INCORPORATING REAL-TIME BIOFEEDBACK CAPABILITIES INTO A VOICE HEALTH MONITOR

Andrés F. Llico[1], Matías Zañartu[1]*, Daryush D. Mehta[2], Jarrad H. Van Stan[2], Harold A. Cheyne II[3], Agustín J. González[1], Marzyeh Ghassemi[4], George R. Wodicka[5], John V. Guttag[4], and Robert E. Hillman[2]

[1] Department of Electronic Engineering, Universidad Técnica Federico Santa María, Valparaíso, Chile
[2] Center for Laryngeal Surgery and Voice Rehabilitation, Massachusetts General Hospital and Harvard Medical School, Boston, MA, USA
[3] Bioacoustics Research Program, Laboratory of Ornithology, Cornell University, Ithaca, NY, USA
[4] Computer Science and Artificial Intelligence Laboratory, Massachusetts Institute of Technology, Cambridge, MA, USA
[5] Weldon School of Biomedical Engineering, Purdue University, West Lafayette, IN, USA

* Corresponding author. Email: matias.zanartu@usm.cl

Abstract:* Many common voice disorders are chronic or recurring conditions that result from abusive and/or faulty patterns of vocal behavior referred to as vocal hyperfunction. Thus, an ongoing goal is the development of long-term ambulatory monitoring for clinical assessment, prevention and modification of hyperfunctional patterns of vocal behavior. This paper reports our initial efforts toward real-time behavioral biofeedback using the smartphone-based Voice Health Monitor that records the high-bandwidth acceleration from the neck skin above the sternal notch. By incorporating real-time estimation of fundamental frequency and sound pressure level, the monitor can provide various types of biofeedback to the user through a vibrotactile alert. The performance of the monitor is compared with that of the commercially available Ambulatory Phonation Monitor (APM) using a bioacoustic transducer tester that generates repeatable vibratory stimuli recorded from ahuman subject. Ambulatory features computed include phonation time, fundamental frequency, sound pressure level, and compliance to a specified threshold level. The results support the implementation of biofeedback in the smartphone-based system and illustrate that the new platform's technology performs more reliably than the APM. Future work calls for the exploration of more sophisticated algorithms to measure vocal behavior and provide clinically meaningful biofeedback.

***Keywords:* Voice use, ambulatory voice monitoring, neck accelerometer, vocal hyperfunction, biofeedback.**

I. INTRODUCTION

It is believed that abusive and/or faulty patterns of vocal behavior lead to functional dysphonia or phonotraumatic lesions, such as nodules and polyps, on the vocal folds. It has been suggested that this type of vocal behavior, often referred to as hyperfunction, could be better characterized and treated by incorporating daily long-term ambulatory voice monitoring and biofeedback into the clinical management process. [1]. Cheyne and colleagues developed such an ambulatory monitoring system that employed a neck surface accelerometer as the phonation sensor that provided a number of advantages over microphone-based systems [2]. This device, the Ambulatory Phonation Monitor (APM, Model 3200, KayPENTAX), has been commercially-available for research and clinical use since 2006 and was used in this study as a reference. The APM does not store the raw accelerometer waveform and thus only operates as a data logger of fundamental frequency (F0) and sound pressure level (SPL) every 50 ms for up to a maximum duration of approximately 14 hours. The APM can also provide biofeedback (via a pager vibrator) based on upper or lower thresholds set for F0 or SPL. Ambulatory biofeedback using the APM has been shown in early case studies to have the potential to facilitate vocal behavioral changes targeted in voice therapy [1].

Our group recently developed an enhanced ambulatory system, referred to as the Voice Health Monitor (VHM), employing the same accelerometer sensor coupled to a smartphone platform [3], as shown in Fig. 1a and 1b. The VHM overcomes many technical limitations of the APM, thus providing the capability to acquire and archive raw acceleration data for over 7 days, with an 80 dB dynamic range, 11.025 Hz sample rate, 16-bit quantization, and processing power to run complex algorithms [3]. Prior to this study, the VHM operated only as a waveform acquisition system, without biofeedback capability. This study aims to expand the VHM operation to incorporate real-time biofeedback features and to compare its performance with that of the APM.

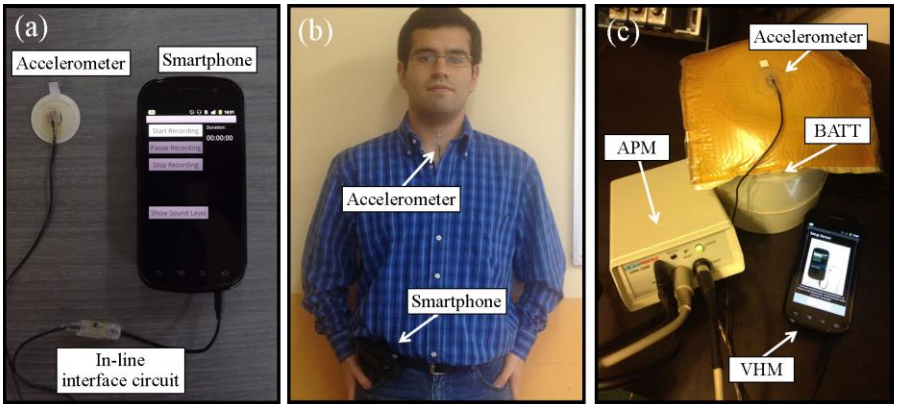

Fig. 1. Mobile Voice Health Monitor: (a) Smartphone, in-line interface circuit, and accelerometer input, (b) Subject wearing the VHM system, and (c) BATT platform for APM and VHM comparison.

II. METHODS

Given the processing capabilities of the smartphone platform, numerous biofeedback targets and approaches can be implemented in the VHM. In this study, we focused on the real-time estimation of F0 and SPL to mimic and contrast the current real-time biofeedback performance of the APM. This comparison was performed using a repeatable excitation signal and biofeedback triggering setup. As shown in Fig. 1c, the same light-weight accelerometer provided the input stimulus to both systems. The accelerometer was mounted on a bioacoustic transducer tester (BATT) [4] that was set to have a flat, band-limited response between 70 Hz and 2 kHz. The BATT was excited with an ambulatory recording previously captured with the VHM from an adult male subject with normal voice (a teacher during a 90-minute lecture), thus providing a signal comparable to that initially obtained with the VHM. Both systems were calibrated with the same subject-specific parameters that related accelerometer level to acoustic SPL.

Each 50 ms frame was divided into two 25 ms subintervals, and the frame was considered voiced if both subintervals exceeded 62 dB SPL. SPL was then re-computed over the entire frame duration. F0 for each voiced frame was equal to the reciprocal of the first peak location in the normalized autocorrelation function if the peak exceeded a threshold of 0.25 [3]. F0 was restricted to the range of 60 to 500 Hz, otherwise, SPL and F0 frame values were set to zero.

A frame counter kept track of the number of frames above-alarm threshold, which is set to 95 dB SPL, increasing when a frame is labeled as voiced, and decreasing when it's not voiced. Biofeedback was triggered when counter reached the equivalent of 300 milliseconds.

III. RESULTS

The summary statistics for the APM and VHM are shown in Table I. With the same conditions and calibration provided to each system, the measured phonation time, percent compliance, and biofeedback time differ slightly. Average differences between F0 and SPL estimates were around 2 Hz and 1 dB, respectively. Although these average measures were similar, differences were observed in the histograms for each parameter (Fig. 2). The slightly greater APM values around the average F0 in Fig. 2a are consistent with the APM labeling more frames as voiced than did the VHM. The SPL histograms in Fig. 2b show that most differences involved the extreme values of the distribution. These findings indicate that the APM labeled more lower-energy and higher-energy frames as voiced and shifted the center of the distribution, thus explaining its increased accumulated phonation time, lower compliance time, and higher biofeedback time. Although the overall differences under the testing conditions are small, the results for the VHM better align with those reported in the literature [1-3,5].

Table I: Summary statistics of ambulatory phonation measures for both systems.

Device	**Total Time (hh:mm:ss)**	**Phonation Time (hh:mm:ss)**	**Average F0 (Hz)**	**Average SPL (dB)**	**% Compliance (SPL <= 95 dB)**	**Biofeedback Time (hh:mm:ss)**
APM	02:08:47	00:38:03 (29.61%)	145.9	82.7	93.5	00:00:46 (2.03 %)
VHM	02:08:52	00:32:45 (25.42%)	148.1	81.6	96.3	00:00:19 (1.01 %)

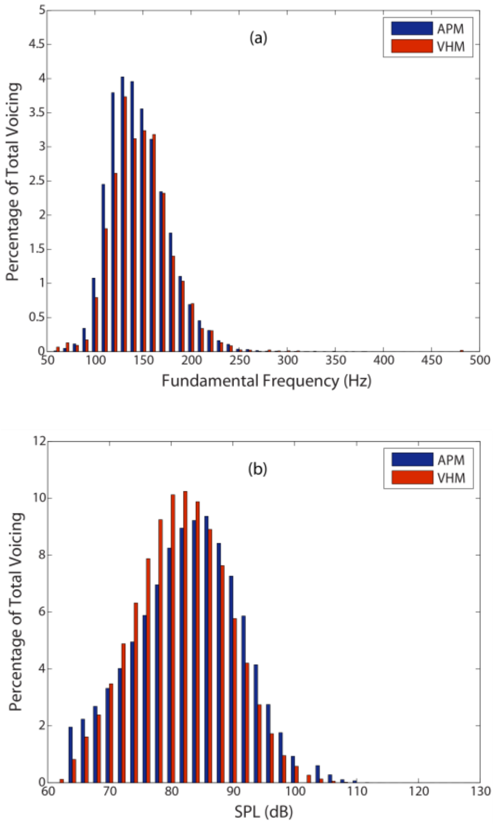

Fig. 2. Histograms for APM (blue) and VHM (red): (a) F0 and (b) SPL

IV. DISCUSSION

Given that the APM is based on technology that is at least 8-10 years old, it is not surprizing that is does not perform as well as the much newer VHM. The VHM provides a 16-bit quantization of the accelerometer signal versus 7-bit quantization employed by the APM. Furthermore, the APM operates with fixed-point arithmetic, where the signal level (in uncalibrated dB units) can only be saved in whole number units that are later converted to whole number SPL values; thus, the level resolution of the APM exhibits round-off error and a potentially coarse representation of SPL. Due to these differences in memory allocation and amplitude quantization, the VHM has approximately 40 dB more in dynamic range than the APM. These factors may explain the differences in resolution for the estimated units of dB SPL and would indicate that the APM is less precise when representing SPL levels for monitoring and biofeedback purposes (in accordance with our clinical observations). The added real-time biofeedback capabilities in the VHM are also more reliable and well suited for professional voice users and patients that have larger vocal ranges.

V. CONCLUSION AND FUTURE WORK

Real-time biofeedback capabilities were added to the VHM based on SPL and F0 thresholds, and its performance was compared with that of the APM. The VHM showed better performance than the APM in terms of its quantization, dynamic range, and computational precision. Subsequent investigations with the VHM in the context of an enhanced real-time biofeedback include the estimation of aerodynamic parameters using impedance-based inverse filtering [6] and z-score assessment [7], as well as wireless connectivity with a server in the clinic. The ability to better facilitate vocal behavioral changes with these new real-time features remains to be tested.

ACKNOWLEDGMENT

This work was supported by NIH-NIDCD grant R33 DC011588, CONICYT grant FONDECYT 11110147, and MIT MISTI grant MIT-Chile 2745333.

REFERENCES

[1] R. E. Hillman and D. D. Mehta, "Ambulatory monitoring of daily voice use," *Perspectives on Voice and Voice Disorders*, vol. 21, no. 2, pp. 56–61, 2011.

[2] H. A. Cheyne, H. M. Hanson, R. P. Genereux, K. N. Stevens, and R. E. Hillman, "Development and testing of a portable vocal accumulator," *J. Speech. Lang. Hear. Res.*, vol. 46, no. 6, pp. 1457-1467, 2003.

[3] D. D. Mehta, M. Zañartu, S. W. Feng, H. A. Cheyne II, and R. E. Hillman, "Mobile voice health monitoring using a wearable accelerometer sensor and a smartphone platform," *IEEE Trans. Biomed. Eng.*, vol. 59, no. 11, pp. 3090–3096, 2012.

[4] S. S. Kraman, G. A. Pressler, H. Pasterkamp, G. R. Wodicka, "Design, construction, and evaluation of a bioacoustic transducer testing (BATT) system for respiratory sounds," *IEEE Trans. Biomed. Eng.*, vol. 53, no. 8, pp. 1711–1715, 2006.

[5] I. R. Titze, E. J. Hunter, and J. G. Švec, "Voicing and silence periods in daily and weekly vocalizations of teachers," *J. Acoust. Soc. Amer.*, vol. 121, no. 1, pp. 469–478, 2007.

[6] M. Zañartu, J. C. Ho, D. D. Mehta, R. E. Hillman, and G. R. Wodicka, "Subglottal impedance-based inverse filtering of speech sounds using neck surface acceleration," *IEEE Trans. Audio Speech Lang. Proc.*, 21(9), pp. 1929-1939, 2013.

[7] R. E. Hillman, E. B. Holmberg, J. S. Perkell, M. Walsh, and C. Vaughan, "Objective assessment of vocal hyperfunction: An experimental framework and initial results," *J. Speech Hear. Res.*, vol. 32, no. 2, pp. 373–392, 1989.