA of the normal subjects' spectra, irrespective of respiratory phase. However, though comparison of spectra in this way, using just two features/spectra (i.e., FMA and MFA), is simple and easy, it is an over-simplification that disregards the complex shape of spectra, which is important when comparing BS spectra.

The frequency banding technique we have described (Fig. 3) was chosen so as to include information about the shape of each spectrum, and produces several features/spectra. Clearly, more features/spectra improve the description of a spectra, and the PCA data-reduction technique lends itself well to the situation when there are many features/spectra. We used 10, 20, and 40 features/spectra to investigate the effect that increasing the number of features/spectra has on the sensitivity of the PCA technique, when used for comparing BS spectra.

The PCA scatter diagrams (Fig. 4) are a clear indication that feature extraction (using both frequency banding and FMA and MFA as features) and the PCA technique are collectively sensitive enough to discriminate between tracheal BS spectra of normal subjects and patients with lung disease. With as few as $M=10$ features/spectra, separation is clearly visible, though separation improves with increased $M$, due to improved description of spectral shape with increasing $M$ (large $M$ may even enable the BS associated with various diseases to be separated, though that is beyond the scope of this paper since very large numbers of patients would be required for such a comparison). We, therefore, conclude that PCA, which as far as we know has not been applied to tracheal sound analysis before, is suitable for formally comparing BS spectra and may become the basis for a new automated method of respiratory disease diagnosis.

The way in which this analysis method can be applied to automate respiratory disease diagnosis is as follows. A reasonably large database of normal BS spectra.
(100 for example) would be constructed; they could be stored as spectra or sets of features. Databases of spectra of abnormal BS recorded from patients with various respiratory diseases would also be constructed, (i.e., one database of sounds from patients with fibrosing lung disease; one database of sounds from patients with emphysema; etc.). Accurate diagnosis of clinical status of each patient's chest condition would be required before any recording is made. A new BS recording (i.e., one to be diagnosed) could then be compared with all those in the database by the PCA method outlined in this paper, to produce a new point on the scatter diagram. If the new point appears within the cluster of points from normal subjects, the new recording is likely from a normal subject. Alternatively, if the new point appears within one of the abnormal groups of points then the new recording is likely from a subject with lung disease.

Our future aims are (1) to improve the hardware by by-passing the tape recorder and directly digitizing the signals, (2) to investigate the technique in sounds recorded over various points on the chest, and (3) to construct the above-mentioned databases of BS.

This project was supported by the Japan Society for the Promotion of Science (JSPS) under the Royal Society/JSPS exchange program.

We also thank Mrs. Yoko Kanazawa for preparing the figures for this paper.

REFERENCES


