Wearable Vital Sign Monitoring System for Reinforced Speech Processing

G. Yilmaz, J.-A. Porchet, M. Crettaz, O. Grossenbacher, M. Frosio, O. Chételat, M. M. Doss •

Existing speech technologies largely focus on modeling the oral cavity information from the acquired acoustic signals. However, there are unexploited aspects such as changes in speaker's behavior or psychological state which could effectively be linked to the variations in the physiological signals. Additionally, it is, according to current trends, foreseen that the future of health care and rehabilitation will be shaped by an amalgamation of novel sensing methods, signal processing, and machine learning. Consequently, there is an increasing interest in reinforcing speech processing by means of multi-modal signal acquisition and this work serves to respond this interest.

Sensor solutions developed at CSEM readily provide high quality electro-physiological signal acquisition, such as electrocardiography (ECG) and impedance plethysmography (IPG). Equipping the available systems with acoustic information, such as cardiac or respiratory sounds, is envisaged to open new paths in body signal monitoring.

In this particular work, we collaborate with Idiap Research Institute (Martigny, Switzerland) to reinforce speech processing algorithms that use speech as a biomarker to detect onsets of diseases or behavioral changes such as Parkinson's disease, dementia, and depression [1]. Moreover, application areas of such a system ranges from remote health care to rehabilitation of people with speech impairments resulting from accidents and medical conditions such as stroke or neuro-degenerative diseases. To this end, a system which can synchronously record speech along with ECG, IPG, and cardiac sounds, is developed. Synchronous recording of these signals enable us to investigate links between temporal changes in speech and physiological signals of interest.

The developed prototype (Figure 1) is built upon a modified version of SENSE, CSEM's proprietary system which can measure ECG, breathing rate, skin temperature and classify activities. The system can be worn by using either the belt or the vest designed for SENSE. Measuring sensor (M) of the SENSE module was modified so that it can create a signal for synchronization of new modules and receive external on/off commands. Data acquisition from SENSE modules is performed via Bluetooth.

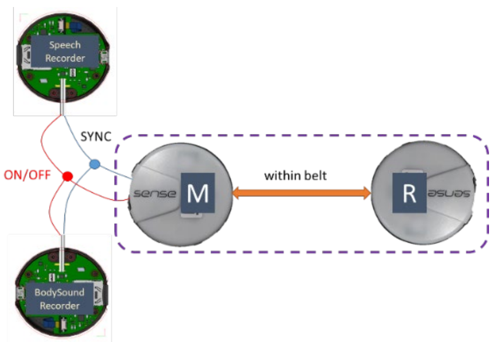

Figure 1: Illustration of the developed prototype composed for multi-modal signal acquisition for reinforced speech processing.

Firstly, a speech recording module is added to the modified SENSE system. This module can synchronously record ambient sound, thus voice of the user as well as others', while SENSE module synchronously records ECG, breathing rate, and skin temperature. It is possible to record for more than 4 hours on a micro SD card inserted into this module. Speech recording unit employs an off-the-shelf bottom-port MEMS microphone interfaced with proprietary electronics developed at CSEM. Positioning of the microphone has been defined and refined with our collaborator to ensure that quality of the recording is suitable for post-processing.

Body sound recording module (stethoscope) incorporates a contact microphone designed and fabricated in-house. The sensing unit is realized by a piezoelectric material which is formed into a specific shape. The microphone has been tuned to capture particularly cardiac signals, in line with stethoscopes available to clinicians, by modifying its shape. Again it is interfaced with a signal-acquisition chain developed at CSEM. Apart from the sensor and frontend electronics, this module has the same capabilities as the speech recording module. Particular attention is paid to the housing design so that this module minimizes interference of external vibrations and common mode electrical signals, e.g., 50 Hz line signal. Both modules are powered by an internal lithium-polymer battery which can be recharged via USB.

In conclusion, the developed system enables synchronous recording of cardiac sound, ambient sound, ECG, and IPG which reveals breathing rate (Figure 2). For cardiac sound acquisition, it is observed that S1 and S2 heart sounds are clearly distinguishable. In the next phase, Idiap will use this system to collect multi-party conversation data and develop tools for non-verbal analysis of conversations integrating physiological signals.

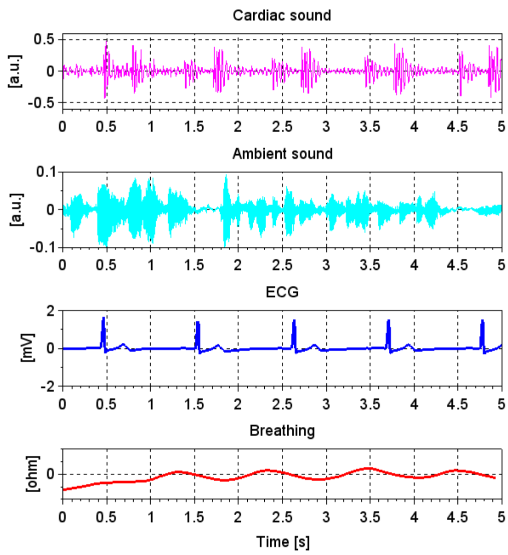

Figure 2: Synchronized waveforms acquired by SENSE and two additional modules designed for speech and body sound recording.

• Idiap Research Institute, Martigny, Switzerland

[1] https://www.tapas-etn-eu.org