A Prosthesis Control System Based on the Combination of Speech and sEMG Signals and Its Performance Assessment

Zheng Wei[1,2], Peng Fang[1,2,*], Lan Tian[1], Qifang Zhuo[1,2], and Guanglin Li[1,2]

[1] Shenzhen Institutes of Advanced Technology,
Chinese Academy of Sciences, Shenzhen, China
[2] University of Chinese Academy of Sciences, Beijing, China
{zheng.wei,peng.fang,lan.tian,qf.zhuo,gl.li}@siat.ac.cn

Abstract. Surface electromyographic (sEMG) signals from the residual limb muscles after amputation have been widely used for prosthesis control. However, for the amputees with high-level amputations, there usually exists a dilemma that the sEMG signal sources for prosthesis control are limited but more limb motions need to be recovered, which strongly limits the practicality of the current myoelectric prostheses. In order to operate prostheses with multiple degrees of freedom (DOF) of movements, several control protocols have been suggested in some previous studies to deal with this dilemma. In this paper, a prosthesis control system based on the combination of speech and sEMG signals (*Strategy 1*) was built up in laboratory conditions, where speech commands were applied for the prosthesis joint-mode switching and sEMG signals were applied to determine the motion-class and execute the target movement. The control performance of the developed system was evaluated and compared with that of the traditional control strategy based on the pattern recognition of sEMG signals (*Strategy 2*). The experimental results showed that the difference between *Strategy 1* and *Strategy 2* was insignificant for the control of a 2-DOF prosthesis, but *Strategy 1* was much better in the control of a prosthesis with more DOFs in comparison to *Strategy 2*. In addition, the positive user experience also demonstrated the reliability and practicality of *Strategy 1*.

Keywords: sEMG, Speech, Prosthesis control, Pattern recognition, Limb Amputee.

1 Introduction

Multifunctional prostheses are very useful for amputees to recover the lost body functions and improve their life quality. Up to now, most modern motorized prostheses are controlled with the surface electromyographic (sEMG) signals from muscles of residual limbs, and several methods have been developed to realize possible practical control of myoelectric prostheses [1-4]. Conventionally, sEMG signals from a pair of

* Corresponding author.

Y. Zhang et al. (Eds.): HIS 2014, LNCS 8423, pp. 72–82, 2014.

residual muscles are applied to control one degree of freedom (DOF) of movements [5]. However, in the case of high-level amputations, the residual muscles are usually limited and cannot supply sufficient sEMG signals for the control of multifunctional prostheses with multiple DOFs. In order to control multiple DOFs with a pair of residual muscle (a muscle pair), a so-called “mode switching” [6] procedure is used for the switching among different joints, which is mostly realized by simultaneous co-contractions of a muscle pair. Take a 3-DOF-prosthesis as example: it has three joint-modes of “hand”, “wrist” and “elbow”, and sEMG signals recorded from the muscle pair of bicep-tricep determine the motion-classes of each joint. The switching among the three joint is performed by the co-contraction of bicep-tricep muscle pair as well. The predetermined switching order is “hand-wrist-elbow…” and the current mode is “hand”. If the user wants to do an elbow movement, he/she has to conduct the co-contraction of the muscle pair twice to switch the joint-mode from “hand” to “wrist” and from “wrist” to “elbow”, and then contracts either the bicep or the tricep to execute an elbow movement. In this way, users have to take a lot of time and efforts in the mode switching and always remember the current joint-mode. As a result, the control method based on the sequential mode switching is strongly limited and commonly rejected by most of the users.

To improve the control performance of the current myoelectric prostheses, a control strategy based on the pattern recognition of sEMG signals has been proposed [7-8]. Here, a pattern recognition algorithm is applied to classify the target motion-classes by decoding the sEMG signals from residual muscles. However, the pattern recognition method is still not much practical if there are not enough residual muscles, especially after high-level amputations. In addition, the signal quality, the operation flexibility, and the real-time patter recognition algorithm are also big challenges which prevent the further improvement of this method.

Some extra non-sEMG signals have been taken into account to overcome the problem of insufficient sEMG signal sources in the present prosthesis control, and one of the possible candidate control information may be the human speech [9]. Speech is a native ability for most people except for the language disabled, and is also independent off the limb functions and amputation conditions. In our pilot study [9], the speech signals were used as additional information and combined together with the sEMG signals for the control of a multifunctional myoelectric prosthesis. A PC-based control system was built up and the primal results demonstrated the practicality of the proposed strategy.

In order to further investigate the performance of the recently proposed control strategy for its practical application, in this work, an embedded myoelectric-prosthesis control system was built up in laboratory conditions by the combination of speech and sEMG signals (*Strategy 1*). For comparison purpose, another system based on the pattern recognition of sEMG signals was also set up. The control performances of both systems were evaluated and compared. The outcomes of this study could make an important progress of the proposed method toward developing a practical prosthesis control system for clinical use.

2 Method

2.1 Subjects

In this study, three able-bodied subjects (marked as *A1*, *A2* and *A3*) and one unilateral transradial amputee (marked as *B1*) were recruited. The demographic information of the subjects is summarized in Table 1. All subjects had full language competence. The protocol of this research was approved by the Institutional Review Board of Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences. All subjects gave written informed consent and provided permission for publication of photographs for a scientific and educational purpose.

Table 1. Demographic information of the subjects recruited in the study

Subjects	Gender	Age	Body situation	Test side
A1	Male	22	Able-bodied	Right
A2	Male	31	Able-bodied	Right
A3	Male	24	Able-bodied	Right
B1	Male	24	Right forearm amputated	Right

2.2 Control Strategies

In the experiments, a commercial myoelectric prosthesis from *Shanghai Kesheng MH32, China* was used. This prosthesis has three joint-modes of "hand", "wrist", and "elbow", and each joint-mode involves two motion-classes as "hand closing/opening", "wrist pronation/supination", and "elbow flexion/extension". Two control systems based on different strategies were built up and examined in the study:

(1) *System 1*(*Strategy 1*): Prosthesis control based on the combination of speech and sEMG signals

In *Strategy 1*, the joint-mode switching was firstly conducted according to the user's speech commands, and then the sEMG signals from a muscle pair were used to determine one of two motion-classes involved in the selected joint-mode and execute the target movement, as shown in Fig 1.

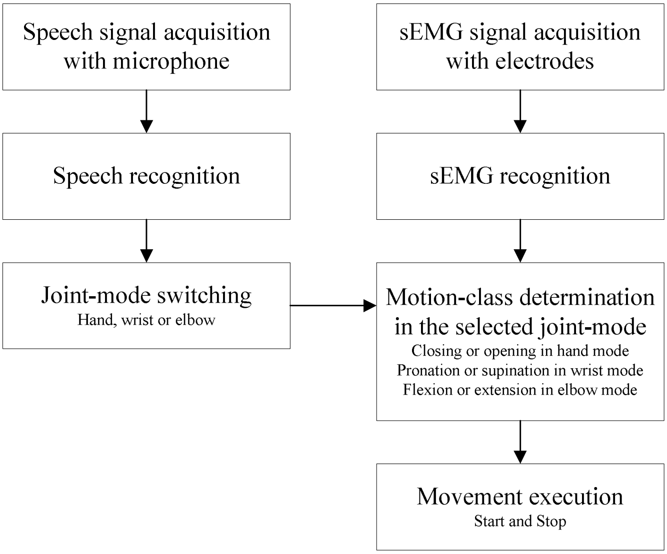

Fig. 1. *Strategy 1:* Prosthesis control based on the combination of speech and sEMG signals

In an office environment (background noise of 55±5 dB), speech signals were acquired with a commercial throat microphone that was attached to the subject's larynx near the vocal folds. Different from normal microphones, the throat microphone only recorded the speech signals transferred through the larynx and was insensitive to the background noise, which might improve the recognition accuracy [10]. Practically, any words could be used as the speech commands depending on users' preference. In this study, three single Chinese characters as shown in brackets, "hand (手)", "wrist (腕)", and "elbow (肘)", were used as the keywords to represent each of the three joint-modes in *Strategy 1*. To avoid the recognition failure due to the dialect or accent of different subjects, an individual recognition template was created for each subject instead of a preset standard speech bank. The speech-signal processor used in this work was SPCE061A (*SUNPLUS Technology*). After amplification, second-order butterworth band-pass filtration (passing band of 340-3700 Hz), and A/D conversion, the speech signals were recognized with the dynamic time warping (DTW) algorithm [11]. In addition, instead of the linear prediction coefficients (LPC) [12], the mel-frequency cepstral coefficients (MFCC) [11] based on auditory mode was used to extract the speech characteristic parameters to improve recognition precision. The acquired speech signals were compared to each template with the DTW algorithm, and the template that had a minimum Euclidian distance to the speech signals was considered as the most matching one and its corresponding keyword was considered as the recognition result to represent the desired joint-mode.

For sEMG signal acquisition, the muscle pair of flexor-extensor was used as the signal source, and each muscle was attached with a bipolar sEMG electrode respectively, as shown in Fig 2. Each muscle corresponded to one of two motion-classes involved in a selected joint-mode, e.g. flexor to "hand closing" and extensor to "hand opening" in the "hand mode", and similarly, flexor to "wrist pronation" and extensor to "wrist supination" if the "wrist mode" was chosen.

Fig. 2. Position of bipolar electrodes on the residual forearm of the transradial amputee in *Strategy 1*

In the embedded system, MC9S12XEP100 (*Freescale semiconductor company, USA*) was used as the micro-controller and responsible for the signal processing and the prosthetic arm driving. sEMG signals were acquired through the bipolar electrodes with a sampling rate of 1000 Hz. After amplification, 50 Hz notch filtration, and A/D conversion, the sEMG signals were transmitted to the micro-controller. The mean absolute value (MAV) was used as the characteristic parameter of the sEMG signals and the K-nearest neighbors (KNN) algorithm was applied for the sEMG signal decoding [13]. Compared to other algorithms, the KNN algorithm is one of the simplest machine learning algorithms with high performance [14]. In this work, it was found that K=3 could achieve the real-time processing and relatively high accuracy.

The control system was composed of five parts as micro-controller, sEMG acquisition module, speech acquisition module, speech recognition module, and motor driver, as shown in Fig 3. The hardware realization is shown in Fig 4.

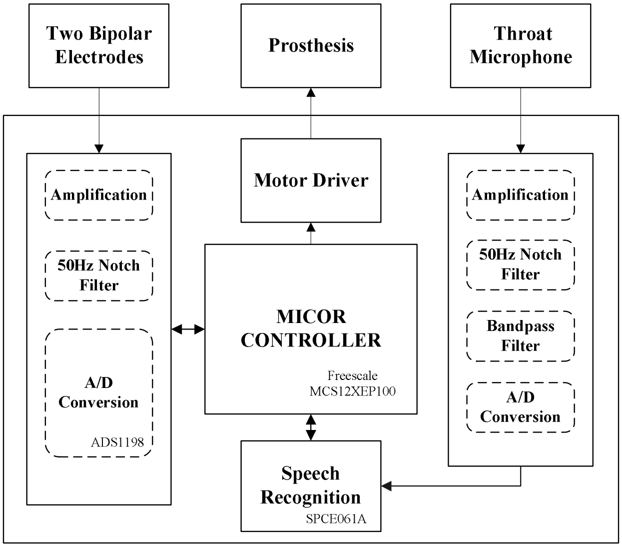

Fig. 3. Diagram of the control system based on the combination of speech and sEMG signals

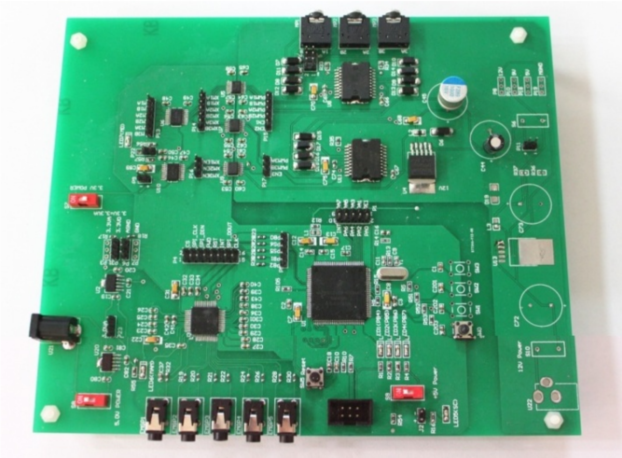

Fig. 4. Hardware realization of the system shown in Fig 3

(2) *System 2* (*Strategy 2*): Prosthesis control based on the pattern recognition of sEMG signals

For a comparison purpose, a control system based on the pattern recognition of sEMG signals was also developed. In this study, it was not necessary to recover the elbow movements for the transradial amputee and therefore the joint-mode of elbow was excluded in *Strategy 2*. Five motion-classes, "hand opening", "hand closing", "wrist pronation", "wrist supination", and "no-movement", were defined. Each subject was asked to accomplish a simple training process to construct his specific pattern template. Four time-domain features, mean absolute value (MAV), waveform length (WL), zero-crossings rate (ZC), and slope sign change (SSC), were extracted as the characteristic parameters of the sEMG signals. The length of the analysis window was 150 ms with an overlap of 100 ms. sEMG signals were decoded with the pattern recognition algorithm of Linear Discriminant Analysis (LDA) [15], and the motion-classes were classified according to the pattern template.

Here, a commercial wireless biological signal acquisition system (*Delsys Trigno Wireless, USA*) was used to acquire sEMG signals. Four sEMG electrodes were placed on the subjects' full/residual muscles of the forearm, as shown in Fig 5. sEMG signals were recorded with a sampling rate of 1000 Hz and transmitted to the computer via a data acquisition card (*USB-6218, National Instruments Corp, USA*).

Fig. 5. Four bipolar sEMG electrodes were placed on the residual forearm of the transradial amputee in *Strategy 2*

2.3 Experiment Protocol

To evaluate and compare the control performance of *Strategy 1* and *Strategy 2*, a measure of *task execution time* was proposed and two different functional tasks were designed.

Task Execution Time**:** The time needed to complete a whole task without any misoperation. A task might contain a series of movements, and the procedure to finish a single movement included the joint-mode switching (in *Strategy 1*) and the movement execution (start and stop, in both strategies).

Task 1 **(Three Joint-Modes Applied):** Subjects were required to complete a task of "water pouring" continuously without any misoperation, which included a series of following movements: "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. Since it was not possible to execute elbow movements with *Strategy 2*, only *Strategy 1* was tested in this task.

***Task 2* (Two Joint-Modes Applied):** Similar to *Task 1* but the elbow-joint movements were excluded, i.e. “hand-closing” to hold a cup with water inside, “wrist-pronation” to pour the water out, and then “wrist-supination” and “hand-opening” to return. Both strategies were tested in this task.

All the tests were repeated at least five times and the results were calculated as the average values over the repeated measurements.

3 Results

In *Task 1* where three joint-modes were applied, all the able-bodied subjects and the transradial amputee could complete the specified “water pouring” successfully with *Strategy 1*, and the *task execution time* of each subject is summarized in Table 2. As can be seen, the able-bodied subjects *A1*, *A2* and *A3* spent 19.4, 21.7, and 21.9 s to complete the task, respectively. The transradial amputee *B1* spent similar time of 21.2 s to complete the same task as the able-bodied subjects did.

Table 2. *Task execution time* to complete *Task 1* with *Strategy 1* for all the subjects

Subjects	*Task execution time (s)*
	Strategy 1
A1	19.4±1.3
A2	21.7±0.5
A3	21.9±1.0
B1	21.2±2.8

In *Task 2* where two joint-modes were applied, all the subjects could also complete the required movement series successfully, and the *task execution time* is shown in Table 3. With *Strategy 1*, the able-bodied subjects *A1*, *A2* and *A3* spent 10.8, 10.5, and 11.4 s to complete the task, respectively, and the *task execution time* for the transradial amputee *B1* was 11.8 s. With *Strategy 2*, the time was 11.1, 10.3 and 9.2 s for the able-bodied subjects, respectively, and 9.5 s for the transradial amputee *B1*. The average value of *task execution time* for *Strategy 2* was slightly less than that for *Strategy 1*. Fig 6 shows the comparison of the statistical analysis for the *task execution time* with different control strategies.

Table 3. *Task execution time* to complete *Task 2* with *Strategy 1* and *Strategy 2* for all the subjects

Subjects	*Task execution time (s)*	
	Strategy 1	*Strategy 2*
A1	10.8±0.8	11.1±0.6
A2	10.5±1.2	10.3±2.4
A3	11.4±0.9	9.2±1.0
B1	11.8±0.4	9.5±0.7
Average	11.1	10.0

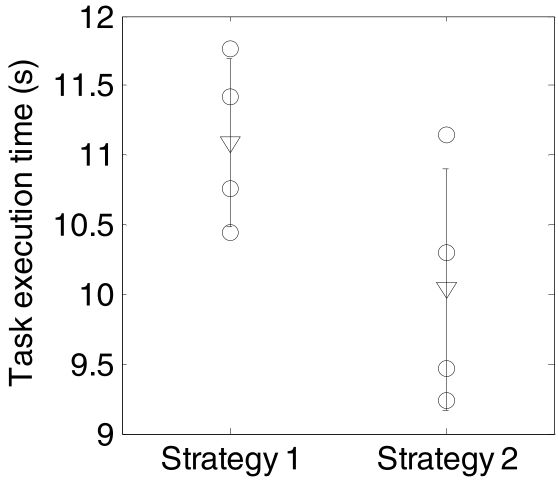

Fig. 6. Comparison of *task execution time* for *Strategy 1* and *Strategy 2* in *Task 2*, where the circles represent the value of each subject, and the triangles represent the calculated average value

4 Discussion

In *Strategy 1*, only a pair of muscles was applied to control the prosthesis with multiple DOFs, which was quite suitable for the amputees without sufficient residual muscles. With *Strategy 1*, the burden from the frequent muscle-pair co-contraction used in the traditional control strategy may be released. What is more, the speech-based joint-mode switching is very flexible and it will be possible to control more DOFs by just adding more speech commands to the system. In *Strategy 2,* high quality sEMG signals and long training process were required for accurate motion-class classification. In addition, it was found that during the experiments the subjects always became tired just after the execution of a few movements. With *Strategy 2*, it will be difficult if more DOFs are required since a more sophisticated recognition algorithm is needed. It should be noted that the transradial amputee still owned the elbow joint, and thus the prosthetic movements of elbow were actually not necessary for them. In the experiments, the joint-modes of “hand”, “wrist”, and “elbow” were just used to represent a 3-DOF experimental prosthesis to assess the control performance. In the case of transhumeral amputation, the users can still control a 3-DOF prosthesis (e.g. hand, wrist, and elbow) by the residual muscles of upper arm (e.g. bicep and tricep) with *Strategy 1*. However, with *Strategy 2*, the transhumeral amputees cannot conduct any hand or wrist movements because no sEMG signals from forearm can be acquired.

With *Strategy 1*, the able-bodied subjects and the transradial amputee achieved similar experimental results of *task execution time* in both *Task 1* and *Task 2* since they all had full language competence. Generally speaking, in *Task 2*, the average *task* *execution time* with *Strategy 1* is slightly longer than that of *Strategy 2*. This is because the pronunciation and recognition of speech signals took some time. Compared with *Strategy 2*, *Strategy 1* does not have obvious advantage for the control of prostheses with less DOFs. Nevertheless, *Strategy 1* can be easily expanded for more DOFs and complicated tasks, but *Strategy 2* will be strongly limited.

5 Conclusion

In this study, a myoelectric-prosthesis control system based on the combination of speech and sEMG signals was built up. Its control performances were evaluated through practical operations and compared with the system based on the pattern recognition of sEMG signals. The strategy of the combination of speech and sEMG signals is flexible and easy to use, which has been approved by the positive user experiences. In addition, it can be expanded in the case that more DOFs are required. The strategy based on the pattern recognition of sEMG signals is practical only if sufficient residual muscles can be obtained and less DOFs are needed. The control systems designed in this work is practical and stable and can be embedded into the present myoelectric prostheses for applications.

Acknowledgments. 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 Governmental Basic Research Grand (#JCYJ20120617114419018), and the Guangdong Innovation Research Team Fund for Low-cost Healthcare Technologies.

References

1. Graupe, D., Cline, W.K.: Functional separation of EMG signals via ARMA identification methods for prosthesis control purposes. IEEE Transactions on Systems, Man, and Cybernetics 5(2), 252–259 (1975)
2. Zardoshi-Kermani, M., Wheeler, B.C., Badie, K., Hashemi, R.M.: EMG feature evaluation for movement control of upper extremity prostheses. IEEE Transactions on Rehabilitation Engineering 3(4), 324–333 (1995)
3. Tenore, F., Ramos, A., Fahmy, A., Acharya, S., Etienne-Cummings, R., Thakor, N.V.: Towards the control of individual fingers of a prosthetic hand using surface EMG signals. In: 29th Annual International Conference of the IEEE EMBS, Lyon, pp. 6145–6148 (2007)
4. Castellini, C., van der Smagt, P.: Surface EMG in advanced hand prosthetics. Biological Cybernetics 10, 35–47 (2009)
5. Parker, P.A., Scott, R.N.: Myoelectric control of prostheses. Critical Reviews in Biomedical Engineering 13(4), 283 (1986)
6. Williams III, T.W.: Practical methods for controlling powered upper-extremity prostheses. Assistive Technology 2(1), 3–18 (1990)
7. Young, A.J., Smith, L.H., Rouse, E.J., Hargrove, L.J.: Classification of simultaneous movements using surface EMG pattern recognition. IEEE Transactions on Biomedical Engineering 60(5), 1250–1258 (2013)

8. Li, G., Li, Y., Yu, L., Geng, Y.: Conditioning and sampling issues of EMG signals in motion recognition of multifunctional myoelectric prostheses. Annals of Biomedical Engineering 39(6), 1779–1787 (2011)
9. Fang, P., Wei, Z., Geng, Y., Yao, F., Li, G.: Using speech for mode Selection in control of multifunctional myoelectric prostheses. In: 35th Annual International Conference of the IEEE EMBS, Osaka, pp. 3602–3605 (2013)
10. Mubeen, N., Shahina, A., Khan, A.N., Vinoth, G.: Combining spectral features of standard and throat microphone signal for speaker recognition. In: 2012 International Conference on Recent Trends in Information Technology (ICRTIT), Chennai, pp. 119–122 (2012)
11. Muda, L., Begam, M., Elamvazuthi, I.: Voice recognition algorithms using mel frequency cepstral coefficient (MFCC) and dynamic time warping (DTW) techniques. Journal of Computing 2(3) (2010)
12. Paul, A.K., Das, D., Kamal, M.M.: Bangla speech recognition system using LPC and ANN. In: 7th International Conference on Advances in Pattern Recognition, Kolkata, pp. 171–174 (2009)
13. Al-Faiz, M.Z., Ali, A.A., Miry, A.H.: A k-nearest neighbor based algorithm for human arm movements recognition using EMG signals. In: 1st International Conference on Energy, Power and Control (EPC-IQ), Basrah, pp. 159–167 (2010)
14. Bay, S.D.: Combining nearest neighbor classifiers through multiple feature subsets. In: ICML, vol. 98, pp. 37–45 (1998)
15. Chen, L., Geng, Y., Li, G.: Effect of upper-limb positions on motion pattern recognition using electromyography. In: 4th International Congress on Image and Signal Processing (CISP), Shanghai, vol. 1, pp. 139–142 (2011)