Non-Invasive Ambulatory Monitoring of Complex sEMG Patterns and its Potential Application in the Detection of Vocal Dysfunctions

N. R. Smith[1,2], T. Klongtruagrok[1,2], G. N. DeSouza[1,2,3], C. R. Shyu[2,3], M. Dietrich[4], *and* M. P. Page[5]

***Abstract*—Voice disorders are non-trivial when it comes to their early detection. Symptoms range from slight hoarseness to complete loss of voice, and may seriously impact personal and professional life. To date, we are still largely missing reliable data to help us better understand and screen voice pathologies. In this paper, we present an ambulatory voice monitoring system using surface electromyography (sEMG) and a robust algorithm for pattern recognition of vocal gestures. The system, which can process up to four sEMG channels simultaneously, also can store large amounts of data (up to 13 hours of continuous use) and in the future will be used to analyze *on-the-fly* various patterns of sEMG activation in the search for maladaptive laryngeal activity that may lead to voice disorders. In the preliminary results presented here, our pattern recognition algorithm (Hierarchical GUSSS) detected six different sEMG patterns of activation, and it achieved 90% accuracy.**

I. Introduction

Occupational voice users such as teachers, singers, etc. are at highest risk for voice disorders largely due to the extraordinary vocal load placed on the laryngeal system while exercising their occupation [1]. Classic symptoms are hoarseness, vocal effort, and vocal fatigue, which are related to vocal hyperfunction [1]. Vocal hyperfunction may lead, in some individuals, to phonotrauma, such as vocal nodules, or to muscle tension dysphonia – i.e. excessive or dysregulated laryngeal muscular activity underlying the vocal changes [2]. To date, we are still largely missing reliable data to help us better understand what dysregulated muscular activity in the laryngeal muscles means. This foundation is necessary to study what differentiates pathological from normal muscular activity during voice for speech. Ambulatory monitoring of voice using surface electromyography (sEMG) of the extralaryngeal muscles can be an innovative approach to advance our understanding of vocal hyperfunction and ultimately to monitor vocal hyperfunction in heavy voice users.

In fact, in the past decade, research intensified the development of devices suitable for ambulatory monitoring of daily speech. Commercial systems were made available to monitor vocal intensity using accelerometers and gyroscopes; voice fundamental frequency (F0) and vocal duration using microphones and frequency transforms, [3], [4]. These non-invasive monitors also have biofeedback capabilities and have recently been designed for smartphone platforms [5]. However, there is a lot more happening during a speech than what can be captured by microphones and inertial sensors.

[1]*Vision-Guided and Intelligent Robotics Lab,*[2]*Electrical and Computer Engineering Department,*[3]*MU Informatics Institute,*[4]*Communication Science and Disorders,*[5]*Otolaryngology – Head and Neck Surgery, University of Missouri, Columbia, MO, USA*

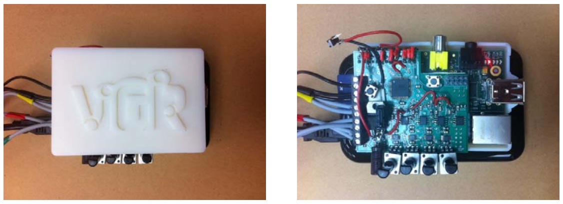
The coordinated contraction of muscle during speech controls a host of biologic functions. As muscles contract, they undergo changes in electrical potentials, which can be monitored by electromyographic (EMG) devices. When studying a single muscle, the optimal signal to noise ratio is typically obtained when the electrodes are placed inside the muscle – a technique available in the healthcare or laboratory setting, but with limited use in people's everyday lives. A less invasive strategy is to place Surface EMG (sEMG) on the skin near the muscle(s) of interest. This method has more real-world practicality, but it comes at the expense of noisier signals and exacerbated occurrences of *cross-talk* between adjacent electrodes. That is because the biologic functions that are subserved by muscular activity do not result from the action of a single muscle, but from the activity of several muscles working in a coordinated fashion.

Recently, the interest in many areas of human-computer interfacing [6], prosthesis [7] and even voice pathology [8], [9] has shifted towards the use of devices that can monitor muscle activity[1]. However, when it comes to recognizing patterns of muscle activity in a reliable, accurate and robust manner, much remains to be done.

In this paper, we present a system for monitoring and recognizing patterns in multiple sEMG signals. A simpler version of the algorithm proposed here has already been used to recognize hand gestures and to operate power wheelchairs [10], [11]. Here, we employed an improved version of that algorithm using four sEMG sensors. We also developed a small device to collect data as subjects go about their daily lives without subjecting them to invasive laboratory testing. The device can wirelessly connect to any smart phone or computer, but it can also provide immediate feedback to the user – either to effect a behavioral change, to monitor progress after an intervention, or to provide early detection of health problems.

The system was connected to the anterior neck of the subject since many complex physiological motor functions occur within the neck (e.g. voicing, speaking, and swallowing). Besides, the muscles in this area are located relatively close to the skin and are quite appropriate for surface EMG. Our goal was to demonstrate that the neck offers an excellent location from which to build our understanding of complex laryngeal patterns underlying voice for speech and non-speech behaviors through sEMG signals. Indeed the preliminary results presented here show that despite the very

[1]http://glneurotech.com/bioradio/bioradio-wireless-physiological-monitor



a) b)

Figure 1. ECHO device depicted from two top angles: a) with and b) without the case cover.

complex muscle groups on the neck, meaningful recognition based on sEMG signals is still possible. For that, data was collected from a single subject for six different vocal patterns, or *gestures,* and using four sEMG channels. The gestures were then classified by our improved hierarchical classifier and an average accuracy of over 90% was achieved.

II. Background and Related Work

A. Voice/Speech Disorders

Classic symptoms of behavioral voice disorders are related to vocal hyperfunction. In some individuals, vocal hyperfunction may lead to phonotrauma, such as vocal nodules, or to muscle tension dysphonia. For example, excessive extralaryngeal muscle activity and a chronic high laryngeal position during speech are characteristic of muscle tension dysphonia [2]. Surface EMG is a noninvasive tool that has been used in voice, speech, and swallowing research, however the body of literature continues to be limited [9], [8]. The new frontier will be to adapt this methodology for ambulatory monitoring of extralaryngeal muscle activity and patterns during voice and speech production. As pointed out earlier, surface EMG of the anterior neck is well suited to capture general information on the muscular activity of the larynx, which can be related to magnitudes and patterns with relevance to upward and downward movements of the larynx in the neck. Presumably, excessive muscle activity or lack of variability in laryngeal movements may be related to vocal fatigue, which in particular plagues occupational voice users such as teachers [1]. The research question to be addressed is whether ambulatory sEMG devices can reliably associate patterns of extralaryngeal muscle activity with voice tasks underlying speech and non-speech behaviors (e.g. voiced sounds, throat clear). If so, the methods will be applied clinically to differentiate between normal and maladaptive laryngeal patterns associated with voice problems.

B. Pattern Recognition of sEMG Signals

The work in [10] introduced the idea of Guided Underdetermined Source Signal Separation (GUSSS) and the GUSSS ratio. In [10], the focus was on discriminating different Muscle Unit Activation Potential Trains, or MUAPT patterns, that emerge when different gestures are performed. As many systems do, it was assumed that an sEMG sensor captures a combination of statistically independent MUAPTs due to cross talk [12], [13]. But unlike most methods in the literature, the system in [10] relied on a single sensor. This was possible because the main characteristic of the GUSSS ratio is that it can indicate the presence or absence of a particular signature or MUAPT pattern within a sensed sEMG signal. The term “Guided” in GUSSS refers to the fact that the sought-out signature is “injected” into the observed signal in order to obtain a corresponding ratio. A low ratio indicates that the signature is most likely present within the sensed signal. A high ratio, on the other hand, indicates that the signature is not being detected in the signal.

Later, a framework for controlling a power wheelchair using the GUSSS method was developed and tested in [11]. The framework proposed a control system based on the recognition of hand gestures. The use of hand gestures was simply to illustrate the fact that any muscle activation pattern or signature derived from a natural and repetitive muscle movement can be employed by the system. In the case of a person with severe impairment, any other muscle movement could be used instead (e.g. eyebrow movement). Compared to other systems found in the literature, which use multiple sEMG sources for classification, the method in [11] compared quite reasonably, reaching up to 92% accuracy for three gestures.

More recently in [14], a hierarchical system based on the GUSSS was developed to achieve higher classification accuracy for a greater number of gestures. The hierarchical method employed tuples of gestures rather than comparing each gesture against every other gesture. This approach allowed the system to compare first the gestures that were easily distinguishable from other gestures, leading to a better accuracy of the system even as the number of gestures increases, reaching up to 86% accuracy for nine gestures.

III. Device Description

The proposed EMG multi-Channel Hardware for Otolaryngology (ECHO) is an Otolaryngology REcording, Analysis and Diagnostic device (OREAD) to log sEMG data from multiple differential sEMG sensor channels. One key feature of the ECHO-OREAD device is that it maintains a small form factor (8.5cm\, x\,6cm\, x\,4.5cm) in order to be portable so that it can be used in a variety of applications. The device is connected to a rechargeable lithium-ion battery to maintain portability. Figure 1 shows two pictures of the ECHO-OREAD device sitting on top of the rechargeable battery and Figure 2a shows two sets of signals from all four channels captured with ECHO-OREAD for the *Cough* and */t/* gestures, with electrode placement as described in Section III-B.

A. Hardware

The hardware of the device consists of a Raspberry Pi board with a custom built PCB docked on top. The custom built PCB contains circuitry for analog to digital conversion and four channels of sEMG inputs. The circuit provides amplification and individual, manual control of the gains for each channel. The channels are also filtered in order to reject undesirable frequencies. Once the signals are amplified and filtered, they are digitized and transferred to the Raspberry Pi through its GPIO connector. Additional buttons on the top of the device can be used to control the behavior of the boards,

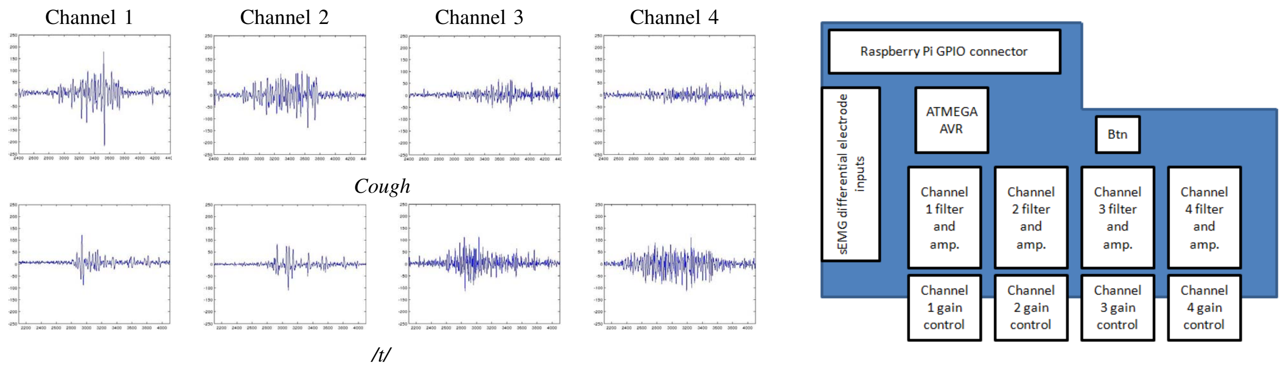


Figure 2. a) Four channels of sEMG signals for gestures *Cough* and */t/*; and b) Basic diagram of the custom PCB for the ECHO-OREAD device.

such as resetting the acquisition and reinitializing the boards. Figure 2b shows a basic diagram of the custom PCB built for the ECHO-OREAD device.

B. Electrode Placement

As seen in Figure 3, surface electrodes were placed according to established guidelines for sEMG recordings [15] with special consideration of recommendations proposed for voice, speech, and swallowing research [9]. Disposable 10mm Ag/AgCl surface electrodes (Bio-Medical Instruments, Warren, MI) were placed in bipolar configurations for single differential recordings from the anterior neck musculature. Two identical electrode pairs were placed on the left and right side of the neck to capture suprahyoid (submental) and infrahyoid muscular activity corresponding to elevations and depressions of the larynx during voice for speech, respectively [16]. The first electrode for the submental muscle site was placed approximately 1cm from midline in the submandibular area superior to the hyoid bone [17], [18], [19], [20] The second electrode of the submental pair was placed in line with the fibers of the muscle and with an interelectrode distance of approximately 1cm [15], [8], [9]. The submental location captures muscle activity from the anterior belly of the digastric, mylohyoid, and geniohyoid muscles.

For the infrahyoid muscle site, the first electrode was centered over the thyroid cartilage just below the thyroid notch and approximately 1cm off midline [17], [18], [8], [9]. The infrahyoid location captures muscle activity from the sternohyoid and omohyoid muscles with additional activity captured from the thin muscle sheath called platysma overlying most of the neck [19], [9]. Due to the small sizes of the individual muscles making up the submental and infrahyoid musculature as well as the multilayered structure of the muscles, sEMG can only capture muscle group activity and not activity from individual muscles. Moreover, it is not realistic to record activity from deeper muscles such as the thyrohyoid and cricothyroid [9]. The ground electrode was placed on the superior bony prominence of the left shoulder. For voice and speech recordings, a placement of the ground electrode close to the electrodes is preferred [9].

The quality of electrode placement was confirmed with tasks that produce target activations such as a swallow (submental and infrahyoid activity) and production of a front vowel (/i/, submental) and back vowel (/u/, infrahyoid).

IV. Proposed Method

This research expands on the classification approach presented in [14] to prove the validity of performing sEMG classification based on extralaryngeal muscle activity in the anterior neck, which underlies voice production for speech and non-speech behaviors (voiced and unvoiced sounds, throat clear, swallowing). A major difference between this approach and the one in [14] is that here four sEMG channels were used instead of just one. The proposed framework is illustrated in Figure 4 and it consists of a two-level hierarchical classifier: 1) a GUSSS-based classifier; and 2) a Multi-Class Support Vector Machine (SVM).

As it can be seen in Figure 4, the first level in the hierarchy involves a number of GUSSS-based classifiers. Basically, these classifiers function as confidence generators, inputting feature vectors extracted from the raw sEMG signal and outputting N confidence vectors $\overrightarrow{\lambda}$, where the elements of the vector indicate the confidence that a sEMG signal contains one of the signatures in the tuples – a tuple is a group with an arbitrary number of signatures: e.g. doubles, triples, etc. All of the obtained confidence vectors are concatenated into a second feature vector, which is then input to the classifier at the second level of the hierarchy. The output of the second level classifier is the final class assigned to the observed sEMG signal. The following sub-sections describe in further detail the classifiers at each level, as well as their training process.

A. Class Signatures and Optimal Choice of Tuples

Let us assume that there is a labeled training set with $C \times T$ signals – i.e. T signals from each of the C possible classes

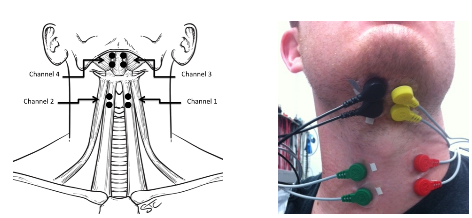


Figure 3. Muscle groups on the human neck: diagram and actual view of electrode placement.

(muscle patterns or gestures). First, a signature for each class is obtained. The current approach is to do an averaging of the T training signals grouped per class. That is, for each class c, a single signature: $s_c = \frac{1}{T}\left(\sum_{x_l\,in\,class\,c} x_l\right)$ is created, where x_l is the l^{th} training signal of class c.

Each GUSSS-based classifier is associated to a tuple of classes, where the sizes and members can be chosen arbitrarily depending on the gestures, user, muscle activity patterns, etc. The rationale behind the tuples is the following: when a large number of C classes are considered at the same time, there might be much confusion between some of the classes. However, it is possible to find subsets of classes for which the confusion between such classes is minimized. So, the goal of the tuples is to allow similar classes to be separated. However, it is also desirable to group as many classes as possible per tuple in order to reduce the complexity of the algorithm. For this paper, the selection of the optimal number and the membership in the tuples was done empirically after trial-and-error. In the future, an automated method for choosing tuples will be explored.

B. sEMG Segmentation and Level 1 Feature Vectors

As mentioned before, the input to each of the GUSSS-based classifiers is a feature vector extracted from the incoming sEMG signal. The features used and the way to obtain the feature vector for a particular tuple, denoted τ_i, is described next. A same procedure applies to all N tuples being considered. Figure 5 depicts a typical sEMG signal and the features considered.

1) GUSSS Ratio: As explained in Section II, the main idea of the GUSSS method is to identify particular signatures within a measured sEMG signal. For any given sEMG signal x, the GUSSS method seeks to identify the presence or not of each possible signatures. This is done by iteratively injecting signatures and obtaining ratios for each one of them. For all $n_i = |\tau_i|$ classes in tuple τ_i, the algorithm obtains the ratios $r_1, \ldots, r_{n_i}$. If signal x contains a pattern in class c, ratio r_c is expected to be smaller than all other ratios r_j, for $j \neq c$.

2) Segmentation of the sEMG Signals: Typically, the sEMG signals for the gestures considered here last from around 250 ms to 500 ms. To capture the structural information of the sEMG signals, we divide them into D segments of equal length. The features described next are calculated for each segment of any given signal.

3) Mean Absolute Value: One feature commonly used for sEMG signals is the Mean Absolute Value (MAV). The MAV of a signal $x(t)$ is obtained by calculating the average of the absolute values of x at all instants t. For a discrete signal:

$$MAV = \frac{1}{K}\sum_{k=1}^{K} |x(k)|$$

where K is the number of samples in a segment of x.

4) Zero Crossing: Another feature extracted from the sEMG signals is the number of Zero Crossings (ZC), which represents how many transitions from positive to negative (or vice-versa) there are in a segment of the signal.

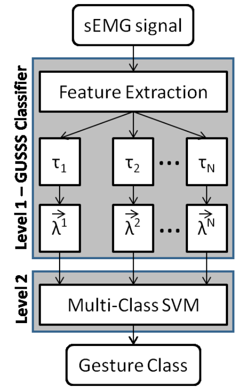


Figure 4. Proposed framework for the Hierarchical GUSSS Classifier used in the ECHO-OREAD.

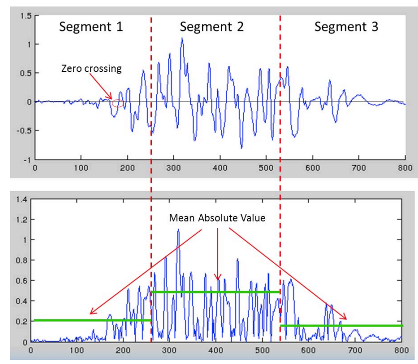


Figure 5. A typical sEMG signal segmented into 3 parts. The zero-crossings are indicated in the top figure. The rectified signal and the MAVs of the segments are shown in the bottom figure.

5) Complete Feature Vector Level 1: After all of the features described above have been extracted, signal x is represented by the following feature vector:

$$\vec{v}_i = [r_1, \cdots, r_{n_i}, m_1, \cdots, m_D, z_1, \cdots, z_D] \quad (1)$$

where $r_1, \ldots, r_{n_i}$ are the GUSSS ratios for each class in tuple τ_i. The MAVs and ZCs for each segment of the signal are m_k and z_k, respectively, for $k = 1, \ldots, D$.

6) Statistics in each Tuple of Gestures: As it will be shown shortly, the system uses the mean vector and covariance matrix of each class within the tuples. So, the above feature vectors are extracted for all T training signals in each class and used to form $\aleph\left(\vec{\mu}^i_j, \sum^i_j\right)$, representing the distribution of class j in the tuple τ_i, where $j = 1, \ldots, n_i$, and $i = 1, \ldots, N$.

C. Distances and Confidence Values

As it was mentioned before, the output of the first level in the hierarchy is a set of confidences that are concatenated to form a second feature vector for the next level. These confidences, which are based on Mahalanobis distances, are obtained by each one of the GUSSS-based classifiers.

First, an input signal y is fed into each one of the tuples described above. Then, for each tuple τ_i, a feature vector $\vec{v}_i$ (eq. 1) is calculated. Finally, the GUSSS-based classifiers calculate Mahalanobis distances to the mean vectors $\vec{\mu}_j^i$ of the classes in tuple τ_i, that is:

$$d_j^i = \sqrt{\left(\vec{v}_i - \vec{\mu}_j^i\right) \left(\sum_j^i\right)^{-1} \left(\vec{v}_i - \vec{\mu}_j^i\right)^T} \quad , j = 1, \ldots, n_i$$

If, for example, distance d_j^i is small (close to zero), the confidence that signal y belongs to class j would be high.

To obtain the actual confidence values, the complementary error function is used:

$$\lambda(d_j^i) = erfc\left(\frac{d_j^i}{\sqrt{2}}\right)$$

where $erfc(x) = 1 - erf(x)$.

For the GUSSS-based classifier corresponding to tuple τ_i, the confidence that signal y belongs to class j is given by $\lambda_j^i = \lambda(d_j^i)$. In the end, the classifier produces n_i confidence levels: $\vec{\lambda}^i = \left(\lambda_1^i, \ldots, \lambda_{n_i}^i\right)$.

Level 2 Feature Vector: After confidence values are obtained for all N tuples, the second feature vector is created as follows: $\vec{u} = \left[\vec{\lambda}^1, \vec{\lambda}^2, , \cdots, \vec{\lambda}^N\right]$

D. Multi-Channel Hierarchical GUSSS

For the enhanced version of the Hierarchical GUSSS with multiple channels used in this research, the steps detailed above are replicated for each channel, leading to a set of vectors $\overrightarrow{u}$. These channel vectors are then averaged in order to form a single confidence feature vector to serve as the input to the multi-class SVM.

E. Level 2 Classifier: Multi-Class SVM

The final classification method consists of a multi-class SVM. To train the SVM, the $\overrightarrow{u}$ vectors are computed for all training signals, for all classes. When it comes to classification, an incoming signal $\overrightarrow{y}$ is fed through level 1 in the hierarchy to obtain the confidences and to create the $\overrightarrow{u_y}$ feature vector. The latter is fed to the multi-class SVM in order to generate the final class assignment.

V. Experiments and Results

In this section, the application of the proposed framework for classification of voice related sEMG patterns is shown. For these experiments, up to six vocal gestures and one resting condition were used. The gestures tested included the following: Cough, Throat Clear, /t/, /s/, /i/, /u/. During the rest period, the subject was asked to be as relaxed as possible, and try to minimize any motion in the throat or mouth area. The sEMG signals of interest, i.e., the ones to be associated with each gesture, are those generated during the transition from the resting condition to the actual vocal gesture and back to resting. In order to further validate the proposed hierarchical approach, two sets of experiments were carried out. The first experiment used a distance classifier and the second, the proposed hierarchical approach.

A. Data Collection

The main goal of these experiments was to validate the fact that meaningful classification can be achieved from extralaryngeal sEMG signals of the anterior neck. Therefore, sEMG signals were collected under well-controlled conditions. One test subject was asked to perform 50 repetitions of each of the six selected gestures. The subject performed all of the gestures of a given type at an interval of about 1.5 seconds per gesture and then rested for a few minutes before attempting the next gesture. Four pairs of sEMG electrodes and a ground electrode were placed as explained in Section III-B and seen in Figure 3.

For the experiments presented here, the data was collected using both the National Instruments digitizer as well as the ECHO-OREAD. Classification was only completed on the data from the National Instruments digitizer as a proof-of-concept. The signals were divided into 3 segments (i.e. $D = 3$), as described in Section IV-B2.

B. Results Using the Distance Classifier

For this experiment, a 10-fold cross validation was performed. Each time 90% of the signals from all collected gestures were used for training. The remaining 10% of the signals were then classified using a simple distance classifier. The distance classifier was used for comparison purposes against our hierarchical method, and the resulting Confusion Matrix can be seen in Table I. This method achieved an overall classification accuracy of about 73%. It should be pointed out here the confusion between both the /i/ and /u/ vocal gestures and the Cough, as indicated by the first entries on the fifth and sixth rows of the confusion matrix (Table I).

C. Results Using Hierarchical GUSSS

In order to validate the classification of the signals, the improved Hierarchical GUSSS classifier was tested with the same data as before, for the distance classifier. Again, for each experiment, a 10-fold cross validation was performed. Each time 90% of the data were used for training and the remaining 10% were then classified using the proposed Hierarchical GUSSS. The results in Table II, also in the form of a Confusion Matrix, demonstrate the higher classification performance of the hierarchical method in comparison with the distance classifier. It is specially noticeable the ability of the Hierarchical GUSSS to distinguish the previously confused /i/, /u/, and Cough. This is attributed to the tuples' ability to isolate similar vocal gestures, which are later classified at level 2 of the hierarchy based on the confidences from the individual tuples.

VI. Discussion, Future Work and Conclusions

The overall accuracy achieved using the four sEMG channels and a simple distance classifier was 73.33%. This result alone validates the claim that meaningful classification can be achieved by applying our GUSSS method to sEMG signals collected from the anterior neck. However, the improved performance of over 90% achieved by the Hierarchical GUSSS further demonstrates the potential of the proposed system, including the ECHO-OREAD device as a tool for detection and diagnosis of voice disorders. The improvement is mostly noticeable due to the better discrimination of similar gestures – i.e. /i/, /u/ and Cough. The reason for this improvement is that the hierarchical method employs tuples of gestures instead of comparing each gesture against every other gesture. This improved discriminant power of the

Table I
CONFUSION MATRIX FOR A SIMPLE DISTANCE CLASSIFIER. THE VALUES ARE AVERAGE PERCENTAGES OVER A 10-FOLD CROSS VALIDATION.

		Classified					
		Cough	Throat Clear	/t/	/s/	/i/	/u/
Actual	Cough	96.97	3.33	0.00	0.00	0.00	0.00
	Throat Clear	16.67	80.00	0.00	3.33	0.00	0.00
	/t/	0.00	0.00	90.00	8.33	1.67	0.00
	/s/	3.33	1.67	8.33	86.67	0.00	0.00
	/i/	35.00	3.33	0.00	1.67	50.00	10.00
	/u/	28.33	13.33	3.33	1.67	16.67	36.67
						Overall	73.33

Table II
CONFUSION MATRIX FOR THE HIERARCHICAL GUSSS CLASSIFIER. THE VALUES ARE AVERAGE PERCENTAGES OVER A 10-FOLD CROSS VALIDATION.

		Classified					
		Cough	Throat Clear	/t/	/s/	/i/	/u/
Actual	Cough	98.33	1.67	0.00	0.00	0.00	0.00
	Throat Clear	3.33	93.33	3.33	0.00	0.00	0.00
	/t/	0.00	0.00	95.00	5.00	0.00	0.00
	/s/	0.00	0.00	6.67	93.33	0.00	0.00
	/i/	1.67	0.00	3.33	0.00	76.67	18.33
	/u/	1.67	0.00	1.67	0.00	13.33	83.33
						Overall	90

Hierarchical GUSSS can be further exploited by an automatic selection of the tuples.

Despite the many promising implications for the future, the first aspect that needs to be explored is the use of the device for both intra- and inter-subject gesture recognition, including data from both sexes and across lifespan. In that sense, reference levels for muscle recordings such as maximum voluntary muscle contractions (MVC) should also be considered as a normalization technique to allow for comparisons between subjects [15], [9]. In that case, each MVC trial should be repeated a number of times with a rest period between trials to limit fatigue.

Also, a larger gesture set should be explored, for example, isometric resistive mandible depression task (depression of jaw against manual resistance provided by the participant with lightly closed lips and teeth [9].

Finally, the effects of different positions of the sEMG sensors and variations on the number of sensors used should also be explored – i.e. a reduced number of sEMG sensors for both aesthetics and subject comfort. All these improvements can make the ECHO-OREAD the first effective, non-invasive and ambulatory tool for the detection and tracking of normal and maladaptive conditions underlying speech and non-speech laryngeal behaviors, including swallowing.

REFERENCES

[1] N. Roy, R. Merrill, S. Thibeault, R. Parsa, S. Gray, and E. Smith, "Prevalence of voice disorders in teachers and the general population," *J. Speech. Lang. Hear. Res*, vol. 47, pp. 281–293, 2004.

[2] K. Verdolini, C. Rosen, and R. Branski, *Classification Manual for Voice Disorders*. Lawrence Erlbaum Associates, 2006.

[3] R. Hillman, J. Heaton, A. Masaki, S. Zeitels, and H. Cheyne, "Ambulatory monitoring of disordered voices," *Ann Otol Rhinol Laryngol*, vol. 115, no. 11, pp. 795–801, 2006.

[4] P. Popolo, J. Švec, and I. Titze, "Adaptation of a pocket pc for use as a wearable voice dosimeter," *J. Speech. Lang. Hear. Res*, vol. 48, no. 4, pp. 780–791, 2005.

[5] D. Mehta, M. Zañartu, S. Feng, H. C. II, and R. Hillman, "Mobile voice health monitoring using a wearable accelerometer sensor and a smartphone platform," in *IEEE Trans. on Biom. Eng.*, pp. 3090–3096, 2012.

[6] S. Joshi, A. Wexler, C. Perez-Maldonado, and S. Vernon, "Brain-muscle-computer interface using a single surface electromyographic signal: Initial results," in *Neural Engineering (NER), 2011 5th International IEEE/EMBS Conference on*, pp. 342–347, April 2011.

[7] K. Singh, R. Bhatia, and H. Ryait, "Precise and accurate multi-functional prosthesis control based on fuzzy logic techniques," in *Communication Systems and Network Technologies (CSNT), 2011 International Conference on*, pp. 188–193, June 2011.

[8] C. Stepp, J. Heaton, M. Braden, T. Stadelman-Cohen, M. Jetté, and R. Hillman, "Comparison of neck tension palpation rating systems with surface electromyographic and acoustic measures in vocal hyperfunction," *Journal of Voice*, vol. 25, 2011.

[9] C. Stepp, "Surface electromyography for speech and swallowing systems: measurement, analysis, and interpretation," *J. Speech. Lang. Hear. Res*, vol. 55, pp. 1232–1246, 2012.

[10] L. A. Rivera and G. N. DeSouza, "Recognizing hand movements from a single semg sensor using guided under-determined source signal separation," in *12th IEEE Int. Conf. on Rehab. Robotics*, pp. 450–455, Jun. 2011. ETH Zurich, Switzerland.

[11] L. A. Rivera and G. N. DeSouza, "A power wheelchair controlled using hand gestures, a single semg sensor, and guided under-determined source signal separation," in *4th IEEE Int. Conf. on BioRob*, pp. 1535–1540, Jun. 2012. BioRob2012, Rome, Italy.

[12] Y. Guangying, "Study of myoelectric prostheses hand based on independent component analysis and fuzzy controller," in *8th Int. Conf. on Electronic Measurmnt and Instrumnts*, pp. 1–174–1–178, Aug. 2007.

[13] J. G. Webster, ed., *Encyclopedia of Medical Devices and Instrumentation. Electromyography.* John Wiley, 2006.

[14] L. A. Rivera, N. R. Smith, , and G. N. DeSouza, "High-accuracy recognition of muscle activation patterns using a single sEMG signal," in *IEEE BioRob*, pp. 579–584, Aug. 2014. Sao Paulo, Brazil.

[15] A. Fridlund and J. Cacioppo, "Guidelines for human electromyographic research," *Psychophysiology*, vol. 23, no. 5, pp. 567–589, 1986.

[16] T. Hixon, G. Weismer, and J. Hoit, *Preclinical speech science. Anatomy, physiology, acoustics, perception, 2nd ed.* Plural, 2013.

[17] M. Dietrich and K. V. Abbott, "Vocal function in introverts and extraverts during a psychological stress reactivity protocol," *Journal of Speech, Language, and Hearing Research*, vol. 55, no. 3, pp. 973–87, 2012.

[18] R. Ding, C. Larson, J. Logemann, and A. Rademaker, "Surface electromyographic and electroglottographic studies in normal subjects under two swallow conditions: Normal and during the mendelsohn manuever," *Dysphagia*, vol. 17, pp. 1–12, 2002.

[19] A. VanBoxtel, "Optimal signal bandwidth for the recording of surface EMG activity of facial, jaw, oral, and neck muscles," *Psychophysiology*, vol. 38, pp. 22–34, 2001.

[20] E. Yiu, K. Verdolini, and L. Chow, "Electromyographic study of motor learning for a voice production task," *J. Speech. Lang. Hear. Res*, vol. 48, no. 6, pp. 1254–1268, 2005.