Tongue Motor Training Support System

Makoto Sasaki, *Member, IEEE*, Kohei Onishi, Atsushi Nakayama, *Member, IEEE*, Katsuhiro Kamata, Dimitar Stefanov, *Senior Member, IEEE*, and Masaki Yamaguchi, *Senior Member, IEEE*

***Abstract*— In this paper, we introduce a new tongue-training system that can be used for improvement of the tongue's range of motion and muscle strength after dysphagia. The training process is organized in game-like manner. Initially, we analyzed surface electromyography (EMG) signals of the suprahyoid muscles of five subjects during tongue-training motions. This test revealed that four types tongue training motions and a swallowing motion could be classified with 93.5% accuracy. Recognized EMG signals during tongue motions were designed to allow control of a mouse cursor via intentional tongue motions. Results demonstrated that simple PC games could be played by tongue motions, achieving in this way efficient, enjoyable and pleasant tongue training. Using the proposed method, dysphagia patients can choose games that suit their preferences and/or state of mind. It is expected that the proposed system will be an efficient tool for long-term tongue motor training and maintaining patients' motivation.**

I. Introduction

Statistics show that in the 2011 fiscal year, pneumonia was the third largest cause of death in Japan. More than 90% of those who died of pneumonia in that period were elderly people. About half of them had developed aspiration pneumonia caused by a swallowing disorder (dysphagia). Around 30% of the people over 75 years of age had dysphagia. The number of dysphagia sufferers in Japan is estimated as 700–800 thousand. The "eating" function has significant role in daily life: people get nutrients necessary to support life, they experience taste and enjoy conversations while eating. Dysphagia affects these fundamental pleasures and degrades significantly patients' quality of life (QOL). Therefore, dysphagia is regarded as a profound social problem in Japan, especially affecting the aging population.

Rehabilitation of dysphagia patients is based on two main approaches: direct training by using food and indirect training without using food. One type of indirect training is tongue motor training, which is aimed to improve the range of tongue motion, muscle force, coordination, and food transport function [1-5]. This training approach for oral function improvement is based on repetitive intentional tongue motions such as moving the tongue to the right and left, and up and down while keeping the mouth open during the exercise. Tongue motion promotes saliva production and lubricates mastication and swallowing during eating. For some dysphagia patients tongue motion training is suggested as a warming-up exercise before eating. However, tongue motor training is a monotonous process and maintaining patient's motivation to such training is difficult. In this study, we present a novel system for tongue motor training that is designed in a game-like manner to maintain patient's interest during training.

M. Sasaki is with Graduate School of Engineering, Iwate University, Morioka, Iwate 020-8551, Japan (corresponding author to provide phone and fax: +81-19-621-6385; e-mail: makotosa@iwate-u.ac.jp).

K. Onishi, and M. Yamaguchi are with Graduate School of Engineering, Iwate University.

K. Kamata is with Pattern Art Laboratory Co., Ltd., Hanamaki, Iwate 025-0054, Japan.

A. Nakayama is with Department of Intelligent Systems Engineering, Ichinoseki National College of Technology, Ichinoseki, Iwate 021-8551, Japan.

D. Stefanov is with the Health Design Technology Institute, Coventry University, Coventry, UK.

II. Tongue Motor Training Support System

Figure 1 presents an outline of the proposed tongue motor training support system. Surface electromyography (EMG) signals of the suprahyoid muscles are analyzed and converted into relevant signals for cursor movement and controlling the therapy computer game. Below we explain our approach for classification of the tongue-training motions and game control organization.

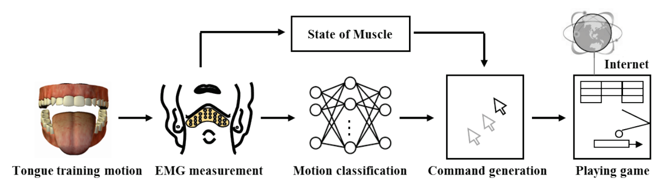


Figure 1. Tongue motor training support system.

A. Classification of tongue training motions

Tongue motor training includes intentional tongue motions such as moving the tongue tip to the right and left and up and down, as illustrated in Figure 2. In this study, tongue training motions and muscle active state during the training session are monitored via analysis of the surface Electromyography (EMG) signals of the suprahyoid muscles (digastric muscle, stylohyoid muscle, mylohyoid muscle, and geniohyoid muscle) captured at the underside of the jaw. In general, tongue motion is developed by the intrinsic muscles of the tongue that control the tongue shape and direction of the tongue tip, and the extrinsic muscles that control the tongue protrusion and retraction. For realization of tongue motion, suprahyoid muscles maintain the hyoid supporting the root of the tongue in appropriate positions according to the movements of the intrinsic muscle of the tongue and the extrinsic muscles of the tongue. Therefore, an efficient

Figure 2. Tongue training motions and swallowing motion.

978-1-4244-7929-0/14/$26.00 ©2014 IEEE

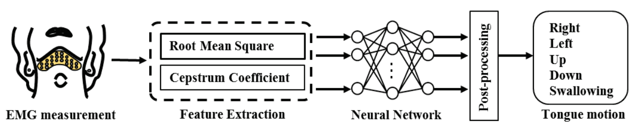


Figure 3. Classification method of tongue-training motions.

classification of the tongue motions can be achieved by careful consideration the coordinated motion of the suprahyoid muscles [6, 7].

Figure 3 shows the procedures for classification of the tongue training motions. In this study we used surface electrodes that include a 22-channel active electrode mount attached at the underside of the jaw, a 1-channel active electrode attached at the right earlobe, and a reference electrode attached at the left earlobe. The mount contained a matrix of 22 active silver electrodes (ϕ 2 mm) placed at intervals of 12.5 mm. Potential differences between the electrode attached to the right earlobe and the electrode at the underside of the jaw were derived by bipolar leading method where the left earlobe was regarded as the reference potential. Then, features were extracted from the EMG signals in the 128 ms frame. The frame was shifted for 16 ms. For extraction of the signal features, we need the root mean square (RMS) and cepstrum coefficients (CC). These were calculated by using the approach presented in [8-10].

The *RMS* corresponds to the magnitude of the EMG signals and can be expressed by their given with the following equation:

$$RMS_l(p) = \sqrt{\frac{1}{N}\sum_{n=0}^{N-1} EMG_l(n)^2} \tag{1}$$

Therein, $EMG_l(n)$ signifies the EMG signals, p stands for frame number, N denotes the number of samples in 1 frame, and l represents the number of channels of EMG signal.

To determine the CC coefficients, we refer to the Fourier transform of $EMG_l(n)$ that can be expressed as:

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

then, the cepstral coefficients can be calculated from the following equation:

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

We use cepstral analysis to separate the spectral envelope from the fine structure. Features of the spectral envelope appear in low-order coefficients and features of the fine structure appear in high-order coefficients. To define the CC features, we use the low-order coefficients up to the fifth coefficient.

Next, we create learning data by defining the motion class as a feature vector composed by the RMS features and the CC features. The motion label is defined by threshold processing of the summation of the RMS features for all EMG channels. The feature vector and the motion label are used as a teacher signal for training of the weights of the connections among all neurons in the neural network. During the procedure for motion classification, threshold processing and majority processing are applied for smoothing the output signals of the neural network output layer.

B. *Game operation using tongue-training motion*

As commented above, classified tongue-training motions are used for control of the PC cursor. For the experiments, we used an input device designed for people with disabilities (Rakurakumouse II wireless; KoKoTo STEP Organization). The Rakurakumouse has eight buttons allowing movement of the cursor in eight directions and buttons for right click, left click, drag, and scroll. The patient can control the interface functions by using one finger. For the tests, we connected photocoupler circuits in parallel to the contacts of the original input device to enable operation of the mouse via digital PC signals. This way, the Rakurakumouse was made to respond to the tongue motions. For testing different control strategies during experiments, some of the relations between the control signals and the cursor movement parameters can be changed easily by simple changes in the computer program. For example, in some tests the mouse pointer could be set to move in inverse direction of the tongue tip movement. The mouse pointer speed can be changed easily to adapt the user's fatigue state.

The games used for tongue motor training can vary depending on one's preference and state of mind. Any game which can be played by mouse might be adapted easily to be used for tongue training. Internet gives a choice of suitable games that can be used. The games can be changed easily to keep the patient's interest in the game and motivation to the rehabilitation process.

III. Experiments of Tongue-Training Motion Classification

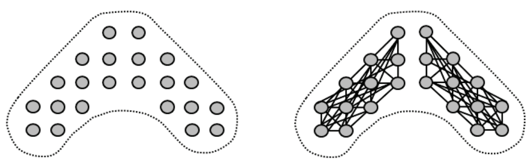
A. *Experimental conditions*

Five healthy male adults with normal tongue function (age 23 ± 2.3 years old, height 170.1 ± 6.6 cm, body weight 60.2 ± 10.3 kg, mean ± SD) participated in the tests. In the experiments, participants were asked to perform with open mouth a set of four training motions that included moving the tongue to the right and left, up and down, as well as a saliva swallowing motion that was performed with closed mouth (Fig. 2). Before each session, the end positions of the tongue for each participant in each motion direction were checked and marked. During the completion of the tongue-training motions it was monitored visually if the end tongue positions are reached. Times for tongue motions were set to about 2 s. A resting time of 2 s was provided between the series. The number of times of measurement was 5 motions × 14 sets. The EMG signals was amplified 2,052 times using an electromyograph (Pattern Art Laboratory Co., Ltd.) with a 14 Hz high-pass filter and a 440 Hz low-pass filter and measured at sampling frequency of 2,000 Hz using a multi-function Data AcQuisition unit (USB-6218; National Instruments Corp.).

B. *Method of learning and assessment*

The EMG signal of each channel is built from to the potential difference between the electrode attached to the right earlobe and the electrode at the underside of the jaw, while the left earlobe is regarded as the reference potential. We conducted preliminary tests for tongue motion classification in the oral cavity by calculating the EMG signal from nine surface electrodes attached at the underside of the jaw at intervals of 20 mm using these combinations [6,7]. For the present experiments, we used a multi-channel electrode mount containing 22 electrodes. The distance between the electrodes was 12.5 mm. In this work we also explored the signal classification accuracy and feature extraction by comparing the classification results for the following two approaches:

- Approach 1: The features are extracted from 22-channel EMG signals (Fig. 4(a)).
- Approach 2: 22 electrodes are divided into two groups (right and left). All potential differences between electrodes are calculated for each group. Then the features are extracted from the EMG signals of 110 channels (${}_{11}C_2 \times 2$) (Fig. 4(b)).



(a) 22-channel EMG (Approach 1) (b) 110-channel EMG (Approach 2)
Figure 4. EMG signal used for feature extraction.

The first 4 sets measurement data were used for learning of the neural network. The remaining 10 sets were used for assessment of the identification accuracy. The neural network was constructed by using Matlab (Neural Network Toolbox; The MathWorks, Inc.). The number of units of the intermediate layer is 50. The learning termination conditions are set to achieve a square error of less than 0.05. Back-propagation was used for learning of the neural networks.

The classification accuracy was assessed by using the following equation:

$$\text{Classification rate} = \frac{\text{number of correct feature vector}}{\text{total number of feature vector}} \times 100 \tag{4}$$

C. Results

An example of the classification results of the tongue-training motions can be seen in Figure 5. These motions contain common phases such as opening the mouth and thrusting the tongue and retrieving the tongue back into the mouth. Therefore, in the common phase, tongue-training motion was sometimes classified incorrectly.

The results from the calculation of the classification rate under approach 1 and 2 are presented in Tables 1 and 2. For approach 1, where the features are extracted from 22-channel EMG signal, the classification rate for five motions exceeded 90% for all participants and the average rate was 93.5%. For approach 2, where the features were extracted from a 110-channel EMG signal, only one subject achieved classification rate above 90% for all five motions and the

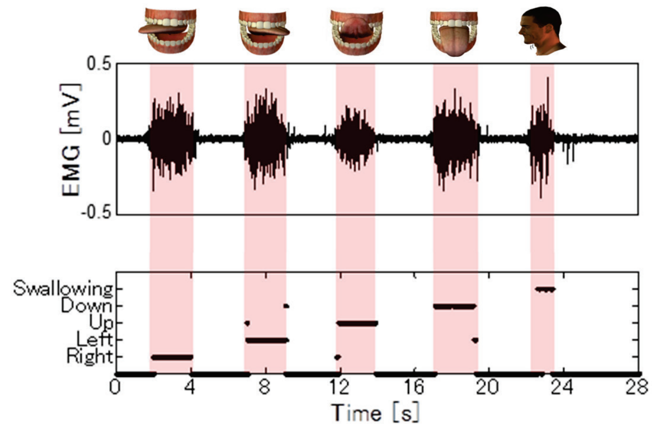


Figure 5. Example of results of tongue-training motion classification.

Table 1 Classification rate for 22-channel EMG signal (Approach 1)

Subject	Right	Left	Up	Down	Swallowing	Total
A	92.5	92.9	99.5	96.2	100.0	96.2
B	87.5	87.1	94.9	80.8	100.0	90.1
C	91.1	88.2	92.7	91.4	88.2	90.3
D	96.7	95.4	93.4	85.9	97.4	93.7
E	99.1	94.9	98.2	93.5	99.0	96.9
Mean	93.4	91.7	95.7	89.5	96.9	93.5

Table 2 Identification rate for 110-channel EMG signal (Approach 2)

Subject	Right	Left	Up	Down	Swallowing	Total
A	92.3	96.1	82.7	94.4	100.0	93.1
B	73.1	92.2	76.0	91.1	90.7	84.6
C	87.8	68.3	84.6	77.0	91.2	81.8
D	91.9	84.7	92.7	77.4	98.6	89.0
E	96.3	88.6	87.8	92.8	82.3	89.5
Mean	88.3	86.0	84.8	86.5	92.5	87.6

average result of all subjects was 87.6%. Approach 2, where the potential differences between the electrodes are combined, allows obtaining much information with a smaller number of electrodes placed at the underside of the jaw. A shortcoming of the same method is that unwanted signal components are generated when many electrodes are used. Approach 2 also requires an increased amount of calculations. As a difference, approach 1 requires fewer calculations and yielded a higher classification rate in this experiment. Future studies will be conducted for optimising the configuration of electrodes and their number.

IV. Game Operation Experiments Using Tongue-Training Motions

Experimental results presented in the previous section showed that all five motions (tongue movement in four positions and swallowing motion) could be classified from the EMG signals of the suprahyoid muscles with accuracy of 93.5%. In the following paragraph we discuss the game experiments for tongue-training motions.

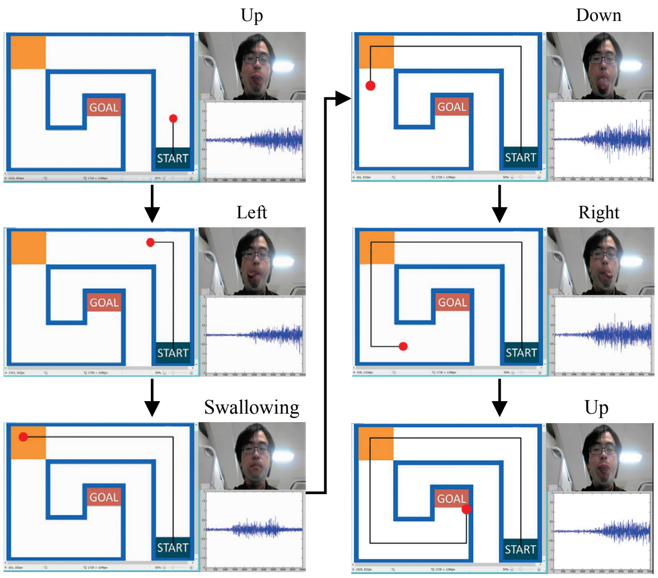


Figure 6. Aspect of game operation by a tongue-training motion.

The directions of the tongue-training motions were aligned to the movements of the mouse pointer. This way, the movement of the tongue to the right or left and up and down resulted in cursor movement is the same directions. In order to avoid eventual false results due to unconscious swallowing motions that occur every 1–3 min the computer program was designed to block all mouse signals at the moment of swallowing and this way, to eliminate eventual false cursor movements due to swallowing. Digital signals for the input support device for people with disabilities were generated by a multi-function Data AcQuisition unit.

The EMG signals in response to these tongue motions were used to train the neural network. Next, we used a simple game as shown in Figure 6, to ascertain whether the subjects were able to control the cursor from the start to the goal by tongue training motions. A notebook PC was placed in front of the subject to display the trajectories of the mouse pointer and EMG signals. On a separate window on the same screen we presented an image from a Web camera directed toward the user's face to provide feedback to the user on the performed tongue movements (mirror mode). Instruction to the user for the times for saliva swallowing was given in an area shown in orange.

Screenshots from the mouse operation experiments are presented in Figure 6. A red circle shows the position of the mouse pointer. The black line shows the trajectory of the mouse pointer. The experiments demonstrated that users were able to navigate accurately the cursor from the start to the goal without protruding from the frame. In the areas colored in orange, the mouse pointer did not move in an unintended direction during saliva swallowing.

The initial results show that tongue training by playing games is accepted well by users. The approach offers great flexibility because dysphagia patients can find various games on the Internet that suit their tastes or mood to perform tongue motor training. The unique features of this method are that patients maintain their motivation for training.

V. CONCLUSIONS

In this study, we developed a novel training support system for improving the tongue's range of motion and muscle strength after dysphagia. Tongue training motions were classified from the surface EMG signals of the suprahyoid muscles. Results reveal that all four tongue-training motions and the swallowing motion were classified with 93.5% accuracy. The approach was tested with a simple PC game. Results demonstrated that simple PC games could be played by tongue-training motions, achieving in this way efficient, enjoyable and pleasant tongue training. Using the proposed method, dysphagia patients can choose games that suit their preferences and/or state of mind. It is expected that the proposed system will be an efficient tool for long-term tongue motor training and maintaining patients' motivation. Training efficiency of the proposed system will be evaluated by further analysis of the evaluation indexes such as tongue pressure test, repetition saliva swallowing test, and the oral diadochokinesis test.

ACKNOWLEDGMENT

This study was supported in part by the Japan Society of Promotion of Science, Japan (Grant-in-Aid for Scientific Research (C) 24500637).

REFERENCES

[1] J.A. Logemann,"*Evaluation and treatment of swallowing disorders. 2nd ed.*," Pro-Ed; Austin, TX, 1998.

[2] T.M. Waters, J.A. Logemann, B.R. Pauloski, A.W. Rademaker, C.L. Lazarus, L.A. Newman, S.K. Hamner, "Beyond efficacy and effectiveness: conducting economic analyses during clinical trials," *Dysphagia*, vol.19, pp.109-119, 2004.

[3] B.R. Pauloski," Pauloski rehabilitation of dysphagia following head and neck cancer, *Phys Med Rehabil Clin N Am*, vol.19, no.4, pp.889-928, 2008.

[4] T. Ooka, T. Haino, S. Hironaka, Y. Mukai, "The Effect of daily oral function training in the Elderly," *J. Den. Health*, vol.58, no.2, pp.88-94, 2008 .

[5] M. Yukihiro, I. Toshimitsu,"*Manual of swallowing therapy on adults*, " Ishiyaku Publishers, Inc., 2002.

[6] M. Sasaki, T. Arakawa, A. Nakayama, G. Obinata, M. Yamaguchi, "Estimation of tongue movement based on suprahyoid muscle activity," in *Proc. of the 2011 IEEE International Symposium on Micro-NanoMechatronics and Human Science*, pp.433-438, 2011.

[7] M. Sasaki, K. Onishi, T. Arakawa, A. Nakayama, D. Stefanov, M. Yamaguchi, "Real-time estimation of tongue movement based on suprahyoid muscle activity," in *Proc. of the 35th Annual International Conference of the IEEE EMBS*, pp.4605-4608, 2013.

[8] M.A. Oskoei, H. Hu, "Myoelectric control systems—A survey," *Biomed. Signal Process. Control*, vol.2, pp.275-294, 2007.

[9] M. Zecca, S. Micera, M.C. Carrozza, P. Dario, "Control of multifunctional prosthetic hands by processing the electromyographic signal," *Crit. Rev. in Biomed. Eng.*, vol.30, no.4-6, pp.459-485, 2002.

[10] M. Yoshikawa, M. Mikawa, K. Tanaka, "A myoelectric interface for robotic hand control using support vector machine," in *Proc. of the 2007 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp.2723-2728, 2007.