Electromyography signal segmentation method based on spectral subtraction backtracking

CAI Huihui[1,2,3], ZHANG Yakun[2,3]*, XIE Liang[2,3], YIN Erwei[1,2,3], YAN Ye[2,3], and MING Dong[1]
1. Academy of Medical Engineering and Translational Medicine, Tianjin University, Tianjin 300072, China
2. Tianjin Artificial Intelligence Innovation Center (TAIIC), Tianjin 300450, China
3. Defense Innovation Institute, Academy of Military Sciences (AMS), Beijing 100071, China

(Received 13 April 2022; Revised 11 June 2022)

Surface electromyography (EMG) is a bioelectrical signal that recognizes speech contents in a non-acoustic form. Activity detection is an important research direction in EMG research. However, in the low signal-to-noise ratio (*SNR*) environment, it is difficult for traditional methods to obtain accurate active signals. This paper proposes a new energy-based spectral subtraction backtracking (E-SSB) method to segment EMG active signal in the low *SNR* environment. Compared with traditional energy detection, the algorithm in this paper adds spectral subtraction (SS) to filter out the clutter, and raises a retrospective idea to improve the classification performance. The experiment results show the proposed activity detection method is more effective than other methods in the low *SNR* environment.
Document code: A **Article ID:** 1673-1905(2022)10-0623-5
DOI https://doi.org/10.1007/s11801-022-2058-x

Electromyography (EMG) is a bioelectrical signal, which is a superposition of muscle and electrical nerve activity on the skin's surface in space and time. The sum of muscle motor units constitutes the intensity of the EMG signal, which reflects the activity characteristics of nerves and muscles[1,2]. A mapping exists between facial muscles and acoustic sequence during the speech, allowing the recognition of speech content without uttering a sound. This silent speech recognition (SSR) technique without vocalization is well suited to avoid external noise interference[3,4]. Therefore, facial EMG research is significant in SSR.

Signal segmentation is a way of making decisions and extracting event ranges in electronic signal processing systems. It is essential for practical applications, especially in communications, speech recognition, and biomedical research[5,6].

Signal segmentation is also known as signal activity detection in biomedicine. Signal activity detection is useful for motion pattern analysis, neuromuscular pathology diagnosis, and auxiliary speech recognition. EMG activity detection is designed to determine an action's start and end position from continuous signals. Thus, it is an important premise of intelligent control, human-computer interaction, and decoding information. An excellent activity detection algorithm can accurately segment the EMG active signal and reduce the computational complexity of model training[7].

Many activity detection methods have emerged in past studies, such as the entropy method[8], energy method[9], cepstrum distance method[10], and so on.

In 2016, CHENG et al[8] used SampEn to characterize the activity endpoints of EMG signals effectively. It performed well in short-term muscle contraction and relaxation and noise resistance as well. In 2021, KANG et al[10] proposed a new algorithm for muscle activity detection. They distinguished the onset state of EMG signal by calculating binarized EMG and polarized EMG to reflect the activation state of arm muscles.

ALMEIDA BRITTO et al[11] used the Teager-Kaiser energy operator to measure the contraction and diastole of muscle activity. Teager-Kaiser highlights muscle movement activity while suppressing background noise to achieve endpoint detection.

The energy detection method is effective in identifying active and inactive signals in a high signal-to-noise ratio (*SNR*) environment. BENGACEMI et al[12] demonstrated that energy detection could not be directly applied to EMG signals and proposed a new method named the adaptive linear energy detector (ALED). Two years later, based on the signal energy, they added a sequential statistical way of signal baseline estimation to deal with the problem of inaccurate detection results[9].

However, it is still challenging to accurately determine the endpoint of an active signal when an EMG signal is captured in low-noise environment.

ZHANG et al[13] used the cepstrum distance method to recognize speech endpoint information and integrated spectral subtraction (SS) to improve the accuracy of speech endpoint detection in low *SNR*. However, this method loses a lot of important information, which affects the speech recognition results.

* E-mail: ykzhang1222@126.com

This paper focuses on these problems and proposes an energy-based spectral subtraction backtracking (E-SSB) method to detect EMG active signal. SSB algorithm ensures the quality of the signal while improving the *SNR*, and the energy thresholding method provides the algorithm's efficiency.

The EMG dataset was acquired by Neuracle's wireless electromechanical acquisition system NEW308M. The sampling frequency is 1 000 Hz with bipolar differential inputs. The device has 8 channels, and we used 6 channels on EMG signal transmission.The dataset consists of 50 subjects and 101 instructions, and each instruction contains 10 pieces of EMG data. Due to the influence of acquisition settings, the non-speech part of the EMG signal has intense noise, which makes the traditional activity detection method invalid, so the E-SSB method is proposed in this paper to solve this problem.

EMG data will carry some interference and noise due to the influence of collection equipment and wires, such as direct current (DC) offset, zero drift, industrial frequency noise, and low-frequency noise. So, before the activity detection, the experiments first use notch filtering to remove the industrial frequency of 50 Hz and filter out the corresponding harmonic interference. Then we use the 4th order Butterworth bandpass filter to control the frequency band range from 10 Hz to 400 Hz[14]. Fig.1 shows the EMG signal of one channel, where (a) is the original EMG signal, and (b) is the signal after notch filtering and bandpass filtering.

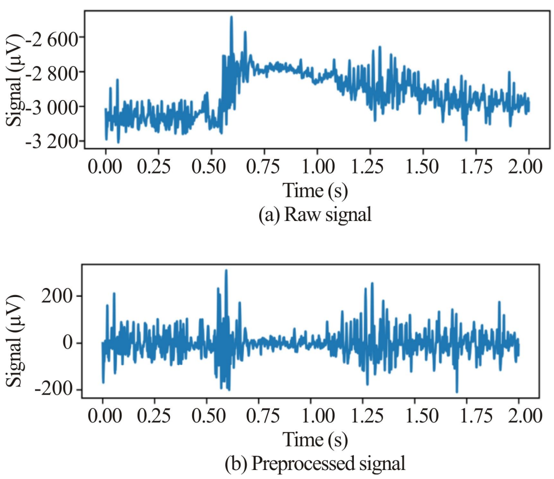


Fig.1 Raw signal and preprocessed signal

Activity detection is the most important part in this paper. The detection algorithm is divided into three steps. The first step is to filter out the noise through SS[15] and improve the *SNR* of EMG signal. The second step is to calculate average energy and find out the endpoint of the active signal[3]. Finally, the endpoint coordinates are merged into the preprocessed signal to reduce signal loss. Fig.2 shows the complete activity detection flow.

The following content details the principle of the E-SSB algorithm.

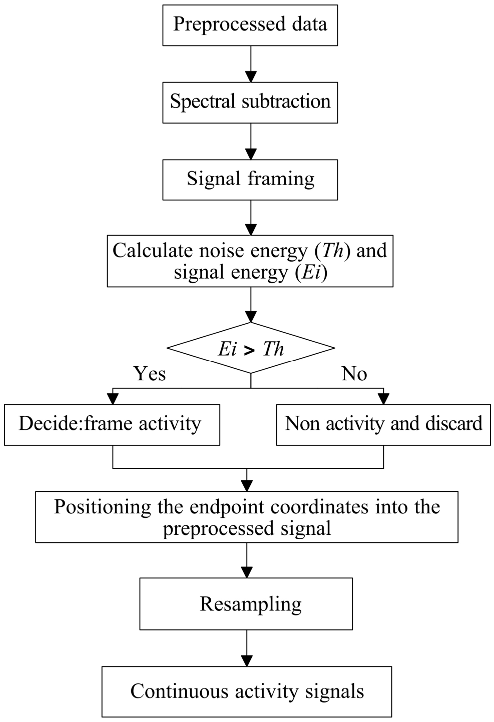


Fig.2 Flow chart of activity detection algorithm

Step 1 The SS method is used to filter out the noise of facial EMG signal.

SS is an efficient signal enhancement algorithm to remove noise and improve *SNR*. The reaction time of human eye is within 200—300 ms. This paper takes the first 200 ms as noise spectrum and then subtracts it from the original EMG signal.

The original EMG signal is defined as $x(t)$, the noise signal is defined as $n(t)$, and the clean EMG signal is defined as $y(t)$, then the relationship among them is

$$y(t)= x(t)- n(t). \tag{1}$$

Since the noise is not related to the EMG signal, the calculation formula of the energy spectrum is as follows

$$| y(w)|^2=| x(w)|^2-| n(w)|^2. \tag{2}$$

The inverse Fourier transform of $y(w)$ is applied to obtain the EMG signal after spectrum reduction.

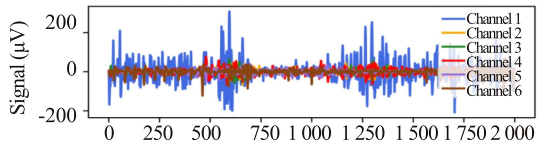
Fig.3(a) shows the signal before spectrum subtraction, in which the non-speech part has intense noise. Fig.3(b) shows the signal after spectrum reduction, and noise in the front of the signal has been removed.

Step 2 Average energy calculation and finding the position where signal energy changes significantly.

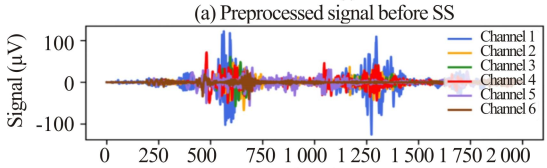
This paper divides EMG signal into windows, sets 25 ms as a window length, and calculates the signal energy value within each window. Then, we take the average energy of 8 windows in the first 200 ms as the threshold value. Because the energy of non-speech signal is minimal after SS, when an active EMG signal emerges, the energy will change abruptly.

The equation for the calculation of the signal energy value is as follows, where N is the number of signal frames, and $x_i(t)$ is the signal frame as

$$E=\frac{1}{N}\sum_{i=0}^{N}\left|x_i(t)\right|^2 . \quad (3)$$



(a) Preprocessed signal before SS



(b) SS signal

Fig.3 Preprocessed and SS signals

The energy value of the signal is compared with the set threshold. If the energy value is greater than threshold value, it is regarded as an active signal.

This paper counted the signal disappearance time to ensure the integrity of algorithm. If the signal disappearance time is greater than 75 ms, the activity detection is judged to be completed. Otherwise, the detection continues.

Step 3 Return the position information into the preprocessed signal.

After SS processing, the facial EMG signal loses much critical information, which reduces the classification effect of different voice commands in the experiments. Therefore, this paper proposes a backtracking method-shown as

$$\boldsymbol{Y}=\{x_{\text{start}}(t),\ x_{\text{start}+1}(t),\ldots,\ x_{\text{end}}(t)\} . \quad (4)$$

The experiments input the endpoint coordinates into the preprocessed signal, where $x_{\text{start}}(t)$ is the starting signal of the EMG, $x_{\text{end}}(t)$ is the ending signal of the EMG, and $\boldsymbol{Y}$ is the final active signal.

To verify the effectiveness of the E-SSB method, the experiment uses classification accuracy as the primary evaluation approach, and combines image verification as auxiliary approach. The combination of two ways can verify the performance of different activity detection methods.

Mel-frequency cepstrum coefficients (MFCCs) are used in EMG features extraction[16]. However, experiments show the classification performance of Mel-frequency spectrum coefficient (MFSC) is better than that of MFCC, so we utilized 58 Mel filter banks to extract MFSC in 6 EMG channels. The MFSC dimension is 58×58×6.

The classification model uses a composite network structure combining convolutional neural network (CNN) and recurrent neural network (RNN). This network was proposed by WANG et al[17] in 2016. In Ref.[18], they verified different configurations of CNN-RNN and proposed the best model framework to identify facial EMG signals. This paper simplifies the model based on Ref.[18], uses CNN for spatial feature learning and gate recurrent unit (GRU) for temporal domain feature learning on the MFSC features, and then connects a fully connected layer and SoftMax function at the end to output probability of 101 labels. The overall architecture of the classification network is shown in Fig.4.

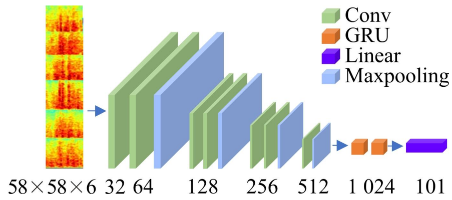


Fig.4 Overall architecture of the classification network

When training the model, the dataset is divided into the training set, validation set, and test set with a ratio of 7: 2: 1. Tab.1 presents the experimental results of three methods for segmenting EMG activity signal, sample entropy, cepstral distance, and energy detection. The energy detection method has the best performance, up to 86%, while the cepstral distance method is only 75.7%, and SampEn's recognition effect is 79%.

Tab.1 Classification accuracy of different methods

Methods	Accuracy (%)
SampEn	79.14
Cepstrum distance	75.73
Energy	86.02
Energy+ SS	85.04
SampEn + **SSB**	86.21
Cepstrum distance + **SSB**	90.80
Energy+ **SSB**	**91.75**

Then, this paper compares the results of these methods with SSB. The energy method with SS reduces the accuracy of the classification results, indicating that the SS method loses the available classification information. However, the classification performance of all these methods is improved after adding the SSB algorithm, which also verifies that the SSB method is effective for facial EMG activity detection. Surprisingly, the cepstral distance method has increased by almost 15%.

Fig.5 shows the activity detection results of 6 methods for the same EMG data. Different colors represent different EMG channels, and the red box in the figure is the EMG signal segmented result by various methods. Fig.5(a) presents the standard EMG activity signal, and (b), (c) and (e) use the original three methods. These methods cannot accurately determine the active signal compared with (a), and it is difficult to determine the appropriate threshold value during the calculation. Fig.5(e) shows the results of the E-SSB proposed in this paper. It accurately gets the active signal and discards the noisy signal to the maximum extent while including the valuable signal. The obtained results largely overlap with the standard results (a), and it more complete than (b). Fig.5(f) shows the cepstrum distance method combined with SSB, which is more accurate than the result of (c), but with much lower accuracy than (e).

Based on Tab.1 and Fig.5, this paper considers the classification results and segmentation accuracy, and finds the E-SSB method works the best.

Fig.6 shows the endpoint detection results for other publicly datasets. (a) shows an EMG-based continuous speech dataset derived from Ref.[19], and (b) indicates the gesture-based NinaProEMG dataset[20]. The results show that the E-SSB algorithm is equally effective in the open dataset, and a suitable threshold is able to detect the EMG active signal accurately. The signal in Fig.6(b) has a high *SNR* and no interference noise, so the conventional method can also obtain accurate results. The E-SSB algorithm mainly solves the activity detection problem in a low *SNR* environment, which is more effective in our dataset.

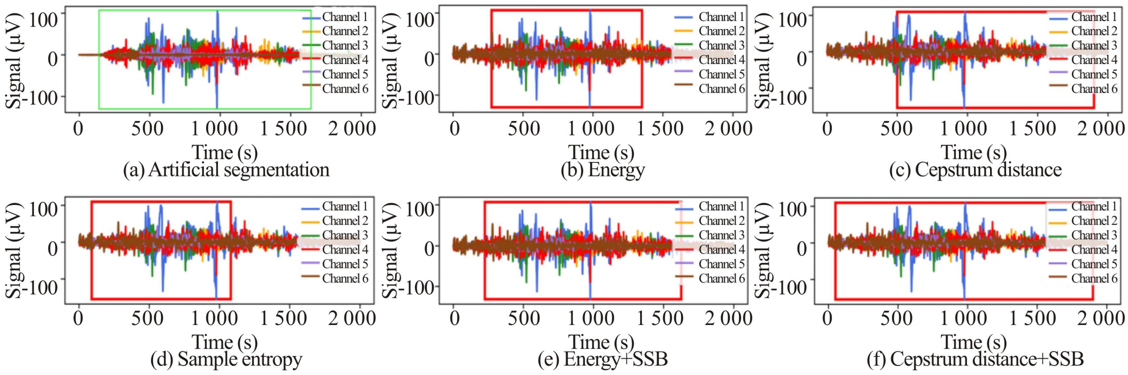


Fig.5 Results of 6 different methods to detect active signal

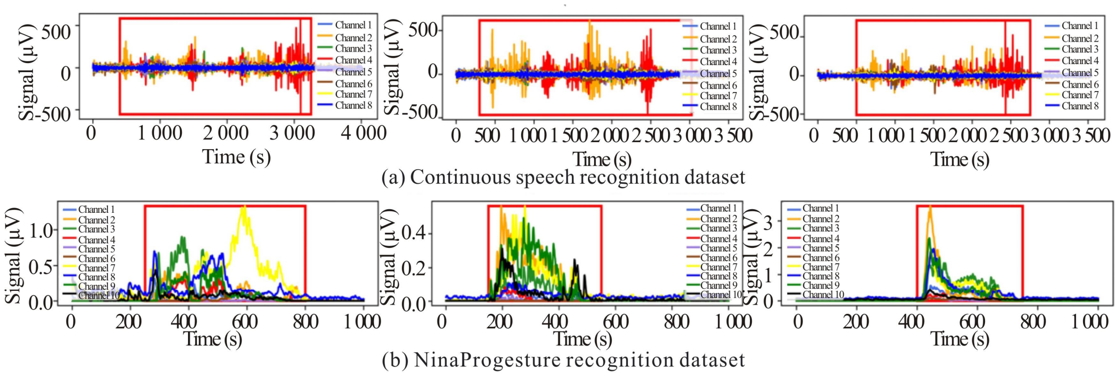


Fig.6 Active signal detection results of 3 different subjects

This paper further compares the influence of different *SNRs* on active signals to verify the effectiveness of the E-SSB in low *SNR* environments. Since our EMG signal generally contains serious noise, it is meaningless for the experiment to add noise in the original high-noise EMG signal. Instead, we use the noise of the EMG signal itself to estimate its *SNR*.

Eq.(5) is the calculation formula of the *SNR*, where V_{signal} is EMG voltage, and V_{noise} is noise voltage,

$$SNR = 20\log_{10}\frac{V_{\text{signal}}}{V_{\text{noise}}}. \tag{5}$$

According to the calculation, the *SNRs* of original signal, preprocessed signal, and SS signal are 0 dB, 3 dB, and 27 dB. As shown in Tab.2, when the *SNR* is 0 dB, the accuracy of E-SSB is 82.03%, but the other two methods fail to recognize the EMG signal. The E-SSB has apparent advantages in a low *SNR* environment.

In a high *SNR* environment of 27 dB, SampEn and cepstral distance methods are effective, and their classification accuracy rates are 86.3% and 90.82%, which are higher than the results at 3 dB. No matter the *SNR* value is 3 dB or 27 dB, the E-SSB method can maintain the highest recognition accuracy rate of more than 90%.

Tab.2 Classification accuracy at different *SNRs*

Methods	*SNR*=0 dB	*SNR*=3 dB	*SNR*=27 dB
SampEn	Error	79.14%	86.30%
Cepstrum distance	10.55%	75.73%	90.82%
Energy+SSB	**82.03%**	**90.02%**	**91.65%**

In conclusion, this paper designs an E-SSB algorithm to solve the problem of inaccurate EMG activity segmentation in low *SNR*. The E-SSB algorithm first uses SS to improve the *SNR*, then uses energy threshold method to calculate the energy value of the EMG signal, and judge the endpoint information for active signal. Finally, the algorithm uses the backtracking method to transfer endpoint information, eliminate the negative impact of SS on the signal and improve the classification accuracy. Furthermore, this paper combines CNN and

GRU frameworks to design a classification model for EMG-based SSR. On a large dataset of 50 subjects, the classification accuracy of 101 classes is as high as 91.65%.

Statements and Declarations

The authors declare that there are no conflicts of interest related to this article.

References

[1] LIU L, YUE W H. Principles of neuro electromyography[M]. Beijing：Science Press, 1983：1-20. (in Chinese)

[2] GONZALEZ-LOPEZ J A, GOMEZ-ALANIS A, DOÑAS J M M, et al. Silent speech interfaces for speech restoration：a review[J]. IEEE access, 2020, 8：177995-178021.

[3] MELTZNER G S, HEATON J T, DENG Y, et al. Development of sEMG sensors and algorithms for silent speech recognition[J]. Journal of neural engineering, 2018, 15(4)：046031.

[4] ZHU M, ZHANG H, WANG X, et al. Towards optimizing electrode configurations for silent speech recognition based on high-density surface electromyography[J]. Journal of neural engineering, 2021, 18(1)：016005.

[5] HIPPENSTIEL R D. Detection theory：applications and digital signal processing[M]. Montrouge：CRC Press, 2017.

[6] HWANG I, CHANG J H. End-to-end speech endpoint detection utilizing acoustic and language modeling knowledge for online low-latency speech recognition[J]. IEEE access, 2020, 8：161109-161123.

[7] MELTZNER G S, HEATON J T, DENG Y, et al. Silent speech recognition as an alternative communication device for persons with laryngectomy[J]. IEEE/ACM transactions on audio speech & language processing, 2017, 25(12)：2386-2398.

[8] CHENG J, CHEN X, PENG H. An onset detection method for action surface electromyography based on sample entropy[J]. Acta electonica sinica, 2016, 44(2)：479.

[9] BENGACEMI H, ABED-MERAIM K, BUTTELLI O, et al. A new detection method for EMG activity monitoring[J]. Medical & biological engineering & computing, 2020, 58(2)：319-334.

[10] KANG K, RHEE K, SHIN H C. A precise muscle activity onset/offset detection via EMG signal[C]//2021 International Conference on Information Networking (ICOIN), January 13-16, 2021, Jeju island, Korea. New York：IEEE, 2021：633-635.

[11] DE ALMEIDA BRITTO P A, DE SOUZA MUNIZ A M, NADAL J. Electromyographic activity of the lower limb in runners with anterior knee pain while running[J]. Research on biomedical engineering, 2021, 37(2)：135-142.

[12] BENGACEMI H, MESLOUB A, OULDALI A, et al. Adaptive linear energy detector based on onset and offset electromyography activity detection[C]//2017 6th International Conference on Systems and Control (ICSC), May 7-9, 2017, Batna, Algeria. New York：IEEE, 2017：409-413.

[13] ZHANG T, ZHANG X B, ZHU X X. Speech endpoint detection with low SNR based on improved cepstrum distance method[J]. Audio engineering, 2017, 41(7)：108-112.

[14] SRISUWAN N, PRUKPATTARANONT P, LIMSAKUL C. Comparison of classifiers for EMG based speech recognition[J]. Journal of physics：conference series, 2020, 1438(1)：012032.

[15] BOLL S. Suppression of acoustic noise in speech using spectral subtraction[J]. IEEE transactions on acoustics, speech, and signal processing, 1979, 27(2)：113-120.

[16] NODERA H, OSAKI Y, YAMAZAKI H, et al. Classification of needle-EMG resting potentials by machine learning[J]. Muscle & nerve, 2019, 59(2)：224-228.

[17] WANG J, YANG Y, MAO J, et al. CNN-RNN：a unified framework for multi-label image classification[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, June 26-July 1, 2016, Las Vegas, USA. New York：IEEE, 2016：2285-2294.

[18] WANG Y, ZHANG M, WU R M, et al. Silent speech decoding using spectrogram features based on neuromuscular activities[J]. Brain sciences, 2020, 10(7)：442.

[19] GADDY D, KLEIN D. Digital voicing of silent speech[EB/OL]. (2020-10-06) [2022-04-10]. https：//arxiv.org/pdf/2010.02960.pdf.

[20] ATZORI M, MÜLLER H. The Ninapro database：a resource for sEMG naturally controlled robotic hand prosthetics[C]//2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, August 30-September 3, 2015, Milan, Italy. New York：IEEE, 2015：7151-7154.