Spectro-Temporal Analysis of Auscultatory Sounds

Tiago H. Falk[1], Wai-Yip Chan[2], Ervin Sejdić[1] and Tom Chau[1]
[1]*Bloorview Research Institute/Bloorview Kids Rehab and the Institute of Biomaterials and Biomedical Engineering,*
University of Toronto, Toronto, Canada
[2]*Department of Electrical and Computer Engineering,*
Queen's University, Kingston, Canada

1. Introduction

Auscultation is a useful procedure for diagnostics of pulmonary or cardiovascular disorders. The effectiveness of auscultation depends on the skills and experience of the clinician. Further issues may arise due to the fact that heart sounds, for example, have dominant frequencies near the human threshold of hearing, hence can often go undetected (1). Computer-aided sound analysis, on the other hand, allows for rapid, accurate, and reproducible quantification of pathologic conditions, hence has been the focus of more recent research (e.g., (1–5)). During computer-aided auscultation, however, lung sounds are often corrupted by intrusive quasi-periodic heart sounds, which alter the temporal and spectral characteristics of the recording. Separation of heart and lung sound components is a difficult task as both signals have overlapping frequency spectra, in particular at frequencies below 100 Hz (6).

For lung sound analysis, signal processing strategies based on conventional time, frequency, or time-frequency signal representations have been proposed for heart sound cancelation. Representative strategies include entropy calculation (7) and recurrence time statistics (8) for heart sound detection-and-removal followed by lung sound prediction, adaptive filtering (e.g., (9; 10)), time-frequency spectrogram filtering (11), and time-frequency wavelet filtering (e.g., (12–14)). Subjective assessment, however, has suggested that due to the temporal and spectral overlap between heart and lung sounds, heart sound removal may result in noisy or possibly "non-recognizable" lung sounds (15). Alternately, for heart sound analysis, blind source extraction based on periodicity detection has recently been proposed for heart sound extraction from breath sound recordings (16); subjective listening tests, however, suggest that the extracted heart sounds are noisy and often unintelligible (17).

In order to benefit fully from computer-aided auscultation, both heart *and* lung sounds should be extracted or blindly separated from breath sound recordings. In order to achieve such a difficult task, a few methods have been reported in the literature, namely, wavelet filtering (18), independent component analysis (19; 20), and more recently, modulation domain filtering (21). The motivation with wavelet filtering lies in the fact that heart sounds contain large components over several wavelet scales, while coefficients associated with lung sounds quickly decrease with increasing scale. Heart and lung sounds are iteratively separated based on an adaptive hard thresholding paradigm. As such, wavelet coefficients at each scale with amplitudes above the threshold are assumed to correspond to heart sounds and the remaining coefficients are associated with lung sounds. Independent component analysis, in turn, makes use

of multiple breath sound signals recorded at different locations on the chest to solve a blind deconvolution problem. Studies have shown, however, that with independent component analysis lung sounds can still be heard from the separated heart sounds and vice-versa (20).

Modulation domain filtering, in turn, relies on a spectro-temporal signal representation obtained from a frequency decomposition of the temporal trajectories of short-term spectral magnitude components. The representation measures the rate at which spectral components change over time and can be viewed as a frequency-frequency signal decomposition often termed "modulation spectrum." The motivation for modulation domain filtering lies in the fact that heart and lung sounds are shown to have spectral components which change at different rates, hence increased separability can be obtained in the modulation spectral domain. In this chapter, the spectro-temporal signal representation is described in detail. Spectro-temporal signal analysis is shown to result in fast yet accurate heart and lung sound signal separation without the introduction of audible artifacts to the separated sound signals. Additionally, adventitious lung sound analysis, such as wheeze and stridor detection, is shown to benefit from modulation spectral processing.

The remainder of the chapter is organized as follows. Section 2 introduces the spectro-temporal signal representation. Blind heart and lung sound separation based on modulation domain filtering is presented in Section 3. Adventitious lung sound analysis is further discussed in Section 4.

2. Spectro-Temporal Signal Analysis

Spectro-temporal signal analysis consists of the frequency decomposition of *temporal trajectories* of short-term signal spectral components, hence can be viewed as a frequency-frequency signal representation. The signal processing steps involved are summarized in Fig. 1. First, the source signal is segmented into consecutive overlapping frames which are transformed to the frequency domain via a base transform (e.g., Fourier transform). Frequency components are aligned in time to form the conventional time-frequency representation. The magnitude of each frequency bin is then computed and a second transform, termed a modulation transform, is performed across time for each individual magnitude signal. The resulting modulation spectral axis contains information regarding the rate of change of signal spectral components. Note that if invertible transforms are used and phase components are kept unaltered, the original signal can be perfectly reconstructed (22). Furthermore, to distinguish between the two frequency axes, frequency components obtained from the base transform are termed "acoustic" frequency and components obtained from the modulation transform are termed "modulation" frequency (23).

Spectro-temporal signal analysis (also commonly termed modulation spectral analysis) has been shown useful for several applications involving speech and audio analysis. Clean speech was shown to contain modulation frequencies ranging from 2 Hz - 20 Hz (24; 25) and due to limitations of the human speech production system, modulation spectral peaks were observed at approximately 4 Hz, corresponding to the syllabic rate of spoken speech. Using such insights, robust features were developed for automatic speech recognition in noisy conditions (26), modulation domain based filtering and bandwidth extension were proposed for noise suppression (27), the detection of significant modulation frequencies above 20 Hz was proposed for objective speech quality measurement (28) and for room acoustics characterization (29), and low bitrate audio coders were developed to exploit the concentration of modulation spectral energy at low modulation frequencies (22). Alternate applications include classification of acoustic transients from sniper fire (30), dysphonia recognition (31), and rotating
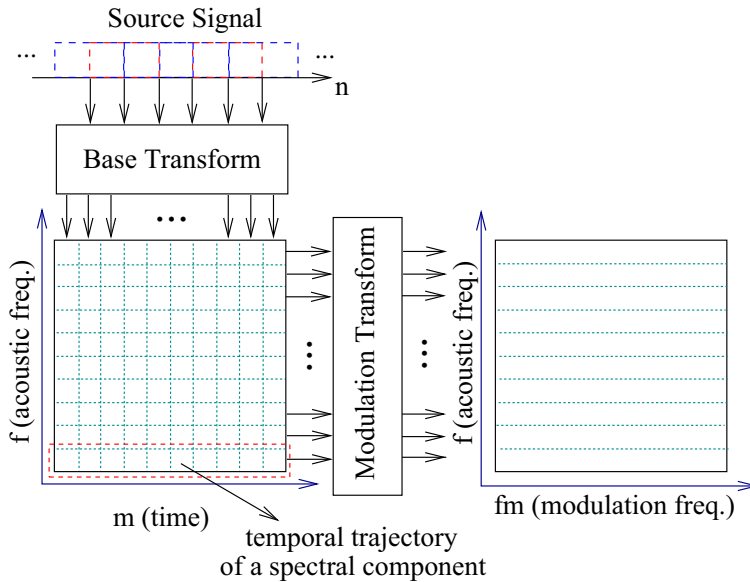
Fig. 1. Processing steps for spectro-temporal signal analysis

machine classification (32). In the sections to follow, two novel biomedical signal applications are described, namely, blind separation of heart and lung sounds from computer-based auscultation recordings and pulmonary adventitious sound analysis.

3. Blind Separation of Heart and Lung Sounds

Heart and lung sounds are known to contain significant and overlapping acoustic frequencies below 100 Hz. Due to the nature of the two signals, however, it is expected that the spectral content of the two sound signals will change at different rates, thus improved separability can be attained in the modulation spectral domain. Preliminary experiments were conducted with breath sounds recorded in the middle of the chest at a low air flow rate of 7.5 ml/s/kg to emphasize heart sounds and in the right fourth interspace at a high air flow rate 22.5 ml/s/kg to emphasize lung sounds. Lung sounds are shown to have modulation spectral content up to 30 Hz modulation frequency with more prominent modulation frequency content situated at low frequencies (< 2 Hz), as illustrated in Fig. 2 (a). This behavior is expected due to the white-noise like properties of lung sounds (33) modulated by a slow on-off (inhale-exhale) process. Heart sounds, on the other hand, can be considered quasi-periodic and exhibit prominent harmonic modulation spectral content between approximately 2-20 Hz; this is illustrated in Fig. 2 (b). As can be observed, both sound signals contain important and overlapping acoustic frequency content below 100 Hz; the modulation frequency axis, however, introduces an additional dimension over which improved separability can be attained. As a consequence, modulation filtering has been proposed for blind heart and lung sound separation (21).
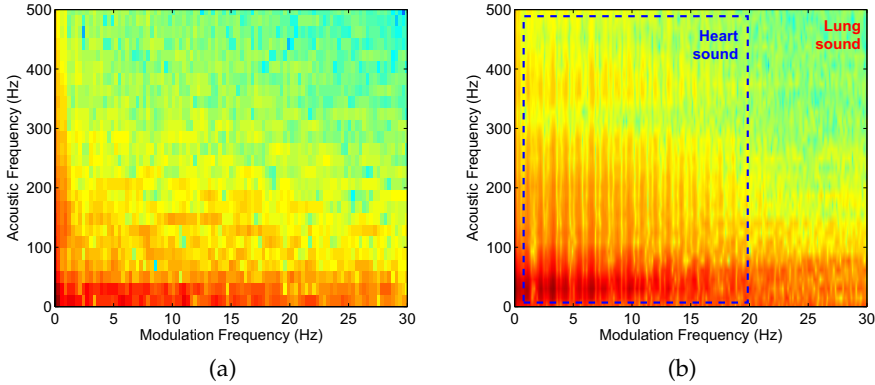
Fig. 2. Spectro-temporal representation of a breath sound recorded at (a) the right fourth interspace at a high air flow rate to emphasize lung sounds, and (b) the middle of the chest at a low air flow rate to emphasize heart sounds. Modulation spectral plots are zoomed in to depict acoustic frequencies below 500 Hz and modulation frequencies below 30 Hz.

3.1 Modulation Domain Filtering

Modulation filtering is described as filtering of the temporal trajectories of short-term spectral components. Two finite impulse response modulation filters are employed and depicted in Fig. 3. The first is a bandpass filter with cutoff modulation frequencies at 1 Hz and 20 Hz (dotted line); the second is the complementary bandstop filter (solid line). Modulation frequencies above 20 Hz are kept as they are shown to improve the naturalness of separated lung sound signals. In order to attain accurate resolution at 1 Hz modulation frequency, higher order filters are needed. Here, 151-tap linear phase filters are used; such filter lengths are equivalent to analyzing 1.5 s temporal trajectories.

For the sake of notation, let $s(f, m)$, $f = 1, \ldots, N$ and $m = 1, \ldots, T$, denote the short-term spectral component at the fth frequency bin and m^{th} time step of the short-term analysis. N and T denote total number of frequency bands and time steps, respectively. For a fixed frequency band $f = F$, $s(F, m), m = 1, \ldots, T$, represents the Fth band temporal trajectory. In the experiments described herein, the Gabor transform is used for spectral analysis. The Gabor transform is a unitary transform (energy is preserved) and consists of an inner product with basis functions that are windowed complex exponentials. Doubly over-sampled Gabor transforms are used and implemented based on discrete Fourier transforms (DFT), as depicted in Fig. 4.

First, the breath sound recording is windowed by a power complementary square-root Hann window of length 20 milliseconds with 50% overlap (frame shifts of 10 milliseconds). An N-point DFT is then taken and the magnitude ($|s(f, m)|$) and phase ($\angle s(f, m)$) components of each frequency bin are input to a "modulation processing" module where modulation filtering and phase delay compensation are performed. The "per frequency bin" magnitude trajectory $|s(f, m)|, m = 1, \ldots, T$ is filtered using the bandpass and the bandstop modulation filters to generate signals $|\hat{s}(f, m)|$ and $|\tilde{s}(f, m)|$, respectively. The remaining modulation processing step consists of delaying the phase by 75 samples, corresponding to the group delay of the implemented linear phase filters. The outputs of the modulation processing modules are the
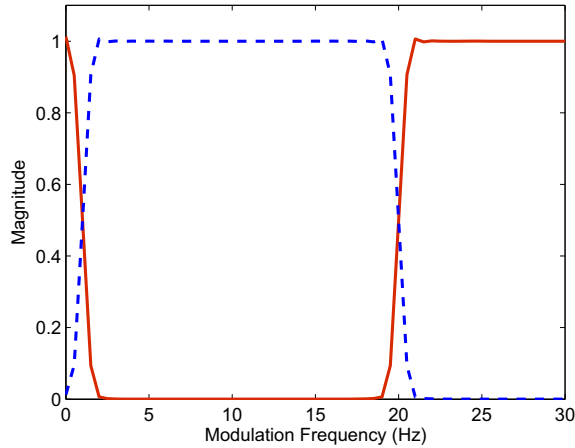
Fig. 3. Magnitude response of bandpass (dotted line) and bandstop (solid line) modulation filters.

bandpass and bandstop filtered signals and the delayed phase components $\angle\bar{s}(f,m)$. Two N-point IDFTs are then taken. The first IDFT (namely IDFT-1) takes as input the N $|\hat{s}(f,m)|$ and $\angle\bar{s}(f,m)$ signals to generate $\hat{s}(m)$. Similarly, IDFT-2 takes as input signals $|\tilde{s}(f,m)|$ and $\angle\bar{s}(f,m)$ to generate $\tilde{s}(m)$. The outputs of the IDFT-1 and IDFT-2 modules are windowed by the power complementary window and overlap-and-add is used to reconstruct heart and lung sound signals, respectively. The description, as depicted in Fig. 4, is conceptual and the implementation used here exploits the conjugate symmetry properties of the DFT to reduce computational complexity by approximately 50%.

It is observed that with bandpass filtered modulation envelopes the removal of lowpass modulation spectral content may result in negative power spectral values. As with the spectral subtraction paradigm used in speech enhancement algorithms, a half-wave rectifier can be used. Rectification, however, may introduce unwanted perceptual artifacts to the separated heart sound signal. To avoid such artifacts, one can opt to filter the cubic-root compressed magnitude trajectories in lieu of the magnitude trajectories. In such instances, cubic power expansion must be performed prior to taking the IDFT. In the experiments described herein, cubic compression-expansion of bandpass filtered signals is used and negligible rectification activation rates ($<2\%$) are obtained.

3.2 Database of Breath Sound Recordings

The University of Manitoba breath sound recordings are used in the experiments; the data has been made publicly available by the Biomedical Engineering Laboratory. Data is obtained from two healthy subjects aged 25 and 30 years on three separate occasions (20). Piezoelectric contact accelerometers were used to record the respiratory sounds from the subjects in sitting position. Accelerometers were secured with double-sided adhesive tape rings at the following five locations: (1) right and (2) left midclavicular, 2nd intercostal space, (3) right and (4) left 4th intercostal space, and (5) center of chest.
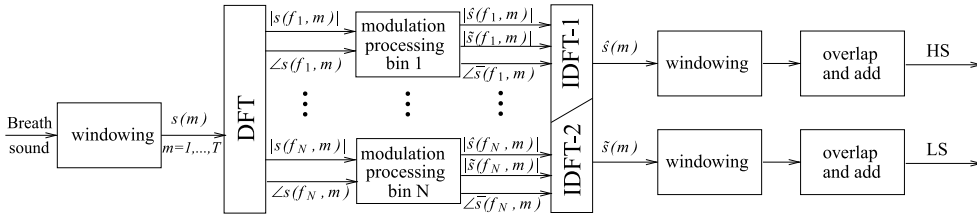
Fig. 4. Block diagram of the modulation filtering approach for blind separation of heart sounds (HS) and lung sounds (LS) from auscultation recordings.

Subjects were asked to maintain their target breathing at low (7.5 ml/s/kg), medium (15 ml/s/kg), and high (22.5 ml/s/kg) flow rates. Subjects were instructed to breathe such that one full breath occurred every two to three seconds at every flow rate and had at least five breaths at each target flow. Three recordings were made per subject and each recording consisted of approximately 20 s at each target flow and concluded with an approximate 5 s of breath hold (total of $\sim$ 65 s). During breath hold, subjects were asked to hold their breath with a closed glottis, thus allowing for a reference heartbeat signal and background noise characterization. Breath sound signals were digitized with 10240 Hz sample rate and 16-bit precision. In our experiments, data is downsampled to 5 kHz in order to reduce computational complexity.

3.3 Benchmark Separation Algorithm

For comparison purposes, a wavelet-based heart and lung sound separation algorithm is used as a benchmark (18); the reader is referred to the following references for a complete description of the algorithm: (18; 34; 35). In the experiments described herein, the threshold used was given by the standard deviation of the wavelet coefficients multiplied by a constant multiplicative factor. As suggested in a previous study (14), values used for the multiplicative factor range from $2.5 - 3.0$ (increments of 0.25) for breath sound segments of low, medium, and high air flow rates, respectively. With the wavelet filtering algorithm, heart and lung sound separation is achieved through an iterative reconstruction-decomposition process. The stopping criterion is set such that the error between two consecutive reconstruction steps drops below 10^{-5} (18).

3.4 Comparative Performance Analysis

Modulation domain and wavelet filtering algorithms are tested on breath sound signals captured at the five locations described in Section 3.2. The plots in Fig. 5 (a)-(b) illustrate short segments of separated heart and lung sounds, respectively, for both systems using signals recorded at the center of the chest during low air flow. Spectral plots of separated signals are further depicted in Fig. 6. Subplot (a) illustrates the spectra of "heart-sound-free" breath sounds and the separated lung sound signals processed by the modulation domain and wavelet filtering algorithms. Power spectra are averaged over 5 s of heart-sound-free breath sounds, which were randomly selected from segments of the breath sound recording between successive heartbeats (selected segments were within $\pm20\%$ of the target low airflow rate). Similarly, subplot (b) depicts average power spectra of breath-hold sounds and the sep-
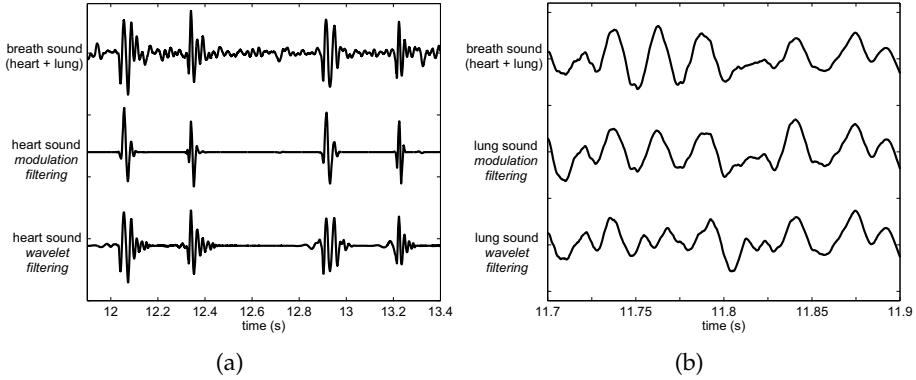
Fig. 5. Breath sound signals (top) and separated (a) heart sounds and (b) lung sounds using modulation domain based filtering (center) and wavelet based filtering (bottom). Subplot (a) depicts a pair of first and second (S1-S2) heart sound tones.

arated heart sound signal. Power spectra are averaged over the approximate 5 s breath-hold duration at the end of the recording session.
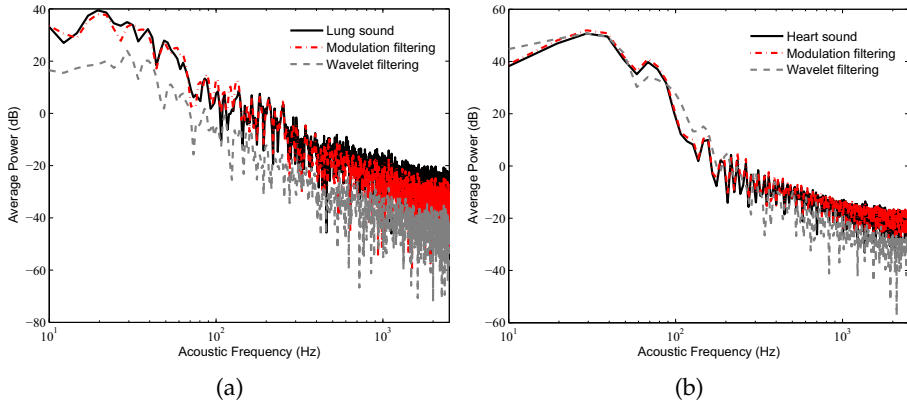

Fig. 6. Spectral plots of breath sounds and (a) separated lung sound and (b) heart sound signals.

In order to quantitatively assess the performance of the blind separation methods, the average log-spectral distance (*LSD*) between the aforementioned breath sound spectra $P(\omega)$ and separated signal spectra $\hat{P}(\omega)$ is used. The *LSD*, expressed in decibel, is given by

$$LSD = \sqrt{\frac{1}{2\pi} \int_{-\omega}^{\omega} \left[10 \log_{10} \frac{P(\omega)}{\hat{P}(\omega)} \right]^2 d\omega}. \tag{1}$$

Filtering	LSD (dB)		Processing time (s)
method	Heart	Lung	
Modulation	0.61 ± 0.13	0.79 ± 0.19	2.44 ± 0.04
Wavelet	1.11 ± 0.21	1.26 ± 0.56	67.2 ± 20.86

Table 1. Log-spectral distances (LSD) and algorithm processing times for wavelet and modulation domain based filtering. Performance metrics are reported as mean $\pm$ standard deviation over the two participants and six recording sessions.

Table 1 reports LSD values obtained for wavelet and modulation domain filtering averaged over the two participants and six recording sessions. In speech coding research, two signals with $LSD < 1$ dB are considered to be perceptually indistinguishable (36). Using this same difference limen for spectral transparency, results in Table 1 suggest that audible artifacts are not introduced by modulation domain filtering; this is corroborated by subjective listening tests conducted with three listeners. For wavelet filtering, however, listeners reported that lung sounds could still be heard in heart sounds and vice versa; such finding is expected given the LSD values greater than unity reported in the table.

Execution time is also an important metric to gauge algorithm performance. Both blind separation algorithms have been implemented using Matlab version 7.6 Release 2008a and simulations were run on a PC with a 2.2 GHz Dual Core processor and 3 GB of RAM. The execution times for heart and lung sound separation, averaged over the five recorded 65 s breath sound signals, are also reported in Table 1. As observed, the computational load of the modulation filtering method is one order of magnitude lower relative to wavelet filtering (approximately 30 times lower processing time). Moreover, with modulation domain filtering, if only the bandstop filter is applied (akin to heart sound cancelation) algorithm processing time can be further decreased by a factor of 1.5. As can be seen, modulation domain filtering allows for fast, yet accurate separation of heart and lung sounds from auscultatory recordings. Separated signals are also shown to be artifact-free, an important factor for accurate clinical diagnosis. In the section to follow, an additional application is presented and shown to also benefit from spectro-temporal signal analysis.

4. Adventitious Lung Sound Analysis

Adventitious lung sounds refer to abnormal sounds present in conjunction with the normal lung sound component (37). Adventitious lung sounds often signal abnormalities in pulmonary conditions (33); representative sounds can include crackles, wheezes, and stridor. Crackles are also referred to as discontinuous sounds as they are brief (in the order of tens of milliseconds) and intermittent. Crackles are caused by fluid obstruction of the small airways often due to inflammation of the bronchi. Crackle sounds, due to their short-term characteristics, are difficult to analyze via spectro-temporal signal processing; crackles have, however, been successfully analyzed via time-frequency wavelet processing (34; 38). Wheezes and stridor, on the other hand, have longer-term behavior that can extend to more than 250 milliseconds (33), hence can be analyzed using modulation spectral analysis.

Wheezes commonly occur in patients with obstructed airways and can have acoustic frequency components ranging from 100 Hz to 1 kHz (33; 39). Wheezes are characterized by
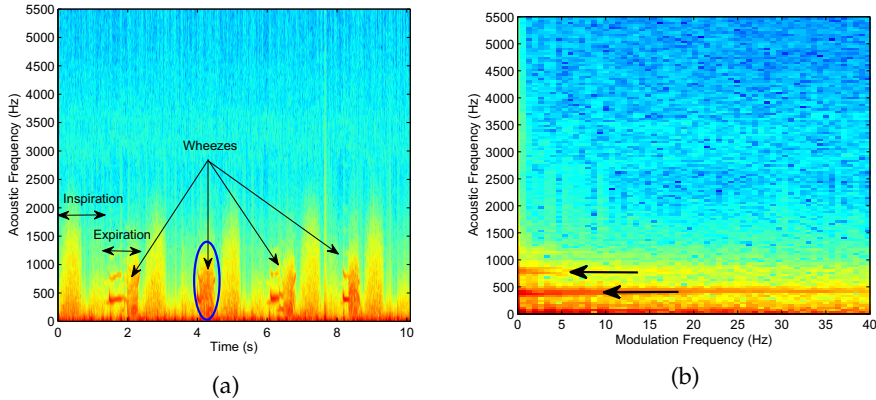
Fig. 7. Subplot (a): spectrogram of a breath sound recording with adventitious wheeze sounds indicated by arrows. Subplot (b) depicts the modulation spectrum of the approximate 0.25 s region highlighted by the ellipsoid in subplot (a).

high-pitched, musical tones manifested most prominently during expiration. Wheezes can be classified as "monophonic" or "polyphonic," if single or multiple tones are present, respectively. The perception and quantification of such properties is difficult if done subjectively via auscultation (39), hence automated methods based on spectrogram processing have been proposed (40). Figure 7 (a) depicts the spectrogram of a breath sound recording with expiratory wheezing taken from the R.A.L.E repository (41). As can be seen, tones are visible during expiration at frequencies around 400 Hz and 800 Hz and such tones are detectable with spectrogram-based methods. The two tones, however, are not easily detectable during the second respiration cycle highlighted by an ellipsoid in Fig. 7 (a). With the use of modulation spectral analysis, the two tones can be easily detected as illustrated by the arrows in Fig. 7 (b), thus can be used to assist inexperienced physicians in detecting pulmonary disorders.

Stridor, in turn, is characterized by a harsh, vibratory noise typically heard during inspiration. Stridor is caused by partial obstruction of the upper airway resulting in turbulent airflow. Figure 8 (a) depicts the spectrogram of a breath sound recording with stridor adventitious sounds taken from the R.A.L.E repository (41). Significant energy is observed at higher acoustic frequencies and tonal sounds can be seen at approximately 200 Hz and in some breath cycles at 1000 Hz. During the last breath cycle, however, the tonal components are not easily observable using spectrogram analysis. The two tonal components, however, are observable using modulation spectral analysis, as depicted by the arrows in Fig. 8 (b).

5. Conclusion

This chapter describes a spectro-temporal signal representation which is shown to be a useful tool for automatic auscultatory sound analysis. The representation, commonly termed "modulation spectrum," measures the rate at which breath sound spectral components change over time. The signal representation is successfully applied to blind heart and lung sound separation and shown to outperform state-of-the-art wavelet filtering both in terms of algorithm
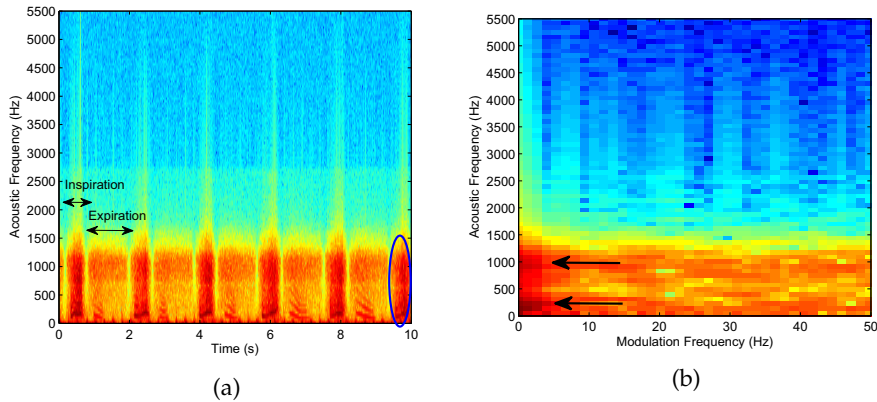
(a)

(b)

Fig. 8. Subplot (a): spectrogram of a breath sound recording with adventitious stridor sounds during inspiration. Subplot (b) depicts the modulation spectrum of the approximate 0.25 s region highlighted by the ellipsoid in subplot (a).

execution time and in separation performance. An alternate application in which the modulation spectrum can be applied, namely, adventitious lung sound detection, is also described.

6. References

[1] R. L. Watrous, "Computer-aided auscultation of the heart: From anatomy and physiology to diagnostic decision support," in *Proc. IEEE Conference of the Engineering in Medicine Biology Society*, 2006, pp. 140–143.

[2] Z. Syed, D. Leeds, D. Curtis, F. Nesta, R. A. Levine, and J. Guttag, "A framework for the analysis of acoustical cardiac signals," *IEEE Trans. on Biomedical Engineering*, vol. 54, no. 4, pp. 651–662, 2007.

[3] R. Murphy, "Computerized multichannel lung sound analysis: Development of acoustic instruments for diagnosis and management of medical conditions," *IEEE Engineering in Medicine and Biology Magazine*, vol. 26, pp. 16–19, 2007.

[4] C.-J. Hou, Y.-T. Chen, L.-C. Hu, C.-C. Chuang, Y.-H. Chiu, and M.-S. Tsai, "Computer-aided auscultation learning system for nursing technique instruction," in *Proc. IEEE Conference of the Engineering in Medicine Biology Society*, 2008, pp. 1575–1578.

[5] A. Marshall and S. Boussakta, "Signal analysis of medical acoustic sounds with applications to chest medicine," *Journal of the Franklin Institute*, vol. 344, no. 3-4, pp. 230–242, 2007.

[6] H. Pasterkamp, R. Fenton, A. Tal, and V. Chernick, "Interference of cardiovascular sounds with phonopneumography in children," *American Review of Respiratory Disease*, vol. 131, no. 1, pp. 61–64, Jan. 1985.

[7] A. Yadollahi and Z. Moussavi, "A robust method for heart sounds localization using lung sounds entropy," *IEEE Trans. on Biomedical Engineering*, vol. 53, no. 3, pp. 497–502, March 2006.

[8] C. Ahlstrom, O. Liljefeldt, P. Hult, and P. Ask, "Heart sound cancellation from lung sound recordings using recurrence time statistics and nonlinear prediction," *IEEE Signal Processing Letters*, vol. 12, no. 12, pp. 812–815, Dec. 2005.

[9] L. Hadjileontiadis and S. Panas, "Adaptive reduction of heart sounds from lung sounds using fourth-order statistics," *IEEE Trans. on Biomedical Engineering*, vol. 44, no. 7, pp. 642–648, July 1997.

[10] T. Tsalaile and S. Sanei, "Separation of heart sound signal from lung sound signal by adaptive line enhancement," in *Proc. European Signal Processing Conference*, 2007, pp. 1231–1234.

[11] M. Pourazad, Z. Moussavi, and G. Thomas, "Heart sound cancellation from lung sound recordings using time-frequency filtering," *Journal of Medical and Biological Engineering and Computing*, vol. 44, no. 3, pp. 216–225, March 2006.

[12] D. Flores-Tapia, Z. Moussavi, and G. Thomas, "Heart sound cancellation based on multi-scale products and linear prediction," *IEEE Trans. on Biomedical Engineering*, vol. 54, no. 2, pp. 234–243, Feb. 2007.

[13] S. Charleston and M. Azimi-Sadjadi, "Multi-resolution joint time delay and signal estimation for processing lung sounds," in *Proc. IEEE Conference of the Engineering in Medicine Biology Society*, 1995, pp. 985–Ű986.

[14] I. Hossain and Z. Moussavi, "An overview of heart-noise reduction of lung sound using wavelet transform based filter," in *Proc. IEEE Conference of the Engineering in Medicine Biology Society*, 2003, pp. 458–461.

[15] J. Gnitecki, I. Hossain, H. Pasterkamp, and Z. Moussavi, "Qualitative and quantitative evaluation of heart sound reduction from lung sound recordings," *IEEE Trans. on Biomedical Engineering*, vol. 52, no. 10, pp. 1788–1792, Oct. 2005.

[16] T. Tsalaile, S. Naqvi, K. Nazarpour, S. Sanei, and J. Chambers, "Blind source extraction of heart sound signals from lung sound recordings exploiting periodicity of the heart sound," in *Proc. International Conference on Audio, Speech, and Signal Processing*, 2008, pp. 461–464.

[17] T. Tsalaile, R. Sameni, S. Sanei, C. Jutten, and J. Chambers, "Sequential blind source extraction for quasi-periodic signals with time-varying period," *IEEE Trans. on Biomedical Engineering*, vol. 56, no. 3, pp. 646–655, 2009.

[18] L. J. Hadjileontiadis and S. M. Panas, "Separation of discontinuous adventitious sounds from vesicular sounds using a wavelet-based filter," *IEEE Trans. on Biomedical Engineering*, vol. 44, no. 12, pp. 1269–1281, 1997.

[19] J.-C. Chien, M.-C. Huang, Y.-D. Lin, and F.-C. Chong, "A study of heart sound and lung sound separation by independent component analysis technique," in *Proc. IEEE Conference of the Engineering in Medicine Biology Society*, Sept. 2006, pp. 5708–5711.

[20] M. Pourazad, Z. Moussavi, F. Farahmand, and R. Ward, "Heart sounds separation from lung sounds using independent component analysis," in *Proc. IEEE Conference of the Engineering in Medicine Biology Society*, Sept. 2005.

[21] T. Falk and W.-Y. Chan, "Modulation filtering for heart and lung sound separation from breath sound recordings," in *Proc. IEEE Conference of the Engineering in Medicine Biology Society*, Aug. 2008, pp. 1859–1862.

[22] M. Vinton and L. Atlas, "A scalable and progressive audio codec," in *Proc. International Conference on Audio, Speech, and Signal Processing*, May 2001, pp. 3277–3280.

[23] L. Atlas and S. Shamma, "Joint acoustic and modulation frequency," *EURASIP Journal on Applied Signal Processing*, vol. 7, p. 668Ű675, 2003.

[24] R. Drullman, J. Festen, and R. Plomp, "Effect of reducing slow temporal modulations on speech reception," *Journal of the Acoustical Society of America*, vol. 95, no. 5, pp. 2670–2680, May 1994.

[25] R. Drullman, J. Festen, and R. Plomp, "Effect of temporal envelope smearing on speech reception," *Journal of the Acoustical Society of America*, vol. 95, no. 2, pp. 1053–1064, Feb. 1994.

[26] H. Hermansky and N. Morgan, "RASTA processing of speech," *IEEE Trans. on Speech and Acoustics*, vol. 2, pp. 587–589, October 1994.

[27] T. H. Falk, S. Stadler, W. B. Kleijn, and W.-Y. Chan, "Noise suppression based on extending a speech-dominated modulation band," in *Proc. International Conference on Spoken Language Processing (Interspeech)*, 2007, pp. 970–973.

[28] D.-S. Kim, "A cue for objective speech quality estimation in temporal envelope representation," *IEEE Signal Processing Letters*, vol. 11, no. 10, pp. 849–852, Oct. 2004.

[29] T. H. Falk and W.-Y. Chan, "Temporal dynamics for blind measurement of room acoustical parameters," *IEEE Trans. on Instrumentation and Measurement*, 2009, in press, (12 pages).

[30] L. Owsley, L. Atlas, and C. Heinemann, "Use of modulation spectra for representation and classification of acoustic transients from sniper fire," in *Proc. International Conference on Audio, Speech, and Signal Processing*, 2005, pp. 1129–1133.

[31] N. Malyska, T. Quatieri, and D. Sturim, "Automatic dysphonia recognition using biologically-inspired amplitude-modulation features," in *Proc. International Conference on Audio, Speech, and Signal Processing*, 2005, pp. 873–876.

[32] S. Sukittanon, L. E. Atlas, and S. G. Dame, "Enhanced modulation spectrum using space-time averaging for in-building acoustic signature identification," in *Proc. International Conference on Audio, Speech, and Signal Processing*, 2006, pp. 153–156.

[33] H. Pasterkamp, S. Kraman, and G. Wodicka, "Respiratory sounds: Advances beyond the stethoscope," *American Journal of Respiratory and Critical Care Medicine*, vol. 156, pp. 974–987, 1997.

[34] L. Hadjileontiadis and S. Panas, "A wavelet-based reduction of heart sound noise from lung sounds," *International Journal of Medical Informatics*, vol. 52, pp. 183–190, 1998.

[35] J. Gnitecki and Z. Moussavi, "Separating heart sounds from lung sounds," *IEEE Engineering in Medicine and Biology Magazine*, vol. 26, pp. 20–29, Jan./Feb. 2007.

[36] W. B. Kleijn and K. K. Paliwal, Eds., *Speech Coding and Synthesis*. Elsevier, 1995.

[37] S. Lehrer, *Understanding lung sounds*. W.B. Saunders, 1993.

[38] S. Selloa, S. kyung Strambib, G. D. Michelea, and N. Ambrosinob, "Respiratory sound analysis in healthy and pathological subjects: A wavelet approach," *Biomedical Signal Processing and Control*, vol. 3, pp. 181–191, 2008.

[39] J. A. Fiz, R. Jane, A. Homs, J. Izquierdo, M. A. Garcia, and J. Morera, "Detection of wheezing during maximal forced exhalation in patients with obstructed airways," *Chest*, vol. 122, pp. 186–191, 2002.

[40] S. Taplidou, L. Hadjileontiadis, T. Penzel, V. Gross, and S. Panas, "WED: An efficient wheezing-episode detector based on breath sounds spectrogram analysis," in *Proc. IEEE Conference of the Engineering in Medicine Biology Society*, 2003, pp. 2531–2534.

[41] "R.A.L.E. repository." Online: http://www.rale.ca