ROBOMECH Journal

RESEARCH ARTICLE

Tongue interface based on surface EMG signals of suprahyoid muscles

Makoto Sasaki[1*], Kohei Onishi[1], Dimitar Stefanov[2], Katsuhiro Kamata[3], Atsushi Nakayama[4], Masahiro Yoshikawa[5] and Goro Obinata[6]

Abstract

The research described herein was undertaken to develop and test a novel tongue interface based on classification of tongue motions from the surface electromyography (EMG) signals of the suprahyoid muscles detected at the underside of the jaw. The EMG signals are measured via 22 active surface electrodes mounted onto a special flexible boomerang-shaped base. Because of the sensor's shape and flexibility, it can adapt to the underjaw skin contour. Tongue motion classification was achieved using a support vector machine (SVM) algorithm for pattern recognition where the root mean square (RMS) features and cepstrum coefficients (CC) features of the EMG signals were analyzed. The effectiveness of the approach was verified with a test for the classification of six tongue motions conducted with a group of five healthy adult volunteer subjects who had normal motor tongue functions. Results showed that the system classified all six tongue motions with high accuracy of 95.1 ± 1.9 %. The proposed method for control of assistive devices was evaluated using a test in which a computer simulation model of an electric wheelchair was controlled using six tongue motions. This interface system, which weighs only 13.6 g and which has a simple appearance, requires no installation of any sensor into the mouth cavity. Therefore, it does not hinder user activities such as swallowing, chewing, or talking. The number of tongue motions is sufficient for the control of most assistive devices.

Keywords: Tongue interface, Motion classification, Support vector machine (SVM), Surface electromyography (EMG), Suprahyoid muscles

Background

A tongue is an intra-oral locomotorium that can be moved quickly and precisely according to one's own will. Anyone can set their own tongue position precisely and can smoothly change the magnitude of the force imposed on the teeth or palate. In fact, tongue motor functions are usually preserved even in people with cervical spinal cord damage. Various studies have demonstrated that people with a high level of movement paralysis can use tongue motions to control home appliances such as a PCs and electric wheelchairs [1, 2].

An interface system based on a small joystick operated by the tongue has been presented in the literature [3]. The joystick is fixed in a suitable position via a special arm mount. The application of such an interface is limited to people with sufficient head motion that is able to reach the joystick. The same design might hinder conversation, eating, and drinking because a part of the joystick is located in the mouth cavity during use. In addition, such a solution might allow flow of excess saliva out of the mouth.

Numerous studies have examined control interfaces containing an artificial palate with buttons activated by the tongue tip [4–6]. In other solutions, a few pairs of light emitting diodes and photodiodes are mounted on an artificial palate to detect the tongue position [7, 8]. By changing the tongue position and thereby activating different sensors, the user sets control commands. A benefit of such solutions is that the tongue interface device remains hidden to others (intra-oral interface). In addition, the number of sensors and their location on the artificial palate can be customized easily. However, because the artificial palate must remain in the mouth cavity for prolonged periods of time, such a design might require

*Correspondence: makotosa@iwate-u.ac.jp
[1] Graduate School of Engineering, Iwate University, Morioka, Iwate, Japan
Full list of author information is available at the end of the article

Sasaki *et al. Robomech J* (2016) 3:9

Page 2 of 11

additional efforts for maintaining oral hygiene and might entail various difficulties related to talking and eating.

Some recent studies present solutions that consist of a small magnet or a piece of ferromagnetic material attached to the tongue tip via gluing or piercing, and a sensor array that detects the tongue tip position [9–13]. Sensor systems that include a small permanent magnet fixed to the tongue tip and an array of magnetic sensors have been presented in the literature [9–11]. Some earlier reports introduced a tongue interface that includes an air-cored coil with inductance changed by a small ferromagnetic stud attached to the tongue [12, 13]. Although such an approach is a simple interface solutions, the (ferro) magnetic stud on the tongue tip might cause some inconvenience to users.

The electromyography (EMG) signals created by skeletal muscles have been used for many years in human movement studies and for control of prostheses [14–26]. Our earlier study specifically addressed the potential of EMG signals and explored the viability of a tongue interface based on surface EMG signals detected at the underside of the jaw [27]. The initial interface system consisted of nine single-surface electrodes attached on the underside of the jaw and connected via multiple lead wires. The proposed tongue interface, which is based entirely on analysis of extra-oral EMG signals, requires no insertion of a palatal plate or a joystick in the mouth, attachment of a magnet or ferromagnetic studs to the tongue, or physical contact of the tongue with any sensor. An artificial neural network (ANN) with three layers of neurons (input, hidden and output) was used as the motion classifier. During the ANN training stage, three thin-film force sensors were installed on an upper jaw mouthpiece to deliver training data for voluntary tongue motions of three types: right, left, and forward. After this initial experiment, a new experiment was conducted for classification of the same voluntary motions without using signals from force sensors [28]. These initial experiments demonstrated that the tongue motions are classifiable from the EMG signals of the suprahyoid muscles. They have some potential for use in control interfaces. However, the initial interface system classified only a small number of tongue motions. Additionally, ANN-based classifiers are well known to have a few important shortcomings: long learning time, local optimal solution depending on the initial value of parameters, and complicated procedures for selection of the number of neurons in the hidden layer. Furthermore, the initial sensor module consists of single electrodes and wires, which can pose severe difficulties. For practical application, the initial interface system required further improvement in few main directions: increased number of classified voluntary tongue motions, improvement of the classification accuracy, and redesign of the electrode module.

This paper proposes a novel tongue interface based on classification of the tongue motions from surface EMG signals of the suprahyoid muscles detectable at the underside of the jaw. The interface allows classification of six tongue motions, which are sufficient for the control of PCs and electric wheelchairs. The new system was evaluated using a computer simulation experiment to assess control of an electric wheelchair.

EMG-based tongue interface
EMG measurement approach
Tongue motions are produced by the coordinated actions of intrinsic muscles, which control tongue posture and tongue tip position, and extrinsic muscles, which control tongue protrusion and retraction [29, 30]. The EMG activity of the lingual muscles has been studied using tungsten microelectrodes and hook-wire electrodes [31] and surface electrodes [32] placed within the oral cavity. However, intra-oral electrodes are unsuitable for the practical control of assistive devices.

The EMG signals of the suprahyoid muscles are detectable via electrodes placed on the skin of the underside of the jaw [33–35]. The suprahyoid muscles comprise several muscle groups such as digastric muscles, stylohyoid muscles, mylohyoid muscles, and geniohyoid muscles, as presented in Fig. 1 [29, 30]. The suprahyoid muscles control the position of the hyoid according to the direction, position, and force of the tongue tip. Therefore, they contain sufficient information about the performed tongue motions. However, the suprahyoid muscles also contribute to motions that are unrelated to the tongue position. Such motions produce EMG signals when such motions are performed. For example, suprahyoid muscles help jaw-opening by pulling the mandible down when the hyoid position is fixed by the infrahyoid muscles. They also pull the hyoid up to assist swallowing when the mandible position is fixed to the muscles used for mastication. A great challenge to the design of a reliable tongue interface is the identification and suppression of EMG signals that do not originate from voluntary tongue motions. That difficulty cannot be resolved merely by electrode positioning because measured EMG signals are always composed of several signals from different muscles around the electrode.

In this study, the EMG signals of the suprahyoid muscles are measured at multiple points of the skin using a multi-electrode array. The multi-electrode approach makes the interface system less sensitive to eventual positioning errors of the electrode unit. Moreover, it enables people with little experience or little knowledge of EMG measurement to apply the sensor. Current research was based on initial experiments conducted for the classification of tongue motions from the EMG signals patterns [27, 28].

Sasaki *et al. Robomech J* (2016) 3:9

Page 3 of 11

Fig. 1 Structure of suprahyoid muscles
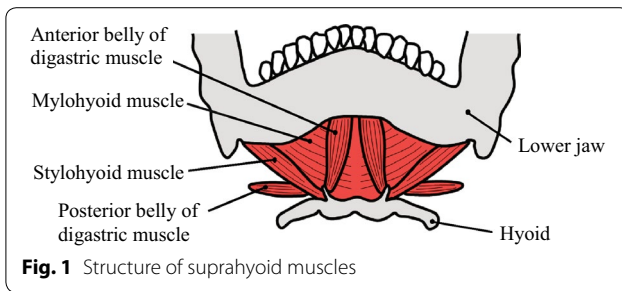
Sensor module and signal pre-processing

The electrode module was designed as a thin flexible boomerang-shaped patch attached to the underside of the jaw (Fig. 2). The prototype sensor dimensions were decided by considering the average size of the lower jaw and curvature near the lower jaw and neck of the subjects in the tests (see "Experiments and data acquisition" section below). The sensor was designed to cover the entire jaw. The number of the electrodes was determined experimentally. The electrodes were positioned on equal inter-electrode distances. The interface assembly, which consisted of 22 active electrodes shaped as $\phi 2 \times 2.5$ mm

a surface view   **b** rear view

c positioning of the electrodes

d schematic view of the position of the pin electrodes

Fig. 2 22-channel active electrode

pure silver rods, was positioned at the inter-electrode distance of 12.5 mm on a polyimide film. The interface unit was 50.0-mm-long and 87.5-mm-wide. The thickness of the entire substrate including the reinforcement film was 0.3 mm. The electrode tips were shaped as hemispheres to facilitate the skin contact. Voltage follower circuits were incorporated into the same interface mount to reduce the output impedance. The electrode base thickness was 1.7 mm. For electric insulation of the electronic parts, both sides of the substrate were covered with a layer of silicon. The interface system was only 13.6 g. For the experiments, the interface module was adhered to the underside of the jaw of the subject. A ground electrode and an active common electrode were connected respectively to the left and right earlobes via ear clips (Fig. 2c). The electric potential between each electrode and the active common electrode was amplified using a separate differential amplifier. The gain of the differential amplifiers was set to 2052. A band-pass filter with a passband from 16 to 440 Hz bandwidth was used to remove the direct current component and high-frequency noise superimposed on the EMG signals. The EMG signals of all 22 EMG channels were digitized by a 16 bit analog-to-digital converter (USB-6218; National Instruments Corp.). In general, the EMG signal frequency range is 0–1000 Hz. Its usable energy is limited to 0–500 Hz [36]. Therefore, the sampling rate was set to 2000 Hz in compliance with the Nyquist theorem.
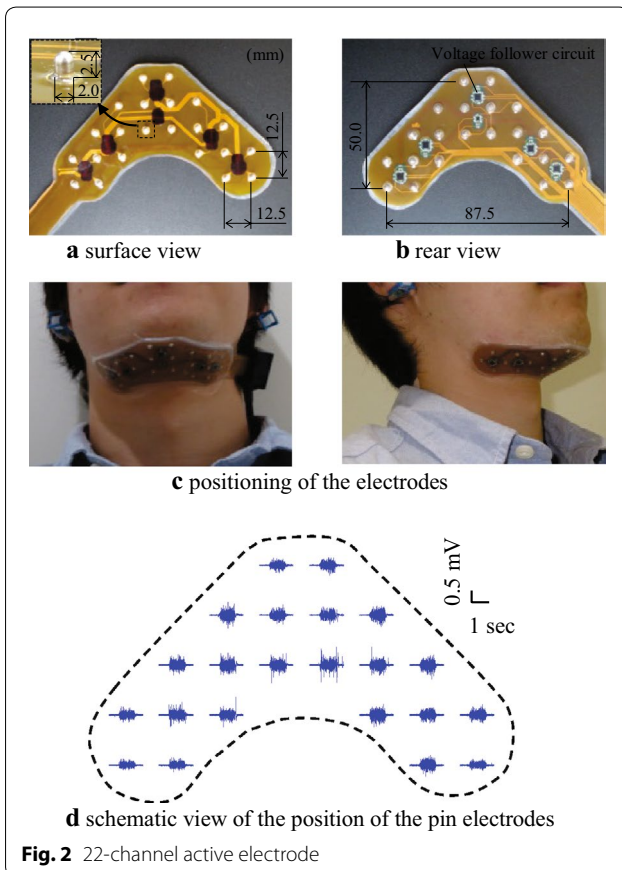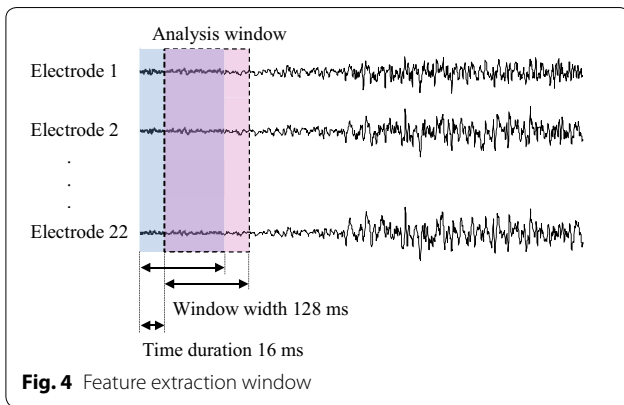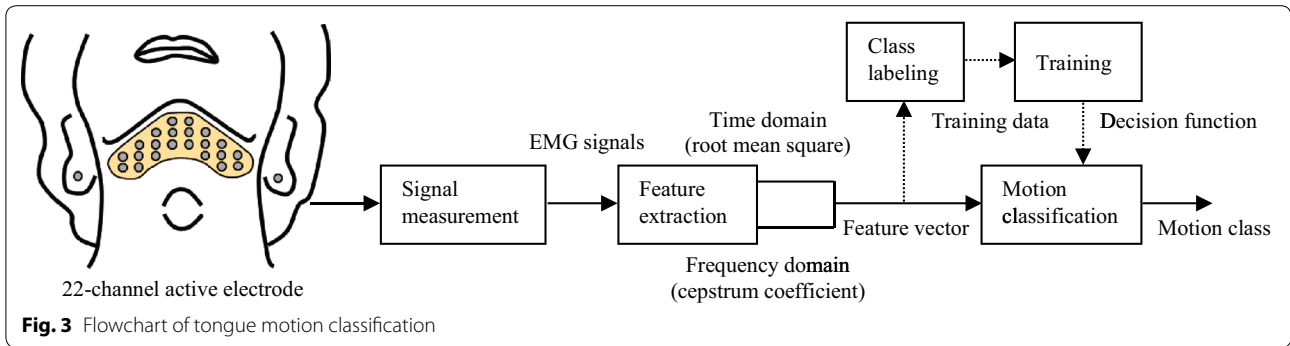
Classification of tongue motions

Figure 3 portrays the tongue motion classification procedure. It comprises the EMG measurement, feature extraction, and motion classification.

Feature extraction

The feature extraction process was based on the overlapped windowing technique proposed by Englehart et al. [37]. It allows faster system response. The EMG signals measured from all 22 channels were segmented for feature extraction into windows consisting of 256 samples, as portrayed in Fig. 4. The length of each window was 128 ms. The next sampling segment slides over the current segment with an increment time of 16 ms. For composition of the feature vector, the root mean square (RMS) and the cepstrum coefficients (CC) of the EMG signals were calculated for each window. The RMS features are characteristics of a time domain. The CC features are characteristics of the frequency domain [28, 38, 39]. Cepstrum analysis techniques have been used for many years for speech recognition because of their fast response and accurate results. Some recent studies have demonstrated that the techniques are useful also for motion classification based on EMG signals [39–41].

Sasaki *et al. Robomech J* (2016) 3:9

Page 4 of 11

Fig. 3 Flowchart of tongue motion classification

Fig. 4 Feature extraction window

The RMS features provide information related to the amplitude of the EMG signals. Let us denote the EMG signals of the l-th electrode in the n-th sample of the p-th analysis window as $EMG_{l,n}(p)$ ($n = 0, ..., N - 1$; $l = 1, ..., L$), where N is the number of samples in one analysis window ($N = 256$), and L is the number of electrodes ($L = 22$). The RMS features can be expressed as the following equation:

$$RMS_l(p) = \sqrt{\frac{1}{N} \sum_{n=0}^{N-1} EMG_{l,n}(p)^2} \qquad (1)$$

The equation above is useful for calculation of the RMS features of all channels.

To calculate the CC features, the Hanning window procedure was applied to each analysis window of the EMG signals. The Fourier transform $X_l^k(p)$ ($k = 0, ..., N - 1$) of $EMG_{l,n}(p)$ can be expressed as shown below.

$$X_l^k(p) = \sum_{n=0}^{N-1} EMG_{l,n}(p)e^{-j2\pi kn/N} \qquad (2)$$

The CC features $CC_l^n(p)$ are calculated from the following equation.

$$CC_l^n(p) = \frac{1}{N} \sum_{k=0}^{N-1} \log\left|X_l^k(p)\right| e^{j2\pi kn/N} \qquad (3)$$

Cepstrum analysis enables separation of the power spectrum of the EMG signals into a smooth component (spectral envelope) and a fine fluctuation component (fine structure). Low-order cepstrum coefficients include information about the spectral envelope whereas the high-order coefficients include fine structure information. The low-order coefficients were calculated using formula (3) and by varying n from $n = 0$ to $n = W - 1$. Here, W is a CC feature parameter (order of the cepstrum coefficients).

The feature vector $\boldsymbol{x}(p)$ for classifying tongue motions can be expressed as

$$\begin{aligned} \boldsymbol{x}(p) = (&RMS_1(p), \ldots, RMS_L(p), \\ &CC_1^0(p), \ldots, CC_1^{W-1}(p), \ldots, \\ &CC_L^0(p), \ldots, CC_L^{W-1}(p))^T \end{aligned} \qquad (4)$$

where the dimension of the feature vector $\boldsymbol{x}(p)$ is $L(1 + W)$.

Motion classification

For this study, the support vector machine (SVM) classifier was used to classify tongue motions. The SVM classifier has the following benefits for this classification:

– The SVM classifier offers excellent recognition performance.
– SVM has high generalization capability because it applies a maximum-margin classification function.
– It converges to a global optimal solution and therefore does not fall into a local optimum solution.
– It has extremely short learning time because of the simple procedures used for calculation of the hyperparameters used for training.

SVM is a method for classification of an unknown feature vector $\boldsymbol{x}(p)$ (hereinafter designated as $\boldsymbol{x}$) into two classes [42]. The decision function is

$$f(\boldsymbol{x}) = sgn\left(\sum_{i=1}^{D} \lambda_i y_i K(\boldsymbol{x}_i, \boldsymbol{x}) + b\right) \qquad (5)$$

Sasaki *et al. Robomech J* (2016) 3:9

Page 5 of 11

where D denotes the number of training samples, y_i signifies the class label that corresponds to the i-th training sample $\boldsymbol{x}$, λ_i is a Lagrangian undetermined multiplier, b is a bias term, and $K(\boldsymbol{x}_i, \boldsymbol{x})$ denotes a kernel function. For this study, the radial basis function (RBF) was selected as the kernel function to map the input data in a high dimensional feature space. The RBF kernel is expressed as

$$K(\boldsymbol{x}_i, \boldsymbol{x}) = \exp(-\gamma||\boldsymbol{x}_i - \boldsymbol{x}||^2) \tag{6}$$

where γ is a kernel parameter. The Lagrangian undetermined multiplier λ_i in the decision function is derived by solving the following equation (quadratic programming).

$$\max_{\lambda_i} \ \sum_{i=1}^{D} \lambda_i - \frac{1}{2} \sum_{i,j=1}^{D} \lambda_i \lambda_j y_i y_j K(\boldsymbol{x}_i, \boldsymbol{x}) \tag{7}$$

$$\text{subject to} \quad \sum_{i=1}^{D} \lambda_i y_i = 0, \quad 0 \le \lambda_i \le C \tag{8}$$

The SVM classification performance depends on the selection of the kernel parameter γ and the penalty parameter C. The optimal combination of γ and C can be obtained using a grid search.

Usually, the SVM classifier is used for classification of features into two classes. In this study, the SVM algorithm was extended to multi-class classification using the one-against-one method [43]. For the classification of M classes tongue motions, $M(M-1)/2$ decision functions are constructed initially for all combinations of these M classes. The feature vector $\boldsymbol{x}$ is classified against each decision function. The final decision on the class is obtained by majority vote.

Experiments and data acquisition
Subjects
This investigation examined five healthy adult male subjects (22.2 ± 1.3 years old, 169.7 ± 7.4 cm tall, 61.0 ± 11.3 kg weight) who were free of musculoskeletal deficits and neurological impairment and who had normal tongue motor functions. Approval for the tests was obtained in advance by the Ethical Review Board of Iwate University. Before the start of the tests, the study objective, experimental protocol and risks were explained to each subject. Written consent was received from each.

Experimental protocol
First, the skin surface of the underside of the jaw was cleaned with alcohol and electrode paste (Elefix; Nihon Kohden Corp.) was applied to reduce the skin-electrode impedance. The 22-channel active electrode was adhered to the underside of the subject's jaw using film dressing (CATHEREEPLUS; Nichiban Co. Ltd.). A ground electrode and an active common electrode were attached on the left and right earlobe of the subject using ear clips.

The tongue motion set included five tongue motions (right, left, up, down, and forward) performed with a closed mouth and a saliva swallowing (Fig. 5). During these motions, subjects were asked to position their tongue tips sequentially in the maxillary right second molar tooth, the maxillary left second molar tooth, the hard palate, the floor of the mouth, and near the maxillary central incisor. Saliva swallowing is an unintentional action that is repeated frequently. The saliva swallowing was included in tests to evaluating its effects on tongue motion classification. In the experiment, each tongue motion was executed for 2 s at a subject's comfortable speed. A resting period of 2 s was given to the subject before the start of the next motion. Consequently, all six motions in the set were completed for 22 s. Each subject was asked to perform the motion set 14 times. The EMG signals during each test were recorded. As a result, 14 datasets were produced for each subject.

Data analysis
Matlab (R2013a; The MathWorks Inc.) was used for data analysis. The SVM classification algorithm was designed using an SVM library: LIBSVM [44]. The programs were executed on a PC (Windows 7 64-bit OS, i7-3770 CPU/3.4 GHz, 16 GB RAM).

To justify the selection of the kernel function, it was confirmed that the RBF kernel matrix calculated from the first four datasets is a symmetric, positive semi-definite matrix (i.e., all eigenvalues of the kernel matrix are non-negative). Then the datasets were used as training data of the SVM. The remaining ten datasets were used for tongue motion classification. The feature vector $\boldsymbol{x}$ for tongue motion classification was defined according to Eq. (4). The values of the RMS features and CC features were calculated, respectively, according to Eq. (1) and Eq. (3). The class labels y_i representing the type of motion in Eq. (5) were obtained using threshold triggering of the EMG signals [45]. The relation between the composition of the feature vector and its classification accuracy was evaluated by comparing the classification results when the CC feature parameter W was varied from 0 to 10. For simplicity in these analyses, $W = 0$ expresses the situation when the CC features are not included in the feature vector.
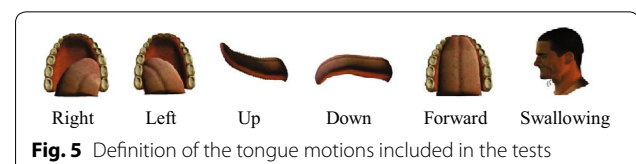

Right   Left   Up   Down   Forward   Swallowing

Fig. 5 Definition of the tongue motions included in the tests

Sasaki *et al. Robomech J* (2016) 3:9

Page 6 of 11

As explained in the section describing "Motion classification", the SVM classification performance depends on selection of kernel parameter γ and penalty parameter C. The optimal combination of γ and C was ascertained using a grid search within the training data. The search included 96 combinations of γ and C for $\gamma = \{2^{-10}, 2^{-9}, \ldots, 2^{1}\}$ and $C = \{2^{1}, 2^{2}, \ldots, 2^{8}\}$. The combination with the highest classification rate was defined using fivefold cross validation. Results showed that the optimum values of γ and C differ for each subject. After training of the SVM with the optimized hyperparameters for γ and C, motion classification of the test data was performed. The predicted class was replaced with a "neutral" tongue position when all EMG signals are under the threshold level (i.e., relaxed state). Next, a majority voting technique was applied to reduce the effect of misclassification. Majority voting was applied to a moving window composed of 20 frames that included the present frame and the prior 19 frames. Classification of the tongue motion was determined from the class with the largest number of wins.

The classification accuracy (CA) of the tongue motions was evaluated using the following equation.

$$CA = \frac{\text{number of correct feature vectors}}{\text{total number of feature vectors}} \times 100\,[\%] \quad (9)$$

Results

Effect of feature parameter selection on classification accuracy

The average classification accuracy and the standard deviation of the classification accuracy for all five subjects are presented in Fig. 6. Results reveal the relation between the feature vector and classification accuracy. In cases where the feature vector was composed of RMS features only ($W = 0$), the classification accuracy of the tongue motions was 84.1 ± 1.5 %. The classification accuracy

Fig. 6 Relation between the composition of the feature vector and classification accuracy. $W = 0$ means that the CC features are not included in a feature vector
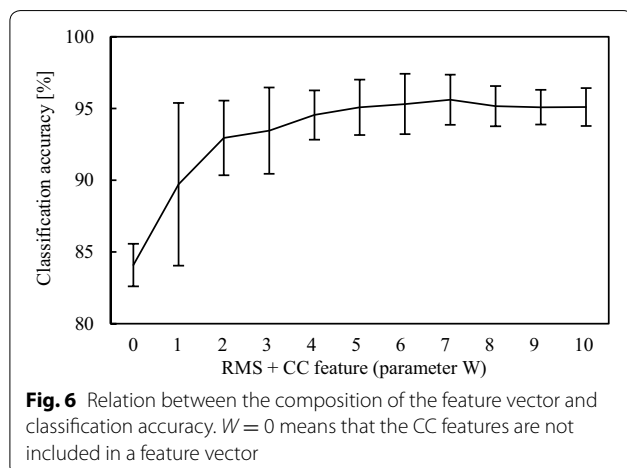
increased substantially when the CC features were added to the feature vector ($W = 1, \ldots, 10$). The classification accuracy exceeded 95 % and remained almost constant when the CC feature parameter was $W = 5$ or higher. The classification accuracy for $W = 5$ was 95.1 ± 1.9 %. For $W = 10$, the classification accuracy was 95.1 ± 1.3 %. No significant difference was found between the classification accuracies calculated with $W = 5$ and $W = 10$.

The dimension of the feature vector $\boldsymbol{x}$ for tongue motion classification was set to $L(1 + W)$ (see Eq. (4)). Because the computational complexity increases significantly for greater values of W, the smallest possible W that gives comparable classification accuracy should be used. As explained above, no significant difference was found between the classification results with $W = 5$ and $W = 10$, which suggests that satisfying classification results are obtainable with a feature vector based on $W = 5$. For that reason, a more detailed examination of the classification results is given here for the case in which the CC feature parameter was selected as $W = 5$.

Tongue motion classification accuracy

Table 1 presents classification results for all five subjects. The lowest total classification accuracy was 91.9 % (for subject A) and the highest total classification accuracy was 96.7 % (for subject B). The average total classification accuracy for all subjects was 95.1 %. Analysis of the classification results for the separate tongue motions demonstrates that the "left" tongue motion was recognized with the highest classification accuracy (97.6 %). The classification accuracy for the "down" tongue motion was slightly lower (96.7 %), followed by results for "saliva swallowing" (95.3 %), "right" tongue motion (95.0 %), "up" tongue motion (94.5 %), and "forward" tongue motion (91.4 %). Table 2 presents details of the classification errors. The "forward" tongue motions were misclassified as "up", "down", and "saliva swallowing". The "up" tongue motion has the second lowest classification accuracy. Frequently, it has been misclassified as "right" tongue motion.

Short signals at the start and the end of the main motion were often misclassified. By applying majority voting technique, the number of these misclassifications was reduced; 1.0 % of all motions were misclassified as a "neutral" tongue position. However, misclassification as a "neutral" tongue position is less important because the "neutral" tongue position is useful as a stop command when the assistive device is controlled by the developed interface. Misclassification of other motions as a "neutral" tongue position cannot create dangerous situations. It will merely cause the controlled device to stop. Overall, the misclassification errors that might affect the operation of the controlled assistive devices were estimated from the total classification accuracy as about 3.9 %.

Sasaki *et al. Robomech J (2016) 3:9*

Page 7 of 11

Table 1 Classification accuracy of tongue motions

Subject	Right	Left	Up	Down	Forward	Saliva swallowing	Total
A	96.1	93.8	84.9	98.3	88.3	90.1	91.9
B	94.1	97.8	97.4	95.8	96.4	98.9	96.7
C	99.5	98.8	97.6	99.9	90.0	92.9	96.4
D	89.1	100	96.0	92.4	95.4	96.2	94.9
E	96.2	97.8	96.5	97.2	87.0	98.4	95.5
Mean	95.0	97.6	94.5	96.7	91.4	95.3	95.1

Computer simulation of wheelchair control

A computer simulation model of an electric wheelchair was developed to evaluate the applicability of the developed tongue interface to control assistive devices. The wheelchair model was controlled virtually by operation commands generated from a confusion matrix presented in Table 2. The error of the commands was set to occur according to the possibility in the confusion matrix. In other words, this is a Monte Carlo method. The error timing was determined using a random number with a uniform probability distribution. The virtual trajectory of the wheelchair's center of gravity was used as an indicator to evaluate the effects of misclassification errors on the wheelchair operability.

Simulation model of an electric wheelchair

Figure 7 portrays a simplified model of the electric wheelchair. The angle θ and the center of gravity position P_G (X_G, Y_G) of an electric wheelchair are defined using the following equations.

$$\theta(t) = \frac{1}{T} \int_0^t (R_r \omega_r(t) - R_l \omega_l(t)) dt \tag{10}$$

$$X_G(t) = \frac{1}{2} \int_0^t (R_r \omega_r(t) + R_l \omega_l(t)) \cos \theta(t) dt \tag{11}$$

$$Y_G(t) = \frac{1}{2} \int_0^t (R_r \omega_r(t) + R_l \omega_l(t)) \sin \theta(t) dt \tag{12}$$

Therein, R_r and R_l respectively denote the radii of the right and left wheel. $\omega_r(t)$ and $\omega_l(t)$ respectively denote the angular velocity of the right and the left wheel. T is the distance between the right and left wheels. For the wheelchair model, wheels with radius 165 mm were selected. The distance between the wheels was 530 mm.

The maximum velocity of the electric wheelchair model V_{max} was set to 4 km/h. The model is based on a trapezoidal model of acceleration and deceleration. The acceleration time T_a and the deceleration time T_d were set to 1 s. In this simulation, a new operation command is sent to the virtual wheelchair every $T_i = 16$ ms because, in the tongue motion classification experiment, the EMG signals were classified at 16 ms intervals (see Fig. 4). Therefore, velocity commands are sent to the right and the left wheel every 16 ms. These commands are based on the rules presented in Table 3. The wheel velocities $R_r \omega_r(t)$ and $R_l \omega_l(t)$ are defined by the following equations.

$$R_r \omega_r(t) = V_{max} S_r(t) T_i / T_a \tag{13}$$

$$R_l \omega_l(t) = V_{max} S_l(t) T_i / T_a \tag{14}$$

Table 2 Confusion matrix for six tongue motions

	Estimated class						
	Right	Left	Up	Down	Forward	Saliva swallowing	Neutral
Actual class							
Right	*95.0*	0.1	1.6	0.4	2.1	0.0	0.8
Left	1.2	*97.6*	0.1	0.2	0.0	0.0	0.9
Up	2.6	0.4	*94.5*	0.4	0.7	0.2	1.2
Down	0.0	0.0	0.7	*96.7*	1.9	0.3	0.4
Forward	0.4	0.0	2.2	3.4	*91.4*	1.2	1.4
Saliva swallowing	0.2	0.0	0.9	0.3	2.2	*95.3*	1.1

Sasaki *et al. Robomech J* (2016) 3:9

Page 8 of 11

Fig. 7 Simplified model of an electric wheelchair
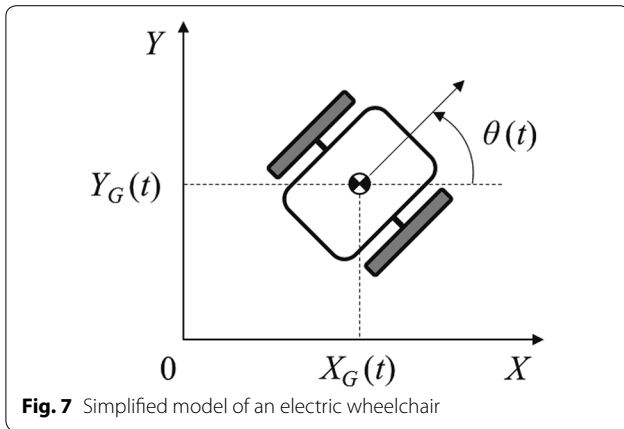
Table 3 Change amount of velocity commands for right and left wheel

Command	$S_r(t)$	$S_l(t)$
Forward	$+1$	$+1$
Back	-1	-1
Right rotation	-1	$+1$
Left rotation	$+1$	-1
Brake	$\begin{cases} +1 & \text{if } R_r\omega_r(t) < 0 \\ -1 & \text{if } R_r\omega_r(t) > 0 \end{cases}$	$\begin{cases} +1 & \text{if } R_l\omega_l(t) < 0 \\ -1 & \text{if } R_l\omega_l(t) > 0 \end{cases}$
None	$+0$	$+0$

Therein, $S_r(t)$ and $S_l(t)$ respectively represent the commands sent to the right and the left wheel in sequential moments of time. $S_r(t)$ and $S_l(t)$ are defined as follows.

$$-T_a \big/ T_i \leq S_r(t) \leq T_a \big/ T_i \tag{15}$$

$$-T_a \big/ T_i \leq S_l(t) \leq T_a \big/ T_i \tag{16}$$

Linking tongue motions with commands for control of the wheelchair model

Commands for the wheelchair model operation are presented in Table 4. They are based on the confusion matrix of tongue motions, as shown in Table 2. Initially, the "Brake" command is set via the "neutral" tongue position.

Table 4 Definition of operation commands

Operation commands	Tongue motions
Forward	Down
Back	Forward
Right rotation	Right
Left rotation	Left
Brake	Neutral
None	Saliva swallowing, up

It is assumed that the "Brake" command is sent to the wheelchair when all EMG signals are under the threshold level (i.e., relaxed state). The "Brake" command causes the wheelchair to decelerate and stop. The "Forward" command was linked with the "down" tongue motion because the probability for misclassification of the "down" tongue motion as "right" or "left" is nearly zero in the confusion matrix. "Right" and "left" tongue motions were used, respectively, as commands for turning of the wheelchair model to the right and left. Reverse wheelchair movement ("back" command) is initiated by "forward" tongue motion.

The classification accuracy of "forward" tongue motion was lower than that of "up" tongue motion. However, the rate of misclassification of "forward" tongue motion as "right" tongue motion is much lower (0.4 %) than that of "up" tongue motion (2.6 %). Its characteristic means that "forward" tongue motion ensures the straight driving performance. In addition, although "forward" tongue motion is misclassified as "down" tongue motion as about 3.4 %, it does not affect the straight driving performance so much because this misclassification reduces the driving velocity while moving backward.

The remaining "saliva swallowing" and "up" tongue motions were defined as no command.

The driving test consisted of six tasks:

E1 Driving the wheelchair forward 5 m
E2 Driving the wheelchair backward 5 m
E3 Turning the wheelchair 360° to the right
E4 Turning the wheelchair 360° to the left
E5 Swallowing saliva while the wheelchair model is stopped
E6 Swallowing saliva while the wheelchair model is moving straight at maximum velocity

The saliva swallowing times in tests E5 and E6 were set to 1 s.

In this situation, the 100 patterns of velocity commands of right and left wheel S_r, S_l considering that the rate of misclassification as shown in Table 2 was generated using a random function. Then, these resultant trajectories were compared with the ideal trajectory, which was calculated as a classification accuracy of all tongue motions is 100 %.

Simulation results

The simulation results of the angle θ and the center of gravity position P_G (X_G, Y_G) of an electric wheelchair are presented in Fig. 7. In addition, the differences between the ideal trajectory and the trajectory including the effect of misclassification are presented in Fig. 8.

Sasaki *et al. Robomech J* (2016) 3:9

Page 9 of 11

Fig. 8 Simulation results of an electric wheelchair
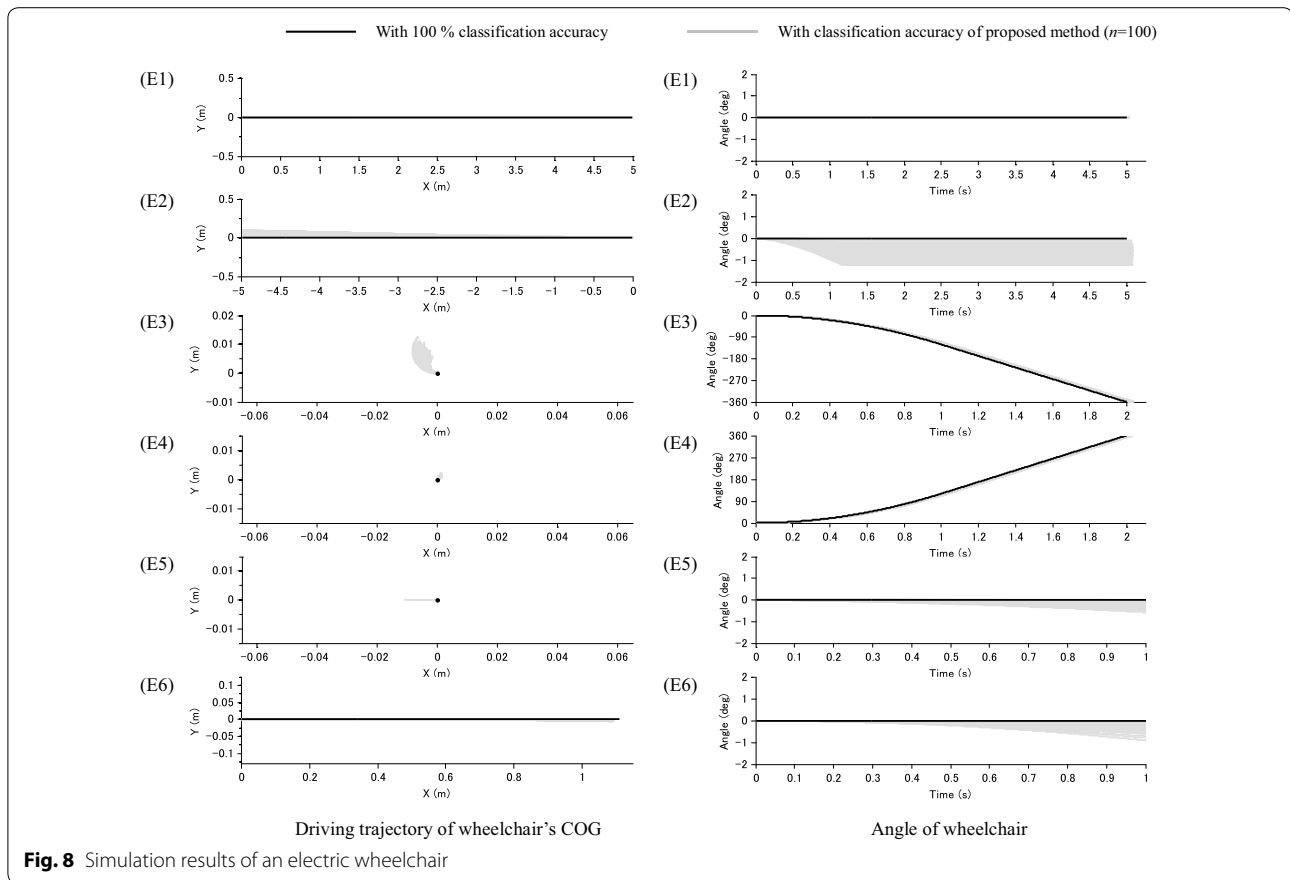
In test E1, both the angle θ and the driving trajectory in y-direction Y_G were, respectively, 0° and 0 mm. The difference between the maximum time required for 5 m moving of the wheelchair and the time for ideal trajectory was only 40.5 ms. In test E2, the maximum deviations of θ and Y_G for moving the wheelchair backward were, respectively, −1.2° and 99.4 mm. Because the distance of moving backward is about 1 m in daily life, the influence of these errors is believed to present no difficulty. These results suggest that the straight driving performance of an electric wheelchair using the proposed tongue interface is sufficient for practical use.

In test E3, maximum X_G and Y_G while turning the wheelchair 360° to the right were, respectively, −8.2 and 12.7 mm. In test E4, maximum X_G and Y_G for turning to the left were, respectively, 1.6 and 2.3 mm. The deviation of the center of gravity position is slight, which suggests good turning performance.

In test E5, the respective variations of the θ, X_G, and Y_G via swallowing saliva while stopping did not exceed −0.6°, −11.3 and 0.0 mm. In test E6, the respective maximum variations of θ and Y_G via swallowing saliva while moving straight with maximum velocity were −0.9° and −5.7 mm. Moreover, the driving velocity was reduced

from 4.0 km/h of maximum velocity to 3.8 km/h. From these results, it was confirmed that the influence of saliva swallowing on wheelchair operation can be inhibited at most to 11.3 mm.

Discussion

The proposed interface, which has simple appearance, can be attached easily and quickly even by a non-experienced caregiver, as depicted in Fig. 2. The prototype tongue interface was extremely lightweight: just 13.6 g. Because the silicon insulation comprises about 60.7 % of the whole sensor mass, further reduction of the sensor mass can be achieved using thinner silicon insulation sheets. Future studies will explore the optimal electrode unit size and the optimal number and location of electrodes for different categories of individuals. Further improvement might include the development of wireless communication between the sensor and the computer.

Tongue motion classification is based on analysis of the EMG activity of the suprahyoid muscles, which contribute not only voluntary tongue motions but also swallowing motion. Therefore, classification must be done of the large number of voluntary tongue motions that might be used for controlling an electric wheelchair and

Sasaki *et al. Robomech J* (2016) 3:9

Page 10 of 11

a PC. Such classification is also necessary for the detection of involuntary motions to inhibit malfunctions of such assistive devices. This study achieved classification accuracy of 95.1 ± 1.9 % using SVM classifier with features of time and frequency domains for five voluntary tongue motions and saliva swallowing. The voluntary tongue motions classified in this study were much more numerous than in our preliminary experiments. They are sufficiently numerous and diverse to control an electric wheelchair and a PC. Our future studies will emphasize further improvement of the tongue motion classification accuracy by optimizing the parameters of classification algorithms such as features and the SVM kernel. In addition, effects of the combination of classifiable tongue motions on the classification accuracy will be clarified.

Computer simulations of driving of an electric wheelchair were conducted to investigate the effectiveness of the proposed classification algorithm. The high performance of straight driving was achieved by finding a voluntary tongue motion that is not misclassified as a "right" or "left" tongue motion from the confusion matrix (Table 2) and by matching this motion with the "forward" command of an electric wheelchair (Table 4). Saliva swallowing during wheelchair driving reduces the velocity slightly. Therefore, this malfunction by saliva swallowing affects driving performance only slightly. However, saliva swallowing while the wheelchair is stopping made the wheelchair back up slightly. To provide a safety margin, some improvement of electric wheelchair control methods must be conducted as future work. As described above, electric wheelchair operation based on the proposed tongue interface has been demonstrated. Future studies will be conducted to evaluate the effects of yawning, talking, drinking, tongue motion speed, muscle fatigue, and head motion on the classification accuracy. The effects of small tongue positioning errors on the classification accuracy of the system will also be assessed. The usability of the tongue interface will be evaluated via new experiments using actual electric wheelchairs, PCs, and other assistive devices, and with testing of people with disabilities.

This study tested the design concept of the new interface through experimentation with five healthy adult male subjects. The results were sufficient to verify the viability of the concept, but a new detailed study will be necessary for evaluation of the developed interface when used by different categories of users. Such a new study will specifically examine the acceptance of the new interface by various users.

Conclusions

This study was conducted to develop and test a novel tongue interface based on the classification of tongue motions from surface EMG signals of the suprahyoid muscles detected at the underside of the jaw. The EMG signals of the suprahyoid muscles were measured via 22 active surface electrodes mounted on a special flexible boomerang-shaped base. The tongue motions were classified from RMS features and CC features of the EMG signals using an SVM classifier. Because the developed interface and this approach require no installation of any sensor into the mouth cavity, the system does not hinder the user's other activities such as eating, chewing, and talking. To verify the effectiveness of the tongue interface, an experiment was conducted with five healthy adult male subjects who had normal motor tongue functions. Results showed that the six tongue motions (i.e., five voluntary tongue motions and saliva swallowing) were classified with high accuracy of 95.1 ± 1.9 %. In addition, the potential of the proposed method was evaluated with a test whereby a computer simulation of an electric wheelchair was controlled using tongue commands. Results from the steering test demonstrated that the computer model was controlled precisely. The developed interface elaborates signals of sufficient number for the control of most assistive devices. This device is therefore useful for people with a high degree of movement paralysis. The tongue control interface can be simplified for use by patients with moderate movement disorders.

Authors' contributions
MS was the main person in charge of conception, design of experiments, acquisition of data, analysis and interpretation of data and drafting of the manuscript. KO took part in the experiments and calculations. DS took part in the conception, interpretation and revision of the manuscript. KK took part in the design of the experimental environment. AN and MY participated in the processing of the experimental data using the SVM. GO took part in the conception and interpretation. All authors read and approved the final manuscript.

Author details
[1] Graduate School of Engineering, Iwate University, Morioka, Iwate, Japan. [2] School of Science and Technology, Middlesex University, London, UK. [3] Pattern Art Laboratory Co., Ltd., Hanamaki, Iwate, Japan. [4] Department of Intelligent Systems Engineering, Ichinoseki National College of Technology, Ichinoseki, Iwate, Japan. [5] Graduate School of Information Science, Nara Institute of Science and Technology, Nara, Japan. [6] Department of Robotics Science and Engineering, Chubu University, Kasugai, Japan.

Competing interests
The authors declare that they have no competing interests.

References
1. Lau C, O'Leary S (1993) Comparison of computer interface devices for persons with severe physical disabilities. Am J Occup Ther 47:1022–1030

Sasaki *et al. Robomech J (2016) 3:9*

Page 11 of 11

2. Ghovanloo M (2007) Tongue operated assistive technologies. In: Proceedings the IEEE 29th engineering medicine biology conference, pp 4376–4379

3. Jouse3, Compusult limited. http://www.jouse.com

4. Clayton C, Platts RGS, Steinberg M, Hennequin JR (1992) Palatal tongue controller. J Microcomput Applicat 15:9–12

5. Terashima SG, Satoh E, Kotake K, Sasaki E, Uekii K, Sasaki S (2010) Development of a mouthpiece type remote controller for disabled persons. J Biomech Sci Eng 5(1):66–77

6. Kim D, Tyler ME, Beebe DJ (2005) Development of a tongue operated switch array as an alternative input device. Int J Hum Comput Interact 18:19–38

7. Wrench A, McIntosh AD, Watson C, Hardcastle WJ (1998) Optopalatograph: real-time feedback of tongue movement in 3D. In: Proceedings the fifth international conference on spoken language processing, pp 1867–1870

8. Saponas TS, Kelly D, Parviz BA, Tan DS (2009) Optically sensing tongue gestures for computer input. In: Proceedings 22nd annual ACM symposium on user interface software and technology, pp 177–180

9. Sonoda Y (1978) Observation of tongue movements employing a magnetometer sensor. IEEE Trans Magn 10:954–957

10. Huo X, Wang J, Ghovanloo M (2008) Introduction and preliminary evaluation of the tongue drive system: wireless tongue-operated assistive technology for people with little or no upper-limb function. J Rehabil Res Dev 45(6):921–930

11. Yousefi B, Huo X, Kim L, Veledar E, Ghovanloo M (2011) Quantitative and comparative assessment of learning in a tongue-operated computer input device: navigation tasks. IEEE Trans Inf Technol Biomed 15(5):747–757

12. Struijk LNSA (2006) An inductive tongue computer interface for control of computers and assistive devices. IEEE Trans Biomed Eng 53:2594–2597

13. Bentsen B, Gaihede M, Lontis R, Andreasen LNS (2014) Medical tongue piercing—development and evaluation of a surgical protocol and the perception of procedural discomfort of the participants. J Neuroeng Rehabil 11(1):1–11

14. Merletti R, Parker PA (eds) (2004) Electromyography: physiology, engineering, and non-invasive applications. Wiley-IEEE Press, New York

15. Hudgins B, Parker PA, Scott RN (1993) A new strategy for multifunction myoelectric control. IEEE Trans Biomed Eng 40(1):82–94

16. Kermani MZ, Wheeler BC, Badie K, Hashemi RM (1995) EMG feature evaluation for movement control of upper extremity prostheses. IEEE Trans Rehabil Eng 2(4):1267–1271

17. Englehart K, Hudgins B, Parker PA (2001) A wavelet-based continuous classification scheme for multifunction myoelectric control. IEEE Trans Biomed Eng 48(3):302–311

18. Ajjiboye AB, Weir RF (2005) A heuristic fuzzy logic approach to EMG pattern recognition for multifunctional prosthesis control. IEEE Trans Neural Syst Rehabil Eng 3(3):280–291

19. Chan ADC, Englehart KB (2005) Continuous myoelectric control for powered prostheses using hidden Markov models. IEEE Trans Biomed Eng 52(1):121–124

20. Chu JU, Moon I, Mun MS (2006) A real-time EMG pattern recognition system based on linear-nonlinear feature projection for a multifunction myoelectric hand. IEEE Trans Biomed Eng 53(11):2232–2239

21. Momen K, Krishnan S, Chau T (2007) Real-time classification of forearm electromyographic signals corresponding to user-selected intentional movements for multifunction prosthesis control. IEEE Trans Neural Syst Rehabil Eng 15(4):535–542

22. Naik GR, Kumar DK, Jayadeva (2010) Twin SVM for gesture classification using the surface electromyogram. IEEE Trans Inf Technol Biomed 14(2):301–308

23. Li G, Schultz AE, Luiken TA (2010) Quantifying pattern recognition-based myoelectric control of multifunctional transradial prostheses. IEEE Trans Neural Syst Rehabil Eng 18(2):185–192

24. Li Z, Wang B, Yang C, Xie Q, Su CY (2013) boosting-based EMG patterns classification scheme for robustness enhancement. IEEE J Biomed Health Inform 17(3):545–552

25. Zhang X, Zhou P (2012) High-density myoelectric pattern recognition toward improved stroke rehabilitation. IEEE Trans Biomed Eng 59(6):1649–1657

26. Stango A, Negro F, Farina D (2015) Spatial correlation of high density EMG signals provides features robust to electrode number and shift in pattern recognition for myocontrol. IEEE Trans Neural Syst Rehabil Eng 23(2):189–198

27. Sasaki M, Arakawa T, Nakayama A, Obinata G, Yamaguchi M (2011) Estimation of tongue movement based on suprahyoid muscle activity. In: Proceeding the 2011 IEEE international symposium on micro-nanomechatronics and human science, pp 433–438

28. Sasaki M, Onishi K, Arakawa T, Nakayama A, Stefanov D, Yamaguchi M (2013) Real-time estimation of tongue movement based on suprahyoid muscle activity. In: Proceeding the IEEE 35th engineering medicine biology conference, pp 4605–4608

29. Ide Y, Koide K (eds) (2004) Fundamental of functional anatomy for chairside evaluation of stomatognathic functions. Ishiyaku Publishers, Tokyo

30. Norton NS (2012) Netter's head and neck anatomy for dentistry, 2nd edn. Elsevier, London

31. Pittman LJ, Bailey EF (2009) Genioglossus and intrinsic electromyographic activities in impeded and unimpeded protrusion tasks. J Neurophysiol 101:276–282

32. Tsukada T, Taniguchi H, Ootaki S, Yamada Y, Inoue M (2009) Effects of food texture and head posture on oropharyngeal swallowing. J Appl Physiol 106(6):1848–1857

33. Coriolano MG, Belo LR, Carneiro D, Asano G, Oliveira AL, Silva DM, Lins G (2012) Swallowing in patients with parkinson's disease: a surface electromyography study. Dysphagia 27:550–555

34. Balata PMM, Silva HJ, Nascimento Moraes KJR, Pernambuco LA, Freitas MCR, Lima LM, Braga RS, Souza SR, Moraes SRA (2012) Incomplete swallowing and retracted tongue maneuvers for electromyographic signal normalization of the extrinsic muscles of the larynx. J Voice 26(6):813.e1–813.e7

35. Yoon WL, Khoo JKP, Liow SJR (2014) Chin tuck against resistance (CTAR): new method for enhancing suprahyoid nuscle activity using a shakertype exercise. Dysphagia 29:243–248

36. Carlo JDL (2002) Surface electromyography: detection and recording. DelSys Incorporated, Boston

37. Englehart K, Hudgins B (2003) A robust, real-time control scheme for multifunction myoelectric control. IEEE Trans Biomed Eng 50(7):848–854

38. Zecca M, Micera S, Carrozza MC, Dario P (2002) Control of multifunctional prosthetic hands by processing the electromyographic signal. Crit Rev in Biomed Eng 30(4–6):459–485

39. Yoshikawa M, Mikawa M, Tanaka K (2007) A myoelectric interface for robotic hand control using support vector machine. In: Proceedings the 2007 IEEE/RSJ international conference on intelligent robots and systems, pp 2723–2728

40. Kang WJ, Shiu JR, Cheng CK, Lai JS, Tsao HW, Kuo TS (1995) The application of cepstral coefficients and maximum likelihood method in EMG pattern recognition. IEEE Trans Biomed Eng 42(8):777–785

41. Lee SP, Kim LS, Park SH (1996) An enhanced feature extraction algorithm for EMG pattern classification. IEEE Trans Rehab Eng 4(4):439–443

42. Cortes C, Vapnik (1995) Support-vector networks. Mach Learn 20:273–297

43. Hsu CW, Lin CJ (2002) A comparison of methods for multi-class support vector machines. IEEE Trans Neural Netw 13(2):415–425

44. Chang CC and Lin CJ (2013) LIBSVM—a library for support vector machines. http://www.csie.ntu.edu.tw/~cjlin/libsvm

45. Micera S, Vannozzi G, Sabatini AM, Dario P (2001) Improving detection of muscle activation intervals. IEEE Eng Med Biol Mag 20(6):38–46