sEMG Data Expansion for Accurate Posture Classification

Minchang Sung, Sunhong Kim, Seulah Lee, and Youngjin Choi

Abstract—The paper presents methods to expand data acquired from a multi-channel sEMG fabric sensor for the dexterous control of robotic prosthesis. It is able to improve a variety of pattern recognition performance in spite of fewer data and less computational time. A multilayer perceptron (MLP) is utilized for the classification of eight postures in order to compare several methods regarding the data expansion such as data expanded with normal distribution (N-dist), data expanded with median operations, and data expansion with median plus normal distribution. Of the methods, an accuracy achieved using the data expanded with median plus normal distribution arrives at 99.32% as the highest, followed by the expansion using median, the expansion using normal distribution.

I. INTRODUCTION

The surface electromyography (sEMG) sensor has been developed to acquire muscular activations, especially, it has been widely used in the biomedical field. According as the number of channels of sEMG sensor has been increased, it is able to extract various information such as muscular coordination for the specific movement or posture of the limb. Since the multiple-channel sEMG signal shows different patterns for each specific posture, the postures can be recognized and classified from the patterns of sEMG signals. There are many application examples using these kinds of patterns, especially, they have been extended to various robotic fields because the robots can produce a variety of motions by putting the desired postures [1], [2], but it is difficult to recognize patterns accurately. Nowadays, many researchers have been trying to apply machine learning to the pattern recognition for more accurate classification, and its performance has been drastically improved thanks to deep learning scheme [3].

In our previous studies, sixteen-channel sEMG fabric sensor has been developed for the control of robotic prosthesis [4]. For eight postures illustrated in Fig. 1, sixteen-channel sEMG signal patterns were recorded from below-elbow amputee of right upper arm with 1,000Hz sampling frequency. Raw sixteen signals were sampled for one second per each posture, and thus the training data set of $16 \times 8000$ were obtained and then applied to the MLP for the posture classification. However, the number of samples was too deficient to train. Thus, three methods for data expansion are introduced to improve the classification performance. How to expand the data is explained in section II with discussions on the learning performance, and several future works are described in section III.

II. METHOD AND RESULT

To begin with, the data were filtered using RMS (root mean squaring) operation as pre-processing because there is an issue that white Gaussian noises of the signal are inevitably introduced [5]. The RMS operation reduces only the process noise collected from the sensor, and do not imply reducing the error caused by the postures. As a matter of fact, the sEMG data corresponding to the postures can be finely changed according to the variations of the sensor position. To alleviate this issue, we make the RMS filtered data be doubled using three methods as follows. First, small error following normal distribution is intentionally added to the RMS filtered data with the constant $\beta$ scaled normal distribution as follow:

$$x_{\text{new}} = x(n) + \beta \cdot N(0, 1), \quad (1)$$

where $x(n)$ implies the RMS filtered sEMG data and $N(0, 1)$ is a normal distribution with mean of 0 and variance of 1.

Second, the data of each channel have some correlation between adjacent samples because the muscle activations are conducted according to the time sequences of action potentials. Thus, the intermediate value is determined through the trapezoidal approximation, and it is inserted between two original data as follow:

$$x_{\text{new}} = \frac{x(n-1) + x(n)}{2}. \quad (2)$$

Third, above two methods are used simultaneously combined as follow:

$$x_{\text{new}} = \frac{x(n-1) + x(n)}{2} + \frac{1}{\gamma} N(0, 1). \quad (3)$$
The data expanded by using the aforementioned three methods and original data without any expansion are partially depicted as shown in Fig. 2.

Normally, if the data are lacking, the MLP experiences over-fitting issue due to a number of weight parameters. As an alternative, it is expected that the aforementioned data expansion makes the over-fitting issue be reduced. Fig. 3 shows a comparison of accuracies with respect to four cases: a) when the original data are used, b) when the data expanded with a normal distribution (N-dist) are used, c) when the data expanded using the median, d) when the data expanded with median plus normal distribution. The result of Fig. 3a) lets us know that the learning did not work well because it made just 75% of accuracy. However, Fig. 3b), c) and d) show us that the learning worked well. The accuracy of Fig. 3b) was 99.12%, c) was 99.02%, and d) was 99.32%. All the results reached over 99% when the data were expanded. The irregular changes of data brought better accuracy along with the data expansion.

III. FUTURE WORK

The paper proposed several methods to expand data acquired from the sixteen-channel sEMG fabric sensor for controlling the myoelectric prosthesis. When the data are lacking, the data expansion with normal distribution noise and median operation must be efficient in improving the accuracy. As future work, the data expansion will be applied to other algorithms such as the convolutional neural network (CNN).

REFERENCES