Medical Engineering & Physics 32 (2010) 940–944

Contents lists available at ScienceDirect

Medical Engineering & Physics

journal homepage: www.elsevier.com/locate/medengphy

Communication

A novel integrated mechanomyogram-vocalization access solution

Natasha Alves [a,b], Tiago H. Falk [a,b], Tom Chau [a,b,*]

[a] Bloorview Research Institute, Bloorview Kids Rehab, Toronto, ON, Canada
[b] Institute of Biomaterials and Biomedical Engineering, University of Toronto, Toronto, ON, Canada

ARTICLE INFO

Article history:
Received 12 April 2010
Received in revised form 4 June 2010
Accepted 6 June 2010

Keywords:
Rehabilitation engineering
Signal processing
Access technology
Mechanomyogram
Eyebrow switch
Non-verbal vocalizations
Multiple switches
Disability
Muscle activity
Silent speech interfaces

ABSTRACT

We introduce a novel dual-switch control paradigm based on the simultaneous measurement of frontalis muscle mechanomyography (MMG) and vocalizations (humming) using a single contact microphone attached to the forehead. Vibrations of the face and skull during vocalization are manifested as periodic high-frequency components in the microphone signal recorded at the forehead. The presence of these periodic components is detected by a normalized cross-correlation function, while muscle contractions are detected using a continuous wavelet transform method. The dual-switch provides two independent binary control signals. Eleven participants, including one individual with severe physical disabilities, participated in a cued activation task in which the dual-switch exhibited sensitivities and specificities of 96.8 ± 3% and 98.4 ± 1%, respectively for vocalizations, and 99.7 ± 0.5% and 99.2 ± 0.5%, respectively for muscle contractions. Since skin vibrations due to voiced sounds and muscle contractions have non-overlapping dominant bandwidths, the performance of the MMG switch was not affected by vocalizations. This new integrated MMG-vocalization access solution affords the user two binary switches from a single access site, and may thus augment access alternatives for certain individuals with severe physical disabilities.

1. Introduction

Historically, individuals with limited functional movements or speech as a result of degenerative motor neuron diseases, severe cerebral palsy, brain stem stroke or brain injuries have had limited means of interacting with the surrounding world. By providing a switching interface that the user is capable of controlling, access technologies offer individuals with severe disabilities alternative means for communication, environmental control or computer access, thus promoting their participation in various daily living tasks. An access solution translates a user's functional intent into a functional activity [1]. Depending on the user's abilities, functional intent may be expressed as a physical movement [2], facial gesture [3], physiological changes [4], or combinations thereof. Individuals with some voluntary muscle activity, but without enough motor function to use a mechanical switch, may benefit from access technologies that harness the voluntary contractile ability of their muscles. Muscle activity is conventionally detected by the electromyogram (EMG), and, more recently, the mechanomyogram (MMG).

MMG is a measure of the low-frequency vibrations on the surface of skin during muscle contraction. The vibratory signal is generated from the gross lateral movement of the muscle at the initiation of a contraction, smaller subsequent lateral oscillations at the resonant frequency of the muscle, and dimensional changes of active muscle fibers [5,6]. We have previously described a binary switch controlled by MMG signals recorded from the frontalis muscle of the forehead during small eyebrow movements [7]. In the study, MMG was measured by a silicone-encased microphone with a sealed air-chamber and a soft-silicone contact membrane [8]. The MMG sensor was attached to the forehead with an elastic head strap. The MMG signal recorded during voluntary contractions was recognized, translated into a switch activation signal, and subsequently used to emulate a single keystroke for computer access [7].

Despite its apparent simplicity and limited functionality, binary switches enable the user to accomplish many tasks when used in combination with a user interface, such as an iconic display for a communication aid or an on-screen scanning keyboard. Nonetheless, expressing one's intent via a single switch is time-consuming and cumbersome. Establishing multiple switches requiring different responses allows the user to enhance their level of activity and environmental stimulation [9–11]. Further, having control of multiple switches increases the user's communication channel capacity, thus making the use of assistive technology for communication or computer access more efficient and comfortable. Multiple switches, however, are often associated with inconveniences, such as multiple mountings, time-consuming configuration, multiple connection points, and less aesthetic appearance [12]. Because

---

* Corresponding author at: Bloorview Research Institute, 150 Kilgour Road, Toronto, ON M4G 1R8, Canada. Tel.: +1 416 425 6220x3515; fax: +1 416 425 1634.
E-mail address: tom.chau@utoronto.ca (T. Chau).

1350-4533/$ – see front matter © 2010 IPEM. Published by Elsevier Ltd. All rights reserved.
doi:10.1016/j.medengphy.2010.06.003

of the complexities involved with configuring multiple switches, recent studies have highlighted the benefit of integrated control systems that allow the user to control multiple functions with the same input device [12,13].

The microphone sensor, designed to monitor MMG signals that lie primarily in the 2–50 Hz range, also detects physiological vibrations, such as pulsations due to blood flow, which are considered to be noise and are therefore filtered. Although attenuated by the silicon, skin and soft tissue, the microphone of the MMG sensor detects tissue and bone vibrations caused by resonance of the vocal tract during phonation [14] and singing [15]. These vibrations are sometimes referred to as non-audible murmur (NAM) [16]. Silicone-embedded microphones, similar in design to the MMG sensor, when attached to specific orofacial regions can been used in silent speech interfaces [16]. Phonatory vibrations are quasi-periodic, with typical fundamental frequencies in adults lying between 85 and 225 Hz [17]. While the high-frequency (>100 Hz) components of the microphone signal associated with vocalization may be considered noise with respect to MMG, the detection of voluntary vocalization from the microphone signal recorded at the forehead may add functionality to the binary MMG-switch.

In this study, we propose the integration of two-switches, an MMG switch and a vocalization switch, measured by a single microphone sensor attached to the forehead. The MMG switch is controlled by small eyebrow movements, while the vocalization switch is activated by voiced sounds (such as vowels) or humming. This paper describes the detection of vocalizations, and the experiments used to estimate the performance of the proposed dual-switch.

2. Methods

2.1. Detection of forehead muscle contractions

The algorithm for the detection of muscle activity due to eyebrow movement has been previously described in detail [7,18], and is summarized here for the sake of brevity.

MMG signals were band-pass filtered with a frequency range of 5–100 Hz. Contractions of the occipitofrontalis muscles during eyebrow movements were detected in real-time using a continuous wavelet transform (CWT) algorithm. The algorithm determines the timing of muscle contractions by comparing CWT coefficients of the MMG signals to scale-specific thresholds derived from the baseline signal.

2.2. Detection of vocalizations

For detecting vocalizations, the microphone signal was pre-processed using a 5th order Butterworth filter with a pass-band frequency range of 100–300 Hz. The lower frequency limit reduces the effect of any MMG components of the microphone signal, which predominantly lie below 50 Hz, on the vocalization switch. The upper frequency limit is the frequency above which vibration levels at the forehead are highly attenuated [15].

Since phonatory vibrations are characterized by periodicity [17], the normalized cross-correlation (NCC) function was used to detect periodicity in the microphone signal [19], and thus identify the incidence of vocalization [20]. For the framed signal, $s(n)$, the NCC is given by:

$$\mathrm{NCC}(k) = \frac{1}{\sqrt{e_0 e_k}} \sum_{n=0}^{L_w-1} s(n)s(n+k) \tag{1}$$

where L is the frame size, $L_w < L$ is the size of the correlation window, $k = 0, 1, \ldots L - L_w$ is the lag, and e_k is the energy of the windowed signal, evaluated by:

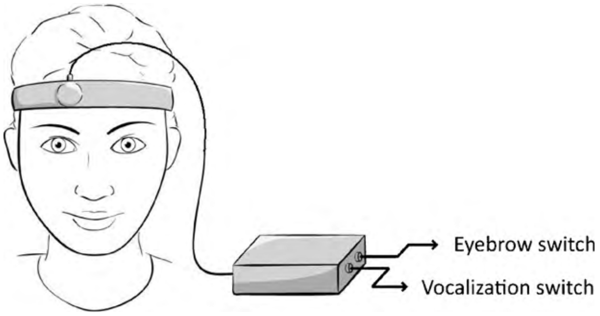

Fig. 1. Schematic representation of sensor placement and switch control.

$$e_k = \sum_{n=k}^{k+L_w-1} s^2(n). \tag{2}$$

The NCC function can assume values between [−1, 1], with values close to unity for lags corresponding to integer multiples of the signal's period [19]. Periodicity, and hence vocalization, is detected when multiple maxima in NCC exist and have values close to unity [20].

In this study, vocalization was detected when the NCC function of the framed signal had at least $N_p \in Z$ peaks with amplitudes greater than a predefined correlation threshold, τ_{NCC}. The parameter settings were similar to those used for other studies on vocal-cord vibration detection [20], with L_w set at 20 ms, $L = 50$ ms, $N_p = 2$, and $\tau_{NCC} = 0.75$. To minimize the false detection of vocalizations due to periodicity in the baseline signal, the vocalization switch was enabled only when the RMS of the filtered high-frequency signal exceeded three standard-deviations of its baseline value.

2.3. Data acquisition

A convenience sample of ten able-bodied individuals (five males, five females), aged 27 ± 2 years, and one adult male with C1–C2 incomplete spinal cord injury (SCI), provided signed consent to participate in the study. Participants were instrumented with an MMG sensor attached to the frontal belly of the occipitofrontalis muscle of the forehead with an elastic headband, as shown in Fig. 1. The position of the sensor on the forehead was not critical, since the frontalis muscle and the frontalis bone are wide, allowing MMG and vocalization to be detected at multiple locations proximal to the eyebrow. For consistency, the sensor was placed 1 cm above the eyebrow, above the inside corner of the right eye.

A program was written in LabView to perform real-time data acquisition, contraction-detection [7], vocalization-detection, switch de-bouncing, and switch activation. Microphone signals were continuously sampled at 1KHz (NI USB-6210, National Instruments), and were processed in non-overlapping 100 ms segments. Each of the two switches was assigned to a separate serial-port, and switch activations asserted the DTR pin of the respective serial-port. The two DTR pins were interfaced to a keyboard encoder (KE-USB36, Hagstrom Electronics) to access two keys of a computer's keyboard independently.

Participants were presented with a pseudo-random sequence of numbers 1–10 at 2 s intervals. They were instructed to activate the MMG switch by raising their eyebrows slightly when the number "1" was presented, and activate the vocalization switch by humming softly when the number "2" was presented. In the first test on the binary MMG switch, 400 stimuli, excluding the number "2", were presented to each participant, with the actionable stimulus ("1") being presented 25% of the time. In the second test on the

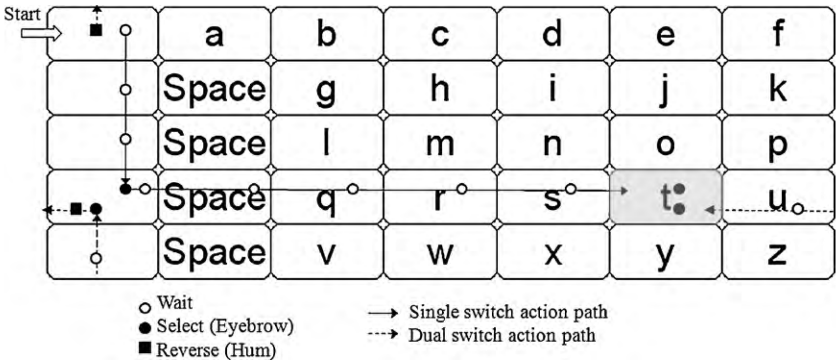

Fig. 2. Shortest paths to select the letter '*t*' using an automatic row-column scanning keyboard with one switch (solid line) and two switches (dashed line).

dual MMG-vocalization switch, 400 stimuli were presented, with each of the two actionable stimuli ("1" and "2") being presented 20% of the time. The sensitivity and specificity of each switch were evaluated, and are given by:

$$\text{Sensitivity} = \frac{\text{True positives}}{\text{True positives} + \text{false negatives}} \times 100 \quad \text{and} \tag{3}$$

$$\text{Specificity} = \frac{\text{True negatives}}{\text{True negatives} + \text{false positives}} \times 100. \tag{4}$$

In order to distinguish between algorithm errors and user errors in the dual-switch paradigm, participants were asked to keep track of their mistakes due to confusion, i.e. humming instead of raising their eyebrow, or vice versa.

In addition to responding to cued stimuli, the switches were tested as the participants typed using an on-screen automatic row-column scanning keyboard (WiVik®). On-screen keyboards can be interfaced with switches to enable individuals with limited mobility to type. When the user has access to a single switch, automatic row-column scanning is a common scanning approach for typing: here the software automatically scans through rows and columns, and the user sequentially selects a row and then a column to type a letter. Additional switches can provide enhancements, such as the ability to cancel an action, adjust scanning rate, and reverse direction, that improve overall effectiveness. In this study, the MMG switch was used to select the row and column of the keyboard. When present, the vocalization switch was used to reverse the row-column scanning direction. The steps required to select a letter using the single MMG switch and the dual MMG-vocalization switch are illustrated in Fig. 2. Participants typed a 31-letter pangram with each of the two switch modalities. The time taken to complete each tasks, and the number of mistakes were recorded. The order of the switch modality and the pangram sentences were randomized to mitigate any familiarization bias.

At the end of the data-collection sessions participants used a five-point linear scale to rate the effort required to activate the dual MMG-vocalization switch. The experimental protocol was approved by the hospital and university research ethics boards, and was in accordance with the Declaration of Helsinki.

3. Results and discussion

Table 1 shows the performance metrics of the switches, averaged across all participants. Results for the participant with SCI were comparable to the able-bodied participants and are not presented separately. Some of the errors in switch activation were due to participants being confused, i.e. humming instead of raising the eyebrow, and vice versa; therefore, switch performance adjusted for user error (last column) are also presented. Both switches in the dual-switch paradigm showed high sensitivity and specificity. The performance of the MMG switch was not affected by introducing the additional vocalization switch, and the sensitivity and specificity of the switch was similar in both the single-switch and dual-switch mode ($p > 0.05$, rank-sum test).

The added functionality reduced the time taken to complete the typing task ($p = 0.008$, rank-sum test). Compared to the single eyebrow switch, the average number of mistakes for the dual-switch typing tasks was higher, but not statistically significant ($p = 0.06$, rank-sum test). Again, most of the mistakes were attributed to participant confusion due to the increased cognitive load. The tests were performed in a randomized order on participants who were unfamiliar with using the switches and the on-screen keyboard. Some of the errors may be mitigated by participant training.

On average, participants rated the effort and fatigue associated with activating the dual-switch as 2 ± 0.7, where [1, no effort at all, not tired; 2, a little effort, not tired; 3, moderate effort, a little tired; 4, a lot of effort, tired; and 5, too much effort, very tired]. In the dual-switch paradigm, the sensitivity of the vocalization switch was lower than the MMG switch ($p = 0.002$, rank-sum test). Participants occasionally had to adjust their tone or pitch in order to activate the vocalization switch. This was expected, since the amplitude of vibrations at the forehead is known to change with phonation during speaking or singing [15]. For example, vibration levels at the forehead are stronger for vowels/i/and/u/, and weak for/a/. Further, high-frequency vocalizations are highly attenuated above 300 Hz and are almost undetectable above 400 Hz [15]. While the signals above 300 Hz were filtered, it is unclear if the switch would be suitable for individuals, such as children, with high-pitch vocalizations. The low-frequency limit of 100 Hz may make the vocalization

Table 1
Performance metrics.

Switch type	Single-switch	Dual-switch paradigm		Dual-switch paradigm corrected for confusion	
	Eyebrow	Eyebrow	Vocalization	Eyebrow	Vocalization
Sensitivity	99.7 ± 0.4	99.7 ± 0.5	96.8 ± 3.8	100 ± 0	97.5 ± 3.5
Specificity	99.9 ± 0.13	99.2 ± 0.5	98.4 ± 1.38	99.8 ± 0.2	98.5 ± 1.3
Typing task					
Time taken (s)	245.6 ± 16.6	217.2 ± 25.4			
No. of mistakes	0.7 ± 0.82	1.6 ± 1.17			

All values are mean ± std. deviation across 11 participants.

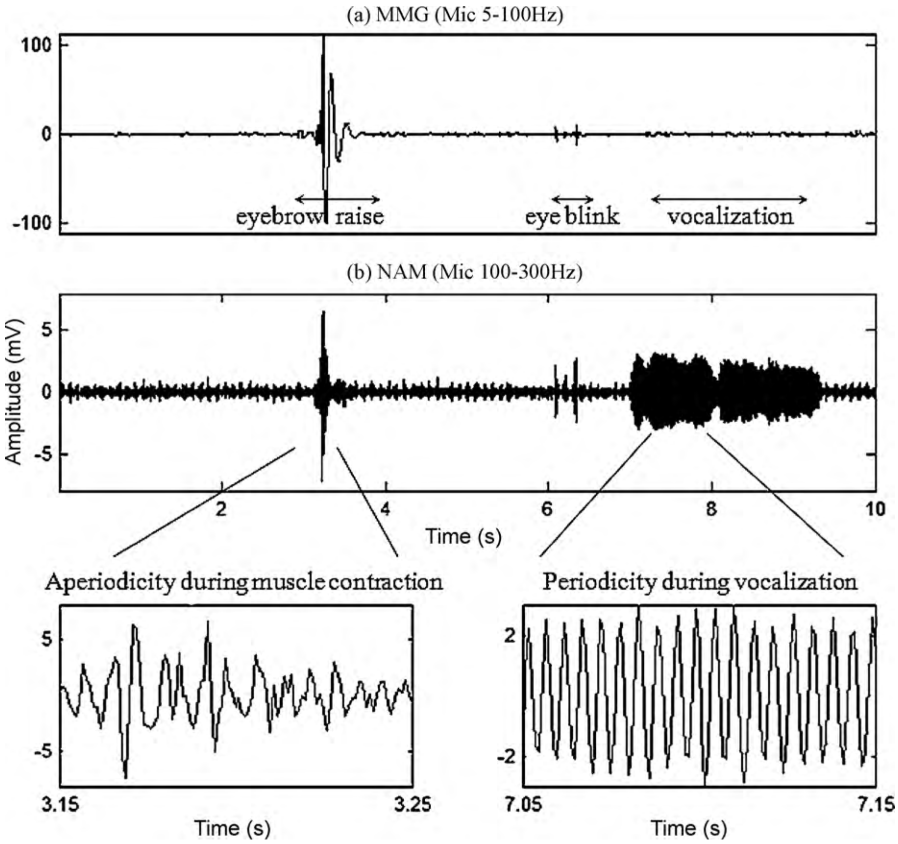

Fig. 3. Microphone-detected MMG and NAM during contraction and vocalization.

switch unsuitable for individuals with very low-pitch vocalizations. However, low sensitivities were not observed for the six adult male participants in the study. Since the dominant frequencies of MMG are well below 50 Hz, the pass-band of the vocalization-detector could be lowered to about 70 Hz to accommodate the lower range of adult male pitches [17].

The differentiation of muscle contractions and vocalization is possible since MMG and NAM occupy non-overlapping bandwidths. However, as seen in Fig. 3, the transient signal at the onset of an eyebrow raise displaces the sensor slightly and introduces an artefact in the microphone signal. While this artefact is usually aperiodic in comparison to the highly periodic NAM during phonation, it occasionally exhibits enough periodicity to falsely trigger the vocalization switch. These false detections could be mitigated by setting a higher correlation threshold (e.g. $\tau_{\text{NCC}} = 0.8$ [20]), or disabling the vocalization switch when the eyebrow switch is triggered.

Since the microphone is attached to the forehead by a headband, unlike non-contact microphones (ex. IST switch by WordsPlus), the transducer has the advantage of being robust to changes in the participant's posture or positioning. Being based on non-acoustically acquired speech cues, the contact microphone is largely insensitive to ambient background noise. In addition, although less sensitive than vocalization-detectors that monitor vocal-fold vibrations at the throat [20], the proposed vocalization switch may offer better comfort, and may be used even when access to the user's neck region is not an option, due, for example, to a tracheotomy.

4. Conclusion

The complementary information provided by a contact microphone during vocalization and muscle contraction facilitates the practical use of a multifunction access solution. While both phenomena are manifested as superficial vibrations, the separation in physiological origins and frequency simplifies their independent control and detection. The integrated detection of muscle contraction and vocalization offers the user two independent bits of control using just one sensor and access site, and may provide expanded access alternatives for certain users with severe and multiple disabilities.

Acknowledgements

This work was supported in part by Natural Sciences and Engineering Research Council of Canada and the Canada Research Chairs program. The authors thank Mr. Ka Lun Tam for his implementing the hardware interfaces, and Mr. Steven Bernstein for assistance with the illustrations.

Conflict of interest

The authors declare that they have no competing interests.

References

[1] Tai K, Blain S, Chau T. A review of emerging access technologies for individuals with severe motor impairments. Assistive Technology 2008;20(Winter):204–19.
[2] Lancioni GE, O'Reilly MF, Singh NN, Sigafoos J, Didden R, Oliva D, Montironi G, La Martire ML. Small hand-closure movements used as a response through microswitch technology by persons with multiple disabilities and minimal motor behavior. Perceptual and Motor Skills Jun. 2007;104:1027–34.
[3] Leung B, Chau T. A multiple camera tongue switch for a child with severe spastic quadriplegic cerebral palsy. Disability & Rehabilitation: Assistive Technology 2010;5. p. 58.
[4] Blain S, Mihailidis A, Chau T. Assessing the potential of electrodermal activity as an alternative access pathway. Medical Engineering & Physics May. 2008;30:498–505.

[5] Barry DT, Cole NM. Muscle sounds are emitted at the resonant frequencies of skeletal muscle. IEEE Transactions on Biomedical Engineering May. 1990;37:525–31.
[6] Orizio C. Muscle sound: bases for the introduction of a mechanomyographic signal in muscle studies. Critical Reviews in Biomedical Engineering 1993;21:201–43.
[7] Alves N, Chau T. The design and testing of a novel mechanomyogram-driven switch controlled by small eyebrow movements. Journal of NeuroEngineering and Rehabilitation 2010;7.
[8] Silva J, Chau T. Coupled microphone-accelerometer sensor pair for dynamic noise reduction in MMG signal recording. Electronics Letters 2003;39:1496–8.
[9] Lancioni GE. Evaluating the use of multiple microswitches and responses for children with multiple disabilities. Journal of Intellectual Disability Research 2002;46. p. 346.
[10] Lancioni GE. Using multiple microswitches to promote different responses in children with multiple disabilities. Research in Developmental Disabilities 2001;22. p. 309.
[11] Realon RE. The effects of making choices on engagement levels with persons who are profoundly multiply handicapped. Education and Training in Mental Retardation 1990;25:299–305.
[12] Nisbet P. Integrating assistive technologies: current practices and future possibilities. Medical Engineering Physics 1996;18. p. 193.
[13] Ding D, Cooper RA, Kaminski BA, Kanaly JR, Allegretti A, Chaves E, Hubbard S. Integrated control and related technology of assistive devices. Assistive Technology: The Official Journal of RESNA 2003;15:89–97.
[14] Kirikae J, Sato T, Oshima H, Nomoto K. Vibration of the body during phonation of the vowels. Revue de Laryngologie, Otologie, Rhinologie 1964;85, p 3.
[15] Sundberg J. Phonatory vibrations in singers: a critical review. Music Perception 1992;9. p. 361.
[16] Nakajima Y, Kashioka H, Shikano K, Campbell N. Non-audible murmur recognition input interface using stethoscopic microphone attached to the skin. In: Proceedings IEEE international conference on acoustics, speech, and signal processing, vol. 5; 5, p. V-708-11. 2003.
[17] Titze IR, Martin DW. Principles of voice production. Acoustical Society of America Journal 1998;104. p. 1148.
[18] Alves N, Chau T. Automatic detection of muscle activity from mechanomyogram signals. Physiological Measurement 2010;31:461–76.
[19] Talkin D. A robust algorithm for pitch tracking (RAPT). Speech Coding and Synthesis 1995;495. p. 518.
[20] Falk T, Chan J, Duez P, Chau T. Assisted communication based on realtime vocal-fold vibration detection. IEEE Transactions on Neural Systems and Rehabilitation Engineering 2010;18(2):159–63.