35th Annual International Conference of the IEEE EMBS
Osaka, Japan, 3 - 7 July, 2013

Using Speech for Mode Selection in Control of Multifunctional Myoelectric Prostheses

Peng Fang, Zheng Wei, Yanjuan Geng, Fuan Yao, and Guanglin Li, *Senior Member, IEEE*

Abstract— **Electromyogram (EMG) recorded from residual muscles of limbs is considered as suitable control information for motorized prostheses. However, in case of high-level amputations, the residual muscles are usually limited, which may not provide enough EMG for flexible control of myoelectric prostheses with multiple degrees of freedom of movements. Here, we proposed a control strategy, where the speech signals were used as additional information and combined with the EMG signals to realize more flexible control of multifunctional prostheses. By replacing the traditional "sequential mode-switching (joint-switching)", the speech signals were used to select a mode (joint) of the prosthetic arm, and then the EMG signals were applied to determine a motion class involved in the selected joint and to execute the motion. Preliminary results from three able-bodied subjects and one transhumeral amputee demonstrated the proposed strategy could achieve a high mode-selection rate and enhance the operation efficiency, suggesting the strategy may improve the control performance of commercial myoelectric prostheses.**

I. INTRODUCTION

Multifunctional prostheses are very useful aids for limb amputees, and a proper control strategy is the decisive factor for their acceptability and practicality. Electromyographic (EMG) signals have been suggested for prosthesis control [1–8], and many myoelectric prostheses have been already available on the market. Traditionally, EMG signals from a pair of residual muscles (agonistic and antagonistic muscle) are used to control one degree of freedom (DOF) of movements. However, the EMG signal source is usually limited, especially for the high-level limb amputees, where few residual muscles are retained but more EMG sources are required for the recovery of their lost limb functions [9–10]. Conventionally, in order to control multiple DOFs with one pair of residual muscles, a so-called "sequential mode-switching (joint-switching)" is utilized, where the switching between different joints of a multifunctional prosthesis is realized by either a simultaneous co-contraction of a muscle pair or an external switch pad. In this way, in order to execute a motion, the corresponding mode should be switched to in advance. As an example, in a 3-DOF-prosthesis with three joints of "hand", "wrist", and "elbow" for transhumeral amputees, the switching sequence is "hand-wrist-elbow-hand-…" and is performed with the co-contraction of a muscle pair, such as residual biceps and triceps. If the present mode is "hand" and the user wants to do an elbow motion, he/she has to co-contract the muscle pair to switch the mode from "hand" to "wrist", then co-contract again to switch from "wrist" to "elbow", and contracts either biceps or triceps to actuate an elbow motion such as flexion or extension. Therefore, switching to different modes is slow and makes the prosthesis control cumbersome. As a result, less than half of the myoelectric-prosthesis owners often use their prostheses due to the long training period, awkward motion, and heavy body burden [11].

A control approach based on pattern recognition of EMG signals may yield a significant improvement over the conventional myoelectric control strategy. The previous studies have shown that this approach has the potential to allow users to naturally operate their myoelectric prostheses with multiple DOFs [1–3, 12–13]. However, it is also limited with the lack of enough EMG sources after amputations. In order to rebuild the lost information sources, a promising nerve-machine interface called Targeted Muscle Reinnervation (TMR) has been proposed and developed recently [13–16], where the residual body nerves are connected to some specific target muscles through surgeries. But the inconvenience of the second surgery and relatively high cost may prevent the further application of the TMR method.

To overcome the difficulty of the awkward mode switching in current commercial myoelectric prostheses, a more convenient and easy-to-realize way is required, which may need some additional control signals. One of the available candidates may be the human speech that has been widely utilized as a kind of simple control signal. The speech signals can be non-invasively acquired and the present speech recognition technique is quite developed [17]. In this paper, we proposed a new control strategy, in which the speech signals would be used as additional information and combined with the EMG signals to realize more flexible control of multifunctional prostheses with multiple DOFs. By replacing the traditional "sequential mode-switching", the speech signals would be used to select a mode of the prosthetic arm, and then the EMG signals would be applied to determine a motion class involved in the selected mode and to execute the motion. The feasibility and performance of the proposed method would be investigated in the study.

*This work was supported in part by the National Natural Science Foundation of China under Grant (#61203209), the Shenzhen Governmental Basic Research Grand (#JC201005270295A), the Shenzhen Public Platform for Biomedical Electronics and Health Informatics, and the Guangdong Innovation Research Team Fund for Low-cost Healthcare Technologies.

P. Fang, Z. Wei, Y. Geng, and G. Li are with the Key Laboratory of Health Informatics of Chinese Academy of Sciences (CAS), Shenzhen Institutes of Advanced Technology, CAS, Shenzhen, Guangdong 518055 China (phone: 86-755-86392219; fax: 86-755-86392299; e-mail: peng.fang@siat.ac.cn).

F. Yao is with the School of Control Science and Engineering, Shandong University, Jinan, Shandong 250061 China.

II. Methods

A. Control Strategy

A 3-DOF-prosthesis (*Shanghai Kesheng MH32, China*) was used. It had three modes of "hand", "wrist", and "elbow", and each mode had two motion classes as "hand-closing/opening", "wrist-pronation/supination", and "elbow-flexion/extension". A control strategy based on the combination of speech and EMG was proposed: The speech signals were used to select a mode (i.e. hand, wrist, or elbow), and then the EMG signals from a pair of muscles were used to determine a motion class involved in the selected mode (i.e. closing or opening for hand; pronation or supination for wrist; flexion or extension for elbow) and execute the motion, as shown in Fig. 1. Here, a pair of muscles (bicep and tricep in this work) with an EMG-electrode on each was used as the EMG source, and each muscle was corresponding to a motion class in a selected mode (i.e. bicep for hand-closing and tricep for hand-opening in the hand mode; bicep for wrist-pronation and tricep for wrist-supination in the wrist mode; bicep for elbow-flexion and tricep for elbow-extension in the elbow mode).

B. Subjects

In the pilot study, four subjects with full language competence were recruited, including three able-bodied subjects (two male and one female, marked as A1, A2, and A3) and one transhumeral amputee (male, marked as B), as summarized in Table I. EMG signals were recorded from the full bicep and tricep of the able-bodied subjects and from the residual bicep and tricep of the transhumeral amputee, as shown in Fig. 2. The protocol of this study was approved by the Institutional Review Board of the Shenzhen Institutes of Advanced Technology, China. All subjects gave the written informed consent and provided the permission for publication of photographs with a scientific and educational purpose.

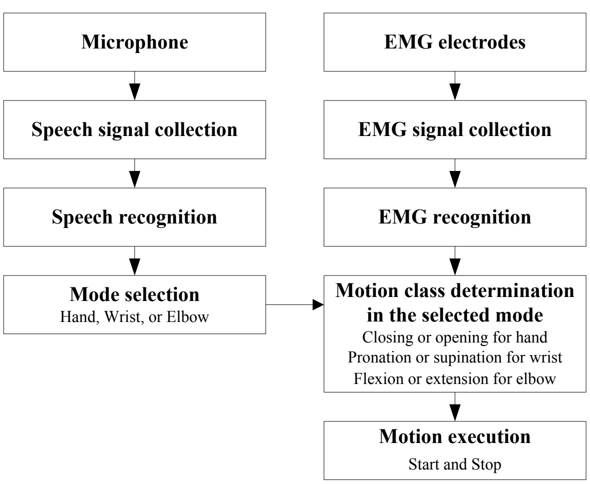

Figure 1. Control strategy for multifunctional prostheses based on the combination of speech and EMG signals, where speech was used for mode selection and EMG was used for motion class determination and motion execution.

TABLE I. Demographic Information of Subjects

Subject	Gender/Age	Body situation	Test side
A1	Male/30	Able-bodied	Left
A2	Male/24	Able-bodied	Right
A3	Female/28	Able-bodied	Right
B	Male/33	Amputated, Right	Right

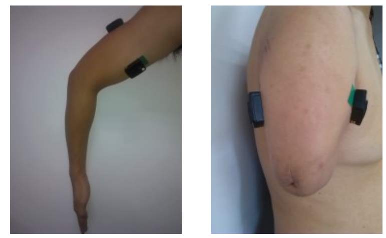

Figure 2. Body situation and EMG recording position for an able-bodied subject (left) and the transhumeral amputee (right).

C. Data Acquisition

For mode selection, simple keywords representing different modes should be recognized. An individual speech template was created for each user instead of a preset standard speech bank, and therefore the system could be applied to any users no matter which language, dialect, or accent he/she used. In practice, any words can be used as the control orders depending on the users' preference. In this study, the words "hand", "wrist", "elbow" in mandarin Chinese were specified for three joint modes. Speech signals were collected in office environment (background noise of around 55 dB) with a commercial omnidirectional electret condenser microphone (sensitivity of –44±3 dB), and processed through amplification, second-order butterworth band-pass filter (passing band of 370 Hz to 3.4 kHz), and acquired with a self-designed acquisition system (sampling rate of 8000 Hz, every 1000 samples were considered as a sample frame). A template was created by pronouncing each keyword five times for each subject. For the recognition of keywords, the dynamic time warping (DTW) algorithm [18] was applied. The mel-frequency cepstrum coefficients (MFCC) based on auditory model was used to extract speech characteristic parameters for a good recognition precision in case channel noise and spectrum distortion existed. In order to reflect the dynamic behavior of speech signals, the MFCC, its first-order difference, and its second-order difference were combined into one vector as the characteristic parameter for the speech signal. The recognition was triggered by signal amplitude, and only the signals that were greater than a given threshold would be compared with the template with the DTW algorithm, and the most matching one was considered as the recognition result. Thereafter, the required mode was selected by the system according to the recognized speech signals.

A commercial EMG collection system (*Delsys Trigno Wireless, USA*) was applied for the EMG recording from a pair of muscles. The amplitude modulation method based on fixed dual-thresholds was used for EMG signal decoding and control. EMG signals from each muscle were recorded (sampling rate of 8000 Hz, every 1000 samples were considered as a sample frame) through the corresponding bipolar electrode and transmitted wirelessly to the processing terminal. When the amplitudes of EMG recordings were greater than a magnitude threshold in a frame, the corresponding samples were counted. Only when the counting number of samples was larger than a given counting threshold, this frame was marked as valid. EMG signals with continuous three valid frames were applied for motion class determination and motion execution.

D. Quantification of Control Performance

Besides the proposed strategy (called Strategy 1 in the following), the traditional "sequential mode-switching" method (called Strategy 0 in the following) was also applied in the study and discussed together with Strategy 1 as a comparison. Three performance measures were proposed and calculated to quantify and compare different strategies:

(1) Mode-selection rate (in Strategy 1), defined as the ratio of correct mode selection over the total selection trials controlled by the speech, i.e. the recognition rate of speech signals;

(2) Mode-switching rate (in Strategy 0), defined as the ratio of correct mode switching over the total switching trials controlled by the co-contraction of a muscle pair, i.e. the recognition rate of EMG signals from muscle pair co-contraction;

(3) Action-execution time (in both strategies), defined as the time needed to complete a whole action without any misoperation. The action might contain a series of motions, and the procedure to finish a motion included the mode selection (in Strategy 1) or mode switching (in Strategy 0), and the motion execution.

All the tests were repeated at least three times and the results (recognition rate and time) were calculated as the average value of the repeated measurements.

III. Results

A. Mode Selection/Switching

As shown in Table II, for Strategy 1, a mode-selection rate of 100% was achieved for all the able-bodied subjects and the transhumeral amputee, indicating the mode could be successfully selected by speech with the present speech recognition system. There was no difference between subjects on mode selection since all the subjects had full language competence. For Strategy 0, however, the mode switching by the co-contraction of bicep and tricep was relatively hard or even impossible. Only the able-bodied subject A1 could conduct the co-contraction completely without any error, since he often took body exercise and was relatively muscular. The able-bodied subject A2 could fulfill the co-contraction partly, and both the able-bodied subject A3 and the transhumeral amputee B could not conduct the co-contraction at all, due to the weakness of their full/residual muscles. Especially for the amputee, his limb function was almost lost after the amputation, and therefore the residual muscles were lack of exercise and became atrophic.

TABLE II. MODE-SELECTION/SWITCHING RATE

Subject	Mode-selection rate in Strategy 1	Mode-switching rate in Strategy 0
A1	100%	100%
A2	100%	86.7%
A3	100%	Not available
B	100%	Not available

B. Functional Action Execution

A functional action of "water pouring" was specified, which included a series of motions: "hand-closing" to hold a cup with water inside, "elbow-flexion" to lift up the cup, "wrist-pronation" to pour the water out, and then "wrist-supination", "elbow-extension", and "hand-opening" to return. The execution time to finish the action continuously without any misoperation was measured, as shown in Table III. It can be seen for Strategy 1 all subjects could finish the required action with execution time of 19.6, 16.0, and 19.2 s for the able-bodied subjects and of 21.9 s for the transhumeral amputee, and the difference between the able-bodied subjects and the amputee was not very obvious. For Strategy 0, on the contrary, a longer execution time of 25.8 and 42.3 s was measured for the able-bodied subjects A1 and A2, respectively. Especially for the subject A2, there was a large increase of execution time compared with Strategy 1. Since the motion switching by co-contraction of muscle pair was not possible for the able-bodied subject A3 and the transhumeral amputee B, there was no measured execution time for them.

IV. Discussion

Compared with the traditional control strategy where the "sequential mode-switching (joint-switching)" is applied, the most difference of the proposed strategy is that the mode was chosen directly by the amputee's speech signals, which enables more easy and flexible control of multifunctional prostheses with multiple DOFs. In Strategy 0 the modes are switched by the co-contraction of a pair of muscles of limbs, however, this simultaneous co-contraction would be difficult to conduct, not only for the amputee but also for the able-bodied, as demonstrated by the experimental results. Only the able-bodied subject with relatively strong muscles could fulfill the co-contraction without much difficulty, and the transhumeral amputee could not do the co-contraction at all due to the weakness of the residual muscles. Besides, the sequential mode-switching is inefficient, and the mode has to be switched frequently even a simple action is desired. It was

TABLE III. ACTION-EXECUTION TIME

Subject	Action-execution time (s)	
	Strategy 1	***Strategy 0***
A1	19.6	25.8
A2	16.0	42.3
A3	19.2	Not available
B	21.9	Not available

found that most of the extra time used in Strategy 0 was consumed for the awkward sequential switching, especially when twice of continuous switching were required, e.g. from “hand” to “elbow” through “wrist”. What is more, the frequent switching may make the muscles tired, and a short rest is therefore unavoidable, which further lowers the efficiency. In Strategy 1, the direct mode-selection by speech is much easier to operate, as confirmed by the high mode-selection rate for all the subjects. In addition, the use of speech requires no long-term training or adaption process. The pronunciation and recognition of a simple keyword is a fast process compared with the co-contraction of muscle pair and the sequential switching, and therefore decreases the action-execution time. In the proposed strategy, EMG is still used as the executive signals for motions. This makes the prosthesis control safer, because a wrong speech order or environmental noise will not execute any motion as long as the prosthesis user does not contract the corresponding muscle. Note that flexible control of prostheses with more DOFs may also be realized with the proposed strategy, since the number of modes represented by speech is unlimited due to the unlimited speech signals.

Considering normal speech signals conducted by air can be easily interfered by environmental noises and it is inconvenient and embarrassing if prosthesis users “talk” to their own “prostheses”, the non-audible murmur (NAM) will be used as control signal instead of the speech signals of current study for possible improvement. NAM is conducted through human body such as muscle tissues and is less interfered by surroundings. Besides, they can not be heard by people around and therefore keep users’ privacy. Further investigations will be conducted on the use of NAM for prosthesis control.

V. Conclusion

A control strategy for multifunctional prostheses based on the combination of speech and EMG signals has been proposed, where the amputee’s speech signals were used to select a mode of a prosthetic arm with multiple DOFs, and then the EMG signals recorded from the residual muscles of limbs were used to determine the motion class involved in the selected mode, and execute the motion. This strategy takes the advantage of speech such as direct, flexible, no long-term training process, and etc, and the muscle burden is released from the awkward mode-switching by the co-contraction of muscle pair which is used in the traditional control method.

The proposed strategy may be implemented to the current myoelectric prostheses by embedding an speech processing module and a mode-selection terminal, and the present EMG-controlled motion-execution part can still be used, which does not increase the cost very much. The practicality of the strategy has been temporarily demonstrated by experimental results with high mode-selection rate and relatively short action-execution time. A positive users’ experience was also reported by all the subjects.

Acknowledgment

The authors would like to thank all the participants in the study for their assistance in the experiment and data processing.

References

[1] P. A. Parker and R. N. Scott, “Myoelectric control of prostheses,” *Crit. Rev. Biomed. Eng.*, vol. 13, no. 4, pp. 283–310, 1986.

[2] B. Hudgins, P. Parker, and R. Scott, “A new strategy for multifunction myoelectric control”, *IEEE Trans. Biomed. Eng.*, vol. 40, no. 1, pp. 82–94, Jan. 1993.

[3] S. H. Park and S. P. Lee, “EMG pattern recognition based on artificial intelligence techniques,” *IEEE Trans. Rehab. Eng.*, vol. 6, no. 4, pp. 400–405, Dec. 1998.

[4] F. H. Y. Chan, Y. S. Yang, F. K. Lam, Y. T. Zhang, and P. A. Parker, “Fuzzy EMG classification for prosthesis control,” *IEEE Trans. Rehab. Eng.*, vol. 8, no. 3, pp. 305–311, Sep. 2000.

[5] A. B. Ajiboye and R. F. Weir, “A heuristic fuzzy logic approach to EMG pattern recognition for multifunctional prosthesis control,” *IEEE Trans. Neural. Syst. Rehabil. Eng.*, vol. 13, no. 3, pp. 280–291, Sep. 2005.

[6] A. D. Chan and K. B. Englehart, “Continuous myoelectric control for powered prostheses using hidden Markov models,” *IEEE Trans. Biomed. Eng.*, vol. 52, no. 1, pp. 121–124, Jan. 2005.

[7] Y. H. Huang, K. B. Englehart, B. Hudgins, and A. D. C. Chan, “A Gaussian mixture model based classification scheme for myoelectric control of powered upper limb prostheses,” *IEEE Trans. Biomed. Eng.*, vol. 52, no. 11, pp. 1801–1811, Nov. 2005.

[8] L. Hargrove, K. Englehart, and B. Hudgins, “A comparison of surface and intramuscular myoelectric signal classification,” *IEEE Trans. Biomed. Eng.*, vol. 54, no. 5, pp. 847–853, May 2007.

[9] G. Li, A. E. Schultz, and T. A. Kuiken, “Quantifying pattern recognition-based myoelectric control of multifunctional transradial prostheses,” *IEEE Trans. Neural. Syst. Rehabil. Eng.*, vol. 18, no. 2, pp. 185–192, Apr. 2010.

[10] Y. Geng, P. Zhou, and G. Li, “Toward attenuating the impact of arm positions on electromyography pattern-recognition based motion classification in transradial amputees,” *J. NeuroEng. Rehab.*, vol. 9, no. 74, Oct. 2012.

[11] M. S. Pinzur, J. Angelats, T. R. Light, R. Izuierdo, and T. Pluth, “Functional outcome following traumatic upper limb amputation and prosthetic limb fitting,” *J. Hand Surg.-Am. Vol.*, vol. 19A, no. 5, pp. 836–839, Sep. 1994.

[12] K. Englehart, B. Hudgins, P. A. Parker, and M. Stevenson, “Classification of the myoelectric signal using time-frequency based representations,” *Med. Eng. Phys.,* vol. 21, no. 6–7, pp. 431–438, Jul-Sep. 1999.

[13] H. Huang, P. Zhou, G. Li, and T. Kuiken, “EMG electrode optimization for targeted muscle reinnervation based neural machine interface,” *IEEE Trans. Neural. Syst. Rehabil. Eng.*, vol. 16, no. 1, pp. 37–45, 2007.

[14] T. A. Kuiken, L. A. Miller, R. D. Lipschutz, B. A. Lock, K. Stubblefield, P. D. Marasco, P. Zhou, and G. A. Dumanian, “Targeted reinnervation for enhanced prosthetic arm function in a woman with a proximal amputation: a case study,” *Lancet*, vol. 369, no. 9559, pp. 371–380, Feb. 2007.

[15] P. Zhou, M. M. Lowery, K. Engleharts, H. Huang, G. Li, L. Hargroves, J. Dewald, and T. Kuiken, “Decoding a new neural-machine interface for control of artificial limbs,” *J. Neurophysiol.*, vol. 98, no. 5, pp. 2974–2982, Nov. 2007.

[16] T. A. Kuiken, G. Li, B. A. Lock, R. D. Lipschutz, L. A. Miller, K. A. Stubblefield, and K. B. Englehart, “Targeted muscle reinnervation for real-time myoelectric control of multifunctional artificial Arms,” *JAMA: J. Am. Med. Asso.*, vol. 301, no. 6, pp. 619–628, Feb. 2009.

[17] Y. Ephraim, “A Bayesian estimation approach for speech enhancement using hidden Markov models,” *IEEE Trans. Sign. Process.*, vol. 40, no. 4, pp. 725–735, Apr. 1992.

[18] L. Muda, M. Begam, and I. Elamvazuthi, “Voice Recognition Algorithms using Mel Frequency Cepstral Coefficient (MFCC) and Dynamic Time Warping (DTW) Techniques,” *J. Comp.*, vol. 2, no. 3, pp. 138–143, Mar. 2010.