2021 43rd Annual International Conference of the
IEEE Engineering in Medicine & Biology Society (EMBC)
Oct 31 - Nov 4, 2021. Virtual Conference

Verification of Normalization Method to Improve Usability and Versatility among Users of Applications that Predict Continuous Motion Using Electromyography

*Taichi Tanaka, Isao Nambu, Yoshiko Maruyama, and Yasuhiro Wada

Abstract —**In applications using electromyography (EMG), it is important to ensure high performance for all users (versatility among users) and to enable use without prior preparation (usability). Some of the current applications that use EMG normalize the signal through methods based on the measured maximum absolute value of EMG (maEMG), such as dynamic contraction (DC). However, usability is low when using DC because the reference value must be measured first every time the application is used. Further, the versatility among users is low because of the nonlinearity of EMG and the fact that maEMG varies among users. This study aimed to improve usability and versatility among users for continuous classification tasks using EMG. To this end, we developed a normalization method using sliding-window and z-score normalization techniques. The results reveal that the proposed method exhibits higher usability and versatility among users than DC. The proposed method does not require any calibration time, suggesting improved usability, while yielding the same classification accuracy as DC (57% for three target tasks) for a model trained using a subject's own data. Further, for a model trained with other users' data, the proposed method yields a classification accuracy of 53%, which is 18% higher than that of DC (35%), suggesting versatility among users. These results demonstrate that the proposed normalization method improves usability and versatility for users of practical applications that use EMG and perform continuous classification, such as prosthetic hands.**

I. Introduction

Electromyography (EMG) is a bioelectrical signal that increases in amplitude when the subject exercises or contracts a muscle. EMG has been used by researchers to study and develop applications that expand or recover human ability [2]. There are two types of applications, namely, continuous motion classification and continuous kinematics/kinetics parameter estimation [1]. Continuous motion classification is used in prosthetic hands, where established motions are predicted using EMG. On the other hand, continuous kinematics/kinetics parameter estimation is used in assist suits and robot arms for estimating joint angles, angular velocities, or torques. However, EMG has disadvantages, such as nonlinearity, in control applications; further, the characteristic magnitude of EMG depends on factors such as sensor placement and fatigue. When commercializing applications, high performance as well as versatility and usability are required for different users. Further, there are two types of versatility, namely, that among individuals and that among users. Here, we consider only the usability and versatility among users, which are defined as follows. Usability is defined as the ability to use the application immediately and without preparation steps such as calibration and wearing. Versatility among individuals is defined the model performance do not vary with variations in the magnitude of EMG depending on factors such as sensor placement and fatigue. To be versatile, the performance of the application should not vary with time. Further, versatility among users implies that high performance is achieved for any user, even if they do not use a model trained with their own data. This is because for commercialization, applications commonly employ a model trained with other people's data.

Several researchers have considered the problem of reduced usability and versatility caused by normalization methods. In machine learning, normalized data are typically used to improve performance and versatility. EMG normalization conventionally employs either maximal voluntary contraction (MVC) or dynamic contraction (DC). A common aspect of MVC and DC is the use of the maximum absolute value of EMG (maEMG) for normalization. MVC normalization uses EMG measured when subjects/users focus on muscles as much as possible. On the other hand, DC normalization uses EMG measured when subjects/users perform a task. Generally, when applying MVC or DC, the corresponding activity (contracting muscles or performing a task) is repeated a few times, and the maEMG is selected over the repetitions. MVC and DC have a disadvantage in that maEMG, which is used as a reference for normalization, varies among users because of factors such as sensor placement, fatigue rate, and subject characteristics. During the training of a machine learning model and measurement for MVC or DC before using an application, differences in maEMG decrease the application performance. Moreover, the fatigue rate increases with the duration for which the application is used; therefore, the application performance usually decreases with time. Therefore, we need a robust normalization or feature extraction method that overcomes EMG magnitude variations and the strong nonlinearity to enable high-performance applications.

For an application that predicts continuous motions using EMG, the waveform is considered better than the magnitude of amplitude [2,3]. This implies that we can predict motions by extracting changes in EMG that occur during exercise and muscle contractions. Therefore, by sufficiently eliminating noise and normalizing the magnitude of the EMG before feature extraction, real-time applications can achieve usability and versatility among individuals and among users.

In this paper, we propose a normalization method using sliding-window analysis (SWA) and z-scoring that does not

T. Tanaka is with the Department of Science Technology of Innovation, and I. Nambu, Y. Maruyama, and Y. Wada are with the Department of Electrical Engineering, Nagaoka University of Technology, 1603-1 Kamitomika Nagaoka 940-2188, Japan (e-mail: taichi_tanaka@stn.nagaoka.ac.jp).

require a calibration and can normalize the magnitude of EMG in real time to improve usability and versatility. The proposed method is referred to as sliding-window normalization (SWN) and is detailed in Section II.B.

Additionally, we predicted continuous elbow-joint motions (rest, flexion, and extension of three targets) on the right arm of different subjects to compare the usability and versatility among users of SWN and DC. The evaluation methods are detailed in Section II.F. In Section III, we describe the effect of window length on model performance by training a model with data from one subject and testing it with all subjects' data on SWN and DC (section III.A). Finally, we compare SWN and DC in terms of the usability and versatility among users for the model showing the best performance on SWN and DC (Section III.B).

II. Methods

A. Data acquisition

The ethics board of Nagaoka University of Technology approved this study. The subjects were two right-handed men that were 22 and 23 years old. They performed a task including 12 types of elbow-joint movements with four start and end points (Fig. 1(A)), moving from one point to another one of the remaining three points. Each trial consisted of pre-rest (2.0 s), task (2.5 s), and post-rest (0.1 s). The subjects executed the task 30 times for each of the 12 elbow-joint movements, for a total of 360 elbow-joint movements. During data acquisition, the following four success conditions were defined:

1) While not moving during the rest period, the elbow-joint angular velocity is not over 2° /s.
2) During the task period, the task ends between 0 and 2 s.
3) The elbow-joint angle is positioned at the starting point (± 2.0°) during the rest period and at the end point (± 6.0°) after the task finishes.
4) The elbow and shoulder are within a circle of 3-cm radius centered on the initial position.

The experimental setup is shown in Fig. 1(B). We used this configuration to measure the EMG and position data, at predetermined locations, with specialized equipment. The position data were measured at three locations, namely, the hand, elbow, and shoulder, using Optotrak (NDI Inc., sampling rate: 500 Hz). The EMG was measured at the biceps brachii (×4), brachialis (×1), brachioradialis (×1), anconeus (×1), triceps brachii (outside) (×2), triceps brachii (long head) (×2), and extensor carpi radialis longus (×1), totaling 12 locations, by using Trigno Lab Avanti (Delsys, sampling rate: 2000 Hz).

B. Sliding-window normalization

To improve usability and versatility, we used SWA and z-scoring as they do not need a calibration and can normalize the EMG magnitude in real time. SWA is often used for signal processing analysis, time-variant parameter estimation, and so on. SWA is based on using signals as long as the window length is established. When acquiring a new sample, sliding-window slides include the newest sample and discard the oldest one. Furthermore, z-scoring is a normalization method for machine learning. This method

(A) (B)

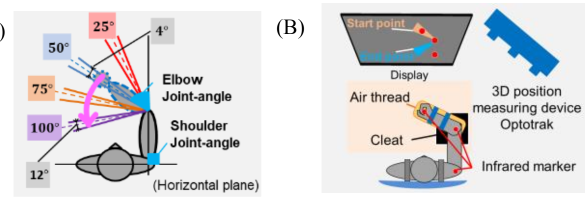

Figure 1. Experimental setup: (A) task, (B) state

makes the mean equal to zero and the standard deviation equal to one. The proposed normalization method, which combines these methods, is referred to as sliding-window normalization (SWN) and is expressed by (1). SWN makes the mean and standard deviation of the samples in the sliding window equal to zero and one, respectively.

$$
\begin{aligned}
SWN\ EMG_n &= \frac{EMG_n - m_t}{\sigma_t} \quad (t-L<n\leq t) \\
m_t &= \frac{1}{L}\sum_{i=0}^{L-1} EMG_{t-i} \\
\sigma_t &= \sqrt{\frac{1}{L}\sum_{i=0}^{L-1}(EMG_{t-i}-m_t)^2}
\end{aligned}
\quad , \tag{1}
$$

where t is the current time, L is the window length, n is the time number in the window length, EMG_i is the i^{th} EMG which is measured or processed, and m_t and σ_t are t^{th} mean and standard deviation of EMG in the sliding-window. The window length is the most important parameter in SWN. Therefore, we investigated its effect on model performance.

C. Normalization during training and testing

In this study, when training a classification model using DC, we normalized the EMG in the training dataset by the maEMG in each training dataset. In addition, when evaluating the performance of the model, we normalized the EMG in the test datasets by the maEMG in each test dataset; the training dataset was also similarly normalized. For the feature extraction, we acquired the same data length as the window length. Subsequently, we calculated the features from the acquired EMG.

In the case of SWN, when training and evaluating the model, maEMG was not used. The window length of the SWN and feature extraction were the same. SWN was applied before feature extraction.

D. EMG processing

For pre-processing, we applied a low-pass filter (3rd-order Butterworth filter, 100 Hz), decimation (from 2000 to 500 Hz) to match the sampling rate of position data, a high-pass filter (3rd-order Butterworth filter, 30 Hz), and normalization (SWN or DC). As a feature for the input of ML, we calculated the mean absolute value (MAV), mean waveform length (MWL), mean zero crossing (MZC), and envelope slope (ES). ES is calculated using (2). For feature comparison, we normalized most features.

$$
\begin{aligned}
ES &= (X^T X)^{-1} X^T Y \\
X &= HilbertTransform(\ SWN\ EMG_{t-L+1\sim t}\ ) \\
Y &= [1, \cdots,\ L]^T
\end{aligned}
\quad . \tag{2}
$$

E. Position processing

Position processing is applied to acquire targets (rest, flexion, and extension, which are elbow-joint motions) for machine learning.

To acquire targets, we applied the zero-phase low-pass filter (2nd-order Butterworth, 20 Hz) in the filtfilt function in MATLAB, converted from position to joint angle, converted from joint angle to joint angular velocity using the differential method, and clustered for rest, flexion, and extension according to (3).

$$f(\dot{\theta}_{elbow}) = \begin{cases} flexion & (\dot{\theta}_{elbow} \geq +2\ [deg./s]) \\ extension & (\dot{\theta}_{elbow} \leq -2\ [deg./s]) \\ rest & (otherwise) \end{cases}, \tag{3}$$

where $\dot{\theta}_{elbow}$ is the elbow-joint angular velocity.

F. Evaluation of usability and versatility

We used different methods to evaluate the usability and versatility among users on the basis of the same performance metric, namely, accuracy, as expressed in (4). The accuracy is used for the easy comparison of multi-class classification results by one overall evaluation value.

$$Accuracy = \frac{Success\ predictions}{Success\ predictions\ +\ Failure\ predictions} \tag{4}$$

Furthermore, we used stratified shuffle-split cross-validation to investigate the effect of maEMG on DC by acquiring the standard deviation of accuracy for each subject. In this study, we divided all data into ten datasets; then, we selected five datasets to train the model and repeated datasets selection ten times for each subject. We used logistic regression as a classification model for three real-time target classifications. The real-time meaning is the models predict the next sample target.

The evaluation criteria of the usability and versatility among users are defined as follows:

1) When evaluating usability, we used models trained with a subject's own data and calculated the accuracy by using all of their own data. If the SWN accuracy for the cross-validation set was as high as or higher than that of DC, SWN was considered to have better usability than DC. This is because the usability is better if the preparation time is shorter. However, SWN does not require calibration, unlike DC.

2) When evaluating the versatility among users, we used models trained with other subjects' data and calculated the accuracy and standard deviation of accuracy by using a subject's own data. If the SWN accuracy for the models trained with other subject's data was higher and the mean of the standard deviation of accuracy of these models was lower than that of DC, SWN was considered to have better versatility among users than DC.

III. Results

A. Effect of window length

To investigate the effect of window length on the performance of SWN and DC, we varied the window length between 50 and 300 ms. Then, we used models that were

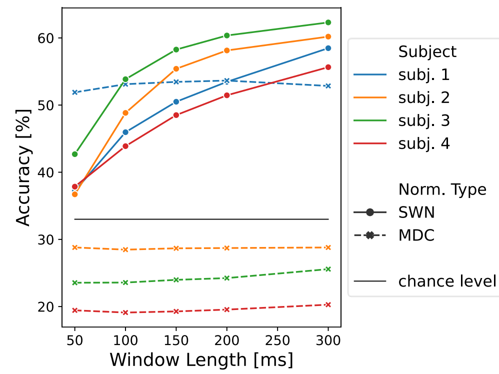

Figure 2. Effect of window length on model performance (using a model trained by the MWL of subject 1).

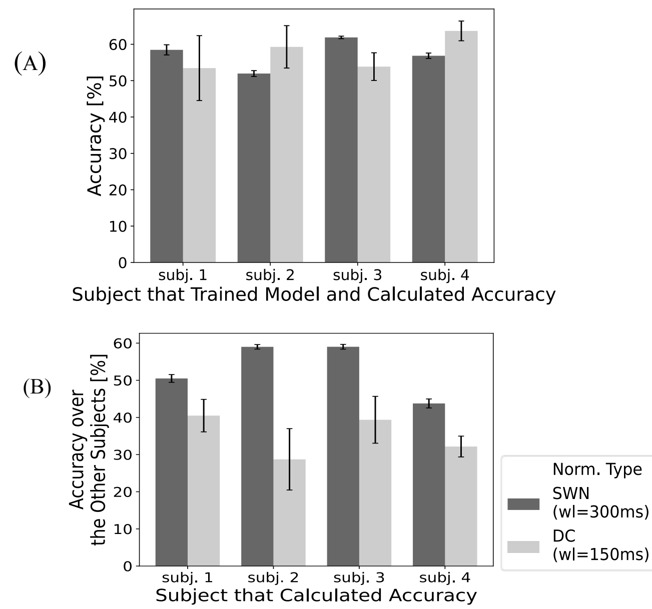

Figure 3. Comparison of SWN and DC in terms of (A) usability and (B) versatility among users. A model trained by a subject's own MWL is used in (A), and accuracy is calculated using the subject's own MWL; the error bar shows the standard deviation of accuracy over the cross-validation set. Models trained by other subjects' MWL are used in (B), and accuracy is calculated using a subject's own MWL; the error bar shows the mean of the standard deviation of accuracy over the models trained with other subject's data.

trained using the MWL of subject 1 and cross-validated by calculating the accuracy for data of all other subjects. Fig. 2 shows these results, from which it is evident that the SWN accuracy increases with the window length for each subject. Further, some SWN accuracies calculated for other subjects' data (especially window length is 300 ms) are higher than those calculated for subject 1's own data. The DC accuracies that are calculated for other subjects' data are lower than those calculated for subject 1's own data as well as the chance level (100% / 3 targets = 33%). These results indicate that a model trained by DC can only fit one individual, while a model trained by SWN shows versatility among subjects.

B. Comparison between usability and versatility

To check the usability and versatility among users when using SWN and DC, we used the method described in Section II.F, i.e., we compared the accuracy and standard deviation of accuracy. The window length was set to 300 ms for SWN and 150 ms for DC, as mentioned in Section III.A (Fig. 2), to compare the best performances of the model. In DC, normalization is applied based on the maEMG in all data for each subject.

To evaluate the usability of SWN and DC, we compared the accuracy and standard deviation of accuracy across the cross-validation set using ten models trained by a subject's own data. Fig. 3(A) shows an example of a usability comparison using MWL. The MAV, MZC, and ES results were similar to those of MWL. Moreover, Fig. 3(A) indicates that SWN has higher usability than DC. There are two reasons for this: the first is that the mean accuracy over the cross-validation set is almost the same for SWN and DC at 57%; the second is that the mean of the SWN standard deviation of accuracy over the cross-validation set is better at 4.5% than that for DC, which is 0.83%.

Further, to evaluate the versatility among users of SWN and DC, we compared the accuracy and the mean of standard deviation of accuracy over the other subjects by using models trained by each of the other subjects' data. Fig. 3(B) shows an example of a versatility comparison using MWL. The error bars indicate the mean of the standard deviation of accuracy over the models trained with other subjects' data. The MAV, MZC, and ES results were similar to those of MWL. Fig. 3(B) indicates that SWN has higher versatility among users than DC. There are two reasons for this. The first is that the mean of SWN accuracy over the models trained with other subjects' data is 53%, which is 18% higher than for DC. The second reason is that the mean of the standard deviation for SWN accuracy over the models trained with other subjects' data is 0.88%, which is 4.6% lower than for DC.

Furthermore, we found that the mean SWN accuracy changes among models trained with data of different subjects. Fig. 4 shows an example of this result using subject 2's MWL. There is a maximum difference of approximately 7% between subject 2's MWL and the MWL of other subjects used to train the model. Fig. 4 indicates that the model performance changes with the quality of training data because the mean SWN accuracy obtained when using the model trained with subject 2's data is lower than that obtained when using models trained with other subjects' data.

IV. Discussion

The results for the usability and versatility among users (Fig. 3) indicate that SWN has beneficial characteristics: it ensures high performance without training the model and without calibration. Therefore, the proposed method can improve the usability and versatility among users of practical applications that use EMG and perform continuous classification, such as prosthetic hands.

In Fig. 2, the SWN accuracies increase with window length, and a sliding-window length of 300 ms is the best in this study. However, previous studies considered a window length between 100 and 200 ms for feature extraction. Therefore, in the case of SWN, the window length for normalization should be long (i.e., 300 ms) and the window length for feature extraction should be short (i.e., 100 and 150 ms); this may potentially improve accuracy. The best window length or window rate for each feature extraction in SWN must be investigated further.

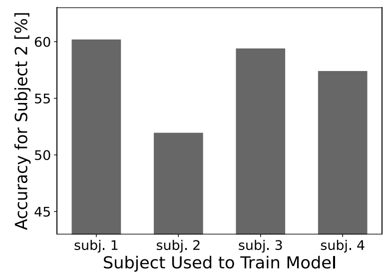

Figure 4. Comparison of mean SWN accuracy for subject 2 among models trained by each subject's MWL.

Fig. 4 indicates that the performance of the model changes with the quality of training data. This implies that a high-performance model can be realized by training with high-quality EMG data. Furthermore, high model performance may potentially be achieved by training a support vector machine, which can select the optimal data for classification, using mixed subjects' data.

V. Conclusions

Herein, we proposed a normalization method using a sliding window and z-scoring for continuous motion classification in order to improve the usability and versatility among users of EMG applications. The experimental results showed that the proposed method has a higher usability and versatility among users than DC.

However, the performance of the proposed method was not high. Therefore, it must be improved by eliminating noise and changing the window length or window rate for feature extraction. Additionally, we identified four other problems to address in future studies. First, we investigated only time-domain features in this study; therefore, we need to investigate other time-domain, frequency-domain, and advanced features. Second, z-scoring is essentially linear normalization, but EMG has strong nonlinearity; thus, nonlinear normalization must be achieved. Third, we have not investigated the versatility among individuals. Therefore, the effect of EMG changes that depend on sensor placement and fatigue rate must be investigated. Finally, as some studies estimated continuous kinematic/kinetic parameters such as joint angle and joint angular velocity, we need a normalization method for regression tasks that does not use the maximum absolute EMG.

References

[1] K. Li, J. Zhang, L. Wang, M. Zhang, J. Li, and S. Bao, "A review of the key technologies for sEMG-based human-robot interaction systems," *Biomedical Signal Processing and Control*, vol. 62, 102074, June 2020.

[2] B. Hudgins and P. Parker, "A new strategy for multifunction myoelectric control," *IEEE Transactions on Biomedical Engineering*, vol. 40, no. 1, pp. 82-94, Jan. 1993.

[3] P. Artameeyanant, S. Sultornsanee, and K. Chamnongthai, "An EMG-based feature extraction method using a normalized weight vertical visibility algorithm for myopathy and neuropathy detection," *SpringerPlus*, vol. 5, pp. 2101, 2016.