

Automatic Cycle Identification in Speaker's Speech Breathing Signals

Y. Ding

Department of Electrical Engineering and Computer Science
Wichita State University
US

G. Tchankwe

Department of Electrical Engineering and Computer Science
Wichita State University
US

D.F. Parham

Department of Communication Sciences and Disorders
College of Health Professions, Wichita State University
US

Y.H. Ding

School of Electronics Information Engineering, Tianjin Key
Laboratory of Film Electronic and Communication Devices
Tianjin University of Technology
Tianjin, China

Abstract--Due to the complexity in respiratory signals, an experienced coder is often required to identify the acceptable cycles and mark the start and end times for each cycle. This algorithm is an important step toward timely identification and coding of more complex respiratory signals and more efficient analysis for physiological signals. Simulation results have shown that the proposed algorithm can identify cycles correctly and efficiently.

Keywords-respiration; speech breathing signal; cycle detection

I INTRODUCTION

Analysis of the speech breathing signals typically involves automatic cycle detections in the waveforms [1]. Although some algorithms are available to isolate specific frequencies of interest in a given waveform [2, 3], in the case of physiological behavior, rate measurements are often used as summary statistics of the waveforms. Unfortunately, cycle-averaging techniques are often inadequate when it comes to describing complex behaviors, particularly when that behavior changes over time, e.g., breaching behaviors during speech [4]. Consequently, an experienced coder who can evaluate each cycle of a given waveform is often required to determine which cycles are acceptable for analysis and manually separate cycles. It is time consuming to separate cycles by human visual coding and measurements. A previous study in [5] provided an algorithm capable of automatically identifying cycles in tidal breath signals. The algorithm used filtering and other signal processing techniques via Matlab programming [6], producing the exact start and end times for each cycle and isolating the distorted cycles due to artifacts, such as motion. It also isolated those portions of the signals negatively impacted by the artifacts. As a continuation of the previous study, this paper extends the cycle identification algorithm to speech signals and automatically detected speech breathing cycles across a variety of humans. Simulations have shown that the proposed algorithm identifies the breath cycles correctly and efficiently, despite the complexity of speech breathing signals. The algorithm will help researchers and clinicians identify and analyze cycles associated with speech breathing in much less time than if a human coder is used.

II SPEECH BREATHING SIGNALS AND DATA COLLECTION SCHEMES

Speech breathing contains a two-phase mechanical process. During the inspiratory phase, passive mechanical forces and active muscle forces draw air into the lungs. During the expiratory phase, passive and active forces interact by sending the tracheal air through the larynx, thus producing the fluid stream needed for phonation. To deliver effective speeches, a speaker must coordinate not only the inspiratory and expiratory muscles, but also the thoracic and abdominal movements. The data in the paper, which are a subset of respiratory kinematic data collected for a study exploring chest wall movement during conversational interactions, were collected using the Inductotracer® system. The test subjects were twenty, American, English-speaking women ranging from ages 22 to 35 (with a mean of 25 years and 4 months). As each participant took turns in the roles of speaker and listener, The Inductotracer® system produced two signals representing rib cage and abdomen excursions, which were digitized at a maximum rate of 2 kHz. This paper focuses on speech breathing signals for each participant in the role of speaker.

A. Examples of Speech Breathing Signals

Fig. 1 highlights the components (respiratory magnitude and respiratory duration) used to properly determine the start times, and identify cycles. Fig. 2 shows the difference in behavior between speech breathing signals and tidal breathing signals. As can be seen, tidal signals appear smoother, thus easier to break down into cycles, whereas speech signals appear to have subcycles, later identified as undesired cycles in this paper. The phenomenon is caused primarily by motion from the test subject during the experiment.

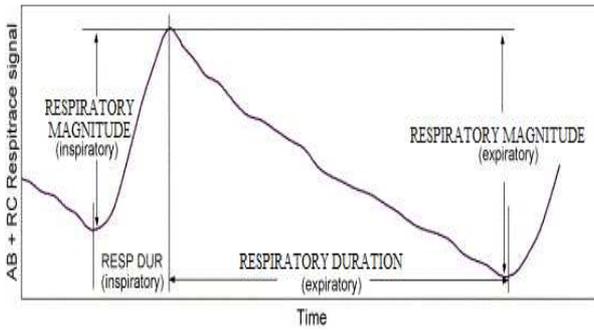


FIGURE I: EXAMPLE OF INDIVIDUAL SPEECH BREATHING, START AND END OF EACH CYCLE MARKED BY VERTICAL LINES.

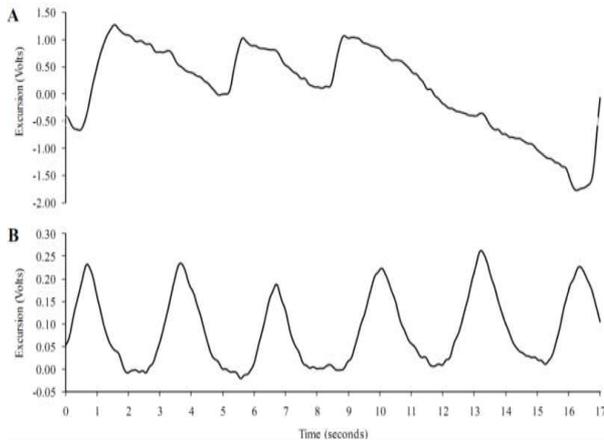


FIGURE II: EXAMPLE OF SPEECH BREATHING SIGNAL (A) AND TIDAL BREATHING SIGNAL (B).

III AUTOMATIC CYCLE IDENTIFICATION ALGORITHM

The automatic cycle identification algorithm includes six steps, as shown in Fig. 3. Two tasks are accomplished: one to identify the cycles, and determine their exact start and end times; and the other to detect and isolate the distorted cycles that cannot be used for the purpose of respiratory signal analysis. To make the proposed algorithm universal to all respiratory speech data collected, detailed criteria have been used for both tasks via multiple steps. To describe the algorithm, the following notations are used: The uppercase letter M represents the total number of data samples in a speech breathing data file. The excursion voltage value at the m -th sample is denoted by $Y(m\tau)$, $0 \leq m \leq M - 1$, where τ is the sampling period in seconds (the sampling period in the analyzed data was 0.01 seconds). The total number of cycles is denoted by N_A , where subscript "A" is used to indicate the step and stage addressed in the proposed cycle identification algorithm; typical values for "A" are I, II, or III. The start time for the n -th breathing cycle is denoted by $T_A[n]$. The cycles are indexed consecutively; therefore, the start time for the n -th cycle is also the ending-time for the $(n - 1)$ -th cycle. Notation $\Phi[n]$ represents the maximum excursion voltage value in the n -th cycle. The minimum values in the inspiratory and expiratory phases in the n -th cycle are denoted by $Y_L[n]$ and $Y_R[n]$, respectively (which are usually not identical), and subscripts L and R are used because the inspiratory phase is in the left portion of a cycle and the expiratory phase is in the

right portion. The terms $Y_L[n]$ and $Y_R[n]$ are also referred to as the left and right minimum in the n -th cycle.

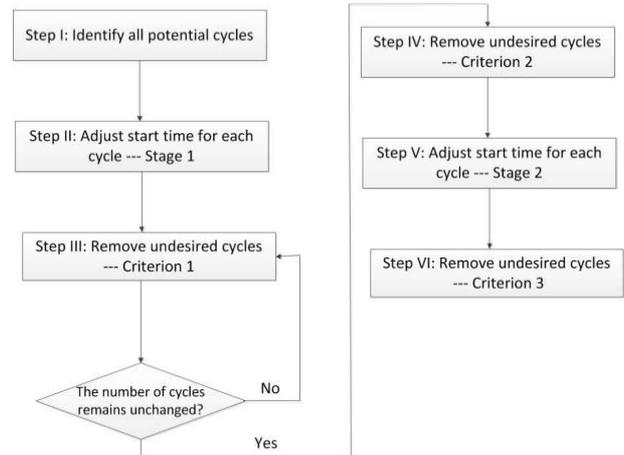


FIGURE III: DIAGRAM OF AUTOMATIC CYCLE IDENTIFICATION ALGORITHM.

A. Step I: Identification for all Possible Cycles

Step I identifies all potential cycles in the speech breathing signals by looking for possible start-time instants via various processing techniques such as decimation, derivation, and interpolation. As done in the previous study of tidal breathing signals [5], derivation is used as an initial means to identify the start time for the potential cycles in speech breathing signals. This is due to the fact that, although almost periodic, the time duration for each cycle is not a fixed constant, and the excursion voltage values at the end of expiratory phases often have small magnitudes yet fluctuate rapidly. The signal is decimated by a factor of 10 (unlike previously at a factor of 30), and a low-pass filter is used to smooth the data. Interpolation is performed on the decimated respiratory signal to recover the original time duration. The processing techniques in Step I also identify a number of cycles that do not represent any normal speech breathing cycles, referred to as "undesired or distorted cycles." Extra processing is required to correctly locate these cycles and remove them from the data file. This task is carried out in Steps III, IV, and VI.

B. Steps II and V: Adjustment of Start Times for Stages 1 and 2

The task for these steps is to find the exact start time of each cycle and adjust the location of the corresponding identification line. The start time of a cycle is often defined as the time-instant at which the next inspiratory phase starts and whose excursion voltage has a locally mathematical minimum or near-mathematical minimum (referred to as a value that is so close to the mathematical minimum that the difference cannot be detected visually). Due to the high complexity of speech signals, a two-stage procedure is implemented in Steps II and V in order to perform a final adjustment on the cycle identification lines such that the exact start time for each cycle can be successfully determined. In Step II of Stage 1, we identify the time-instant at which the inspiratory phase starts and the excursion voltage increases consistently. Step II increases the accuracy of next task, i.e., detection and removal of undesired cycles. However, the start times of cycles

identified in Step II are often over adjusted. These offsets in start times often occur in situations where there are many small dips in the ending segment of the expiratory phases. Consequently, the cycle identification lines are drawn between these dips. In Step V, Stage 2 is introduced to undo the over adjustment by looking for the time-instant at which the previous expiratory phase ends. Detailed descriptions of the two stages in Steps II and V are as follows.

1) Step II Stage 1

The diagram for Step II Stage 1 is shown in Fig. 4. We search for the time-instant at which the inspiratory phase starts and the excursion voltage increases consistently over a certain number of samples. We choose 15 samples in the algorithm. As can be seen, the start times for some cycles identified in Stage 1 are slightly over adjusted from the correct desired start times because there is usually a small bump in the trough before the cycle. In Stage 1, the start time is searched such that all excursion voltages in the next 15 samples increase consistently. As a result, these small bumps are excluded from the current cycle's inspiratory phase; in other words, they are included into the previous cycle's expiratory phase. Further adjustments are performed in Step V of Stage 2.

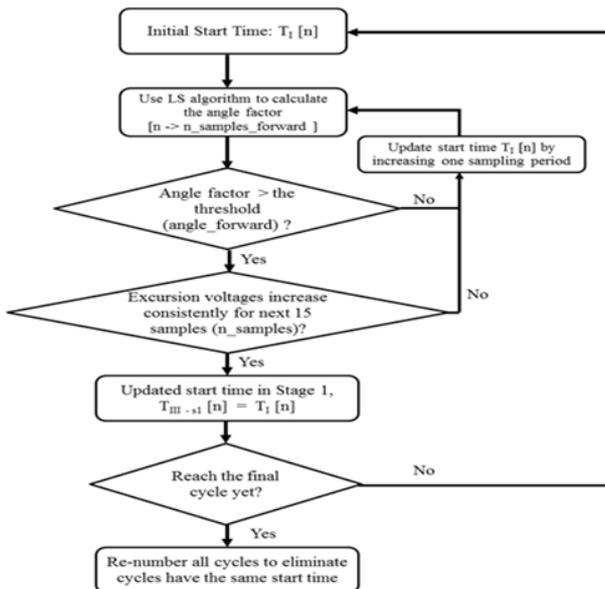


FIGURE IV:STEP II, STAGE 1, ADJUSTMENT OF START TIME FOR EACH CYCLE.

2) Step V Stage 2

Step V is designed to further adjust the exact start time of each cycle. To undo the over adjustment in Step II, we identify the time at which the next inspiratory phase starts and examine the local minimum in the neighborhood of the trough between two adjacent cycles. A least squares method is introduced to accurately detect the valley of each cycle by calculating the angle factor for an n-sample back-tracked line. Fig. 5 illustrates the idea of a back-tracked line such that the small bumps cannot be ignored when determining the correct start time of a cycle. The value of n is controlled by a variable called back_step in the algorithm. The least squares method applied is described as follows:

First, the least squares method is used to calculate the angle factor of the best fit straight line from the present sample to n-samples (n_samples_forward=30) after it.

If this angle factor is smaller than a certain threshold (angle_forward=0.01), then this sample is discarded and the algorithm moves to the next sample.

Until the algorithm reaches a point that has a large enough angle factor, the n-samples after this point (n_samples=15) will continue to be examined if they increase continuously or not.

If not, the algorithm moves to the next sample and calculates the angle factor after this sample again.

If yes, this point is the start time of this cycle. The algorithm will move on to the next cycle.

The least squares method is used to step over long bumps (thus, it uses a large number of samples, n_samples_forward=30), while the previous method is used to step over short bumps (thus, it uses a smaller number of samples, n_samples=15).

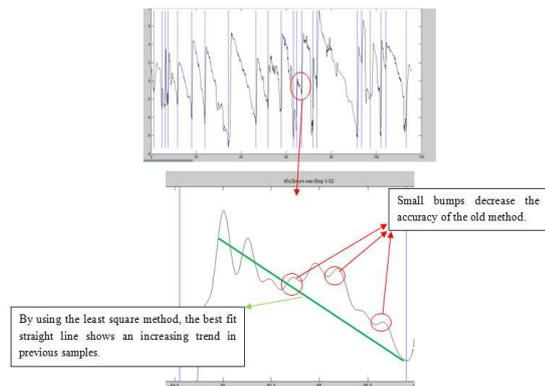


FIGURE V: STEP V, STAGE 2, BACK-TRACKED LINE.

3) Steps III, IV, and VI: removal of undesired cycles—criteria 1, 2, and 3

The undesired cycles are often caused by artifacts in the measurement data. These cycles can be generally classified into two patterns, based on their locations, as shown in Fig. 6. Steps III, IV, and VI are proposed to remove these undesired cycles in each pattern using three criteria, in which various characteristics of the data—local maxima, minima, and slope—are applied.

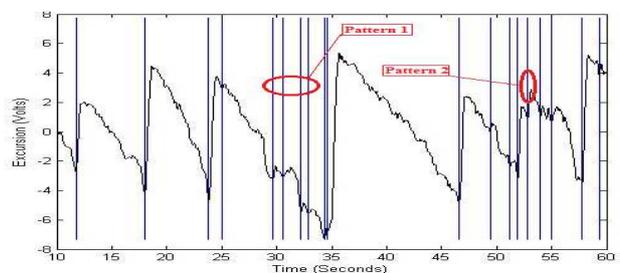


FIGURE VI: TYPICAL UNDESIRED CYCLE PATTERNS.

After the undesired cycles are identified and removed, the cycles are renumbered; therefore, the index of the last cycle is also the total number of desired cycles in the data file.

IV CONCLUSION

We developed an algorithm capable of identifying and isolating the quasiperiodic patterns (cycles) in speakers' speech breathing signals. Despite the higher complexity in the speech breathing signals, the proposed algorithm not only identified the exact start time of the cycles but also effectively detected the undesired cycles.

ACKNOWLEDGMENTS

This work was supported by the National Natural Science Foundation of China (No.51101113), Tianjin Natural Science Foundation (No.14JCYBJC16200).

REFERENCES

- [1] P. de Chazal, C. Heneghan, E. Sheridan, R. Reilly, P. Nolan, and M. O'Malley, "Automated processing of the single-lead electrocardiogram for the detection of obstructive sleep apnea," *IEEE Trans. Biomedical Engineering*, vol. 50, no. 6, pp. 686–696, June 2003.
- [2] R. Bailon, L. Sornmo, and P. Laguna, "A robust method for ECG based estimation of the respiratory frequency during stress testing," *IEEE Trans. Biomedical Engineering*, vol. 53, no. 7, pp. 1273–1285, July 2006.
- [3] O. Meste, B. Khaddoumi, G. Blain, and S. Bermon, "Time-varying analysis methods and models for the respiratory and cardiac system coupling in graded exercise," *IEEE Trans. Biomedical Engineering*, vol. 52, no. 11, pp. 1921–1930, Nov. 2005.
- [4] J. W. Hamner and J. A. Taylor, "Automated quantification of sympathetic beat-by-beat activity, independent of signal quality," *Journal of Applied Physiology*, vol. 91, no. 3, pp. 1199–1206, Sept. 2001.
- [5] Y. Ding, D. F. Parham, Z. Wang, and K. Lee, "Automatic cycle identification in tidal breathing signals," *Journal of Bioengineering & Biomedical Science*, 2012.
- [6] S. M. Schulz, E. Ayala, B. Dahme, and T. Ritz, "A MATLAB toolbox for correcting within-individual effects of respiration rate and tidal volume on respiratory sinus arrhythmia during variable breathing," *Behavior Research Methods*, vol. 41, no. 4, pp. 1121–1126, Nov. 2009.