sEMG feature extraction using Generalized Discrete Orthonormal Stockwell Transform and Modified Multi-Dimensional Scaling

Somar Karheily, Ali Moukadem, Jean-Baptiste Courbot, Djaffar Ould Abdeslam
IRIMAS UR 7499, Université de Haute-Alsace, Mulhouse, France
somar.karheily@uha.fr

***Abstract*—This paper proposes a method based on a generalized version of the Discrete Orthonormal Stockwell Transform (GDOST) with Gaussian window to extract features from surface electromyography (sEMG) signals in order to identify hand's movements. The features space derived from the GDOST is then reduced by applying a modified Multi-Dimensional Scaling (MDS) method. The proposed modification on MDS consists in using a translation in kernel building instead of the direct distance calculation. The results are compared with another study applied on the same dataset where usual DOST and MDS are applied. We achieved significant improvements in classification accuracy, attaining 97.56% for 17 hand movements.**

***Index Terms*—sEMG classification, GDOST transform, MDS dimension reduction**

I. Introduction

The loss of a limb has a big impact on the life of a person, as it affects the ability of doing daily activities. The number of amputees in the world is estimated to be 57 million living people with a limb amputation due to a trauma [1]. Many research and work were done to develop prosthetic that replace some functionalities with many challenges, such as the degree of freedom (DOF), accuracy and design. The main parts in prosthetic are the acquisition unit, processing unit, and the moving parts. Currently, most of the prosthesis use the surface electromyography (sEMG) signals to identify the intended move of the amputee [3]. The sEMG signals are widely used in clinical upper-limb prosthetic control as they are noninvasive and relatively easy to measure. Typically, the amplitude of sEMG signals ranges from 50 µV to 10 mV and the frequency range is between 20 Hz and 500 Hz [4].

In this paper, we use a generalized version of the Discrete Orthonormal Stockwell Transform (DOST) that we refer to as GDOST to extract Time-Frequency (TF) domain features of the sEMG signals, and then we use the Multi-Dimensional Scaling (MDS) method to reduce the features dimension. We optimize the kernel calculations of this method, which significantly improves the classification accuracy measured by using a K Nearest Neighbor (kNN) classifier.

A. Related Work

Because sEMG signals are non-stationary, the use of TF features is more relevant as they preserve the variance of the frequency over the time which could be distinctive property of patterns in sEMG.

The related works in sEMG classification area show that TF features overcome the Time Domain (TD) features as we see by looking on different studies applied to the same data. In [5] different sets of TD features were applied on a 17-gestures-dataset and the average accuracy using kNN classifier were 85%. In a previous study, using the same dataset [10], we also proposed using TF methods and the results obtained exceeded 90% for two TF methods: Short-Time Fourier Transform (STFT) and Stockwell Transform (ST), exceeding thus the outcomes of TD features while working on the same dataset and classifier than in [5].

The STFT [7] is a TF transform which divides the signals into shorter segments of fixed length and then calculates the Fourier transform on each of it. STFT then determines the local frequency content of the signal which varies over the time. STFT was successfully used to extract sEMG features with good results. In [8] STFT with a Gaussian window was used on a dataset containing 17 movements and the average results of the classification was 92% using kNN and SVM classifiers. In [23], the STFT was used with the deep learning, where its spectrogram images have trained with 50-layer Convolutional Neural Network (CNN) based on Residual Networks (ResNet) architecture and the test accuracy was 99.59% for 7 different hand gesture.

Stockwell Transform (ST) improves the STFT by using a multi-resolution Gaussian window. Unlike the STFT, the Gaussian window in ST changes its width over frequencies. The ST is successfully used for features extraction for sEMG as in [9], where the classification accuracy of 6 different movements was 98.12%.

The calculation complexity in TF methods makes it hard to apply them in real prosthetic due to the required response time restrictions. We also applied the DOST in sEMG classification in our previous study [10] which was 10 times faster than STFT and ST but with a lower classification accuracy (88.08%, instead of 90.05% for STFT and 90.96% for ST).

The TF features yield a fine data representation, but lie in high dimension, making necessary the use of dimension reduction method. The most known method is the Principal Component Analysis (PCA) [11]. In [12] PCA was used with wavelet packet transform features and the classification accuracy for 9 hand movements was 96%. Different studies used also non-linear dimension reduction methods with TF features as

in [13] which compared PCA to Diffusion Maps (DM) and showed that DM outperforms PCA when less training data is used.

B. This Paper

The purpose of this paper is to improve the accuracy of DOST and MDS combination when classifying sEMG signals in order to take advantage of:

- the time-efficiency of DOST,
- the reduced amount of training data required when using non-linear dimension reduction methods.

These improvements were achieved by:

- Using GDOST with a Gaussian window instead of the classical DOST which is equivalent to apply a rectangular window.
- Improving the kernel of MDS to better present the dissimilarities between the observations.

In the remaining of the paper, section II details these improvements and the numerical experiments are reported in section III, while section IV concludes the study.

II. Methods

The sEMG signals we consider is recorded on several channels (electrodes), and each movement is repeated several times. As first step, we normalize this recorded signals by making the mean value is equal to 0 and the standard deviation is 1 on each channel. Then all signals are stored into one matrix S. The second step is to segment this data into observations, where each observation is the recorded sEMG signals on each channel in a specific time range. This yields m windows of signals for this observation, where m is the number of channels. For an observation X, $X = \{x_1, x_2, ..., x_m\}$, where $X \in \mathbb{R}^{m \times a}$, $x_i \in \mathbb{R}^a$ is the signal recorded on electrode i. and a is the number of samples in this signal.

A. DOST and GDOST

The DOST is the orthonormal version of the ST. It avoids redundancy in the time-frequency plane and paves the way to compute time-frequency representation in lower algorithmic complexity [16]. Let $f(t)$ be a signal $\in L^2([0,1])$, p is the number of the frequency bands, ν indicate the center of a frequency band, β indicates the width of the frequency band and τ for the time localization. The DOST coefficients $f_{p,\tau}$ can be calculated from an inner product between the signal $f(t)$ and the basis functions $D_{p,\tau}$:

$$f_{p,\tau} =< f, D_{p,\tau} > \tag{1}$$

where $D_{p,\tau}$ is given as [15]:

$$D_{p,\tau}(t) = \frac{1}{\sqrt{\beta(p)}} \sum_{f=\nu(p)-\beta(p)/2}^{\nu(p)+\beta(p)/2-1} e^{2\pi i f t} e^{-2\pi i f \tau/\beta(p)}, \quad t \in \mathbb{R} \tag{2}$$

The basis $D_{p,\tau}$ is not equivalent to the classical ST with Gaussian window. Indeed as R.G. Stockwell pointed in [16] this is equivalent to ST with boxcar window. In order to propose

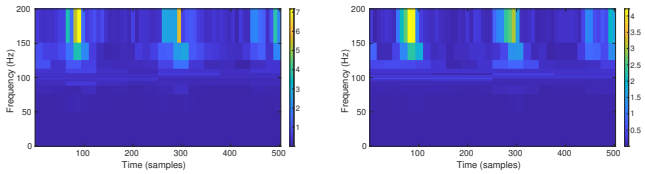


Fig. 1: DOST representation (left), and its GDOST counterpart of the same movement. Both are rearranged to yield a TF representation. In this figure and in the following, depictions cover a 250ms window length.

a generalized version of the DOST that allows to apply an admissible generalized window φ, authors in [20] propose the following basis for the GDOST:

$$E^{\varphi}_{p,\tau}(t) = \frac{1}{\sqrt{\beta(p)}} \sum_{j=0}^{\beta(p)-1} \left[c^{\varphi}_{p,j}(\nu(p))\right]^{-1} e^{2\pi i(\beta(p)+j)\left(t-\frac{\tau}{\beta(p)}\right)} \tag{3}$$

For the special case $\varphi = \check{\chi}$ which corresponds to boxcar window, the authors [20] proved that :

$$E^{\check{\chi}}_{p,\tau} = D_{p,\tau} \tag{4}$$

In this paper, we propose to use a Gaussian window as originally used in the ST with $\sigma = 0.1$. In this case, $\varphi = g(t)$ which can be given as:

$$g(t) = \frac{1}{\sigma\sqrt{2\pi}} e^{\frac{-t^2}{2\sigma^2}}. \tag{5}$$

An example of DOST and GDOST on the same signal is given in Fig. 1.

B. Feature extraction from GDOST

For an observation X recorded on m channels: $X = \{x_1, x_2, ..., x_m\}$. The GDOST transform of this observation is represented as:

$$GDOST(X) = \{F_{x_1}, F_{x_2}, ..., F_{x_m}\}, \tag{6}$$

where F_{x_i} is the GDOST transform of x_i.
The GDOST yields a number of features equals to the number of the samples in the signal, so that $\forall i \in \{1, ..., m\}, F_{x_i} \in \mathbb{R}^a$, This means that the features number of observation X on all channels will be $K = m \times a$.
We extract the GDOST features for every observation in the dataset S, so that the features matrix of the all observations in the dataset is $F \in \mathbb{R}^{N \times K}$, where N is the overall existing observations in the dataset.

C. Multidimensional Scaling (MDS)

MDS [18] replaces the observations' features by a new set of features measuring the dissimilarities between each pair of observations [17]. The steps of MDS are [18]:

- From the features matrix F, we calculate the paired distances which will be the MDS kernel matrix D. For two movements X and Y, their features are represented

as: $\{F_{x_1}, F_{x_2}, ..., F_{x_m}\}$, $\{F_{y_1}, F_{y_2}, ..., F_{y_m}\}$ The paired distances between these two observations (X, Y) will be calculated as:

$$d_{xy} = \sum_{i=1}^{m} \|F_{x_i} - F_{y_i}\|_2 \tag{7}$$

- From D, we compute the distance matrix A, which represents the Euclidean distances between the observations, so that $\forall i, j \in \{1, \ldots, N\}$: $a_{ij} = -\frac{1}{2} d_{ij}^2$.
- Then we apply double centering: $B = HAH$, where H is the centering matrix, *i.e.* $H = I - \frac{1}{N} 1_N 1_N^T$ with 1_N a vector of ones of size N.
- A spectral decomposition of B is performed, such that $B = \mathcal{V} \Lambda \mathcal{V}^T$, where Λ is the diagonal matrix which has eigenvalues of B, and $\mathcal{V}$ contains the eigen-vectors.
- Finally, the embedded features matrix Z is:

$$Z = \mathcal{V}_q \Lambda_q^{\frac{1}{2}} \tag{8}$$

where q is the number of embedded features. We choose $q \ll K$, so that the embedded features still have the intrinsic characteristics, and the less-related or redundant features are removed.

D. Distance synchronization

The Euclidean distances in MDS are sensitive to the time shift in GDOST features, which leads to erroneous dissimilarity calculations (Eq. 7), while better dissimilarity calculations should consider the relative positions of the energy in the time-frequency plane. To achieve that when calculating the distance between X and Y, we shift features of Y step-by-step and calculate the distance after each shifting.

Let the operator $\mathcal{T}^j F$ be the circular shift with step $j \in \mathbb{Z}$ on the sequence of features F. Then, the distance between X and Y is taken as the minimal distance between Y and shifted versions of X. The proposed operation can be expressed as follow:

$$d_{xy} = \min_{j \in \{0, \ldots, a-1\}} \sum_{i=1}^{m} \|(\mathcal{T}^j F_{x_i}) - F_{y_i}\|_2 \tag{9}$$

The final distance matrix is $D \in \mathbb{R}^{N \times N}$, will contains all paired distances d_{xy}.

This optimization improves the dissimilarities calculations between pairs making them shift-invariant.

E. Main Algorithm

The main steps of our algorithm are the following:

- *Normalization*, so that the raw data have a 0-mean and a unitary standard deviation.
- *Data segmentation*. We divide the data on each channel into fixed-size windows. An observation of a movement consists of m windows (one window on each channel): $X = \{x_1, x_2, ..., x_m\}$ where $x_m \in \mathbb{R}^a$.
- *TF features extraction*. For each observation, we calculate the GDOST transform. The output of this step (Eq. (6)) will contains features of all observations represented as $F \in \mathbb{R}^{N \times K}$, where N is the number of all observations in the dataset, and K is the total number of TF features for each observation: $K = m \times a$.
- *Dimension reduction.* By applying the MDS with our enhancement on the kernel calculations (Eq. (9)), we reduce the dimension of $F \in \mathbb{R}^{N \times K}$ from previous step, into $Z \in \mathbb{R}^{N \times q}$ where $q \ll K$.
- *Classification*. We use k-fold cross-validation and kNN as a classifier to evaluate our methods.

In Fig. 2, we can see the main steps of our work.

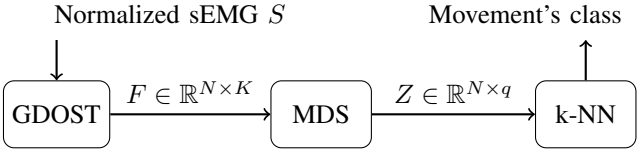


Fig. 2: Main algorithm overview.

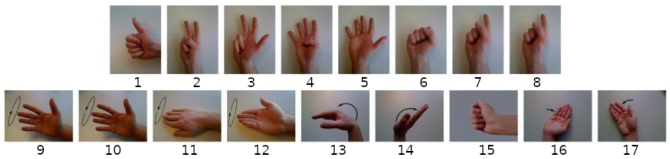


Fig. 3: Visual depiction of the 17 hand gestures considered in our study, based on the NinaPro database [19]

III. Numerical experiment

A. Data

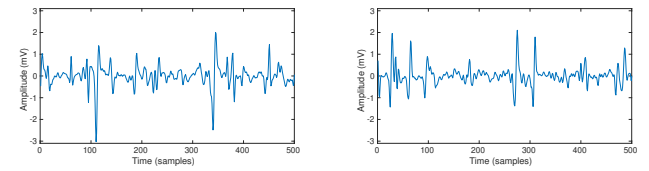
We use the public database from the Ninapro Project [19] as it provides many kind of movements recorded on many different subjects. These signals are saved in unified format that simplifies the testing on different subjects and different datasets. We apply our methods on the dataset 2 - exercise 1, which contains raw sEMG data for 17 different basic movements of fingers and wrist, depicted in Fig. 3.

The subject is asked to perform the movement for 5 seconds followed by 3 seconds of rest, then repeating that 6 times. The sampling rate is 2 kHz and the number of electrodes is $m = 12$. Fig 3 shows the list of these movements. For segmentation, we chose the window length equal to 250 ms with a 125 ms overlap between windows.

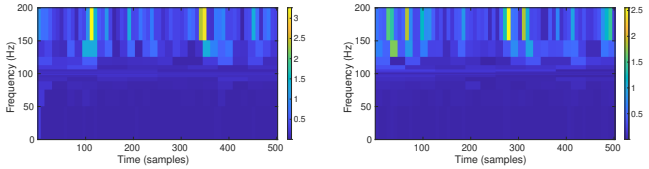
B. Feature extraction

The features we used in our work is the TF features calculated using GDOST. As we have 12 channels in our used data, that leads to 12 signals for each observation; therefore The GDOST features are computed and combined from all these signals. Fig. 4 shows the GDOST transform of two different movements on the same channel.

The window length we used is 250ms, and sampling rate in the used database is 2kHz, which means we have $a = 500$ samples for each window on each channel, and we have $m = 12$ the number of channels; therefore an observation X will have

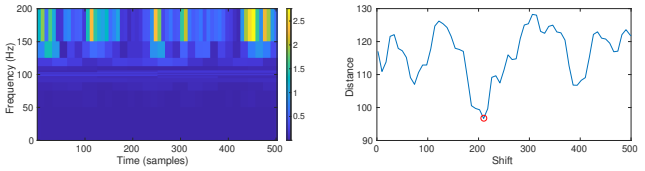


(a) Move 3, channel 1 (b) Move 7, channel 1

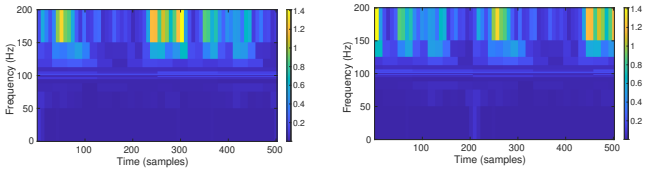


(c) GDOST of move 3, channel 1 (d) GDOST of move 7, channel 1

Fig. 4: GDOST transform example for two different movements (3 and 7, see Fig. 3) acquired on the same channel.



(a) $GDOST(X_0)$ (b) Distance by shift value



(c) $GDOST(X_1)$ (d) $\mathcal{T}^{210}(GDOST(X_1))$

Fig. 5: The figure shows GDOST(X_0) in (a). (c) shows GDOST(X_1) without any shifting and (d) shows GDOST(X_1) with circle-shifting by value 210. (b) shows the distances between GDOST(X_0) and GDOST(X_1) while shifting X_1.

$500 \times 12 = 6000$ samples in total. Applying GDOST on this observation on all channels will give $K = 6000$ features as TF features that represent this movement. The total number of features for each movement is high and needs to be reduced to remove redundancy and avoid classification over-fitting.

C. Dimension Reduction

We apply distances' synchronization between observations pairs to get better dissimilarity representations. The Fig. 5 shows GDOST representations of two different observations of the movement 9 recorded on channel 1. In Fig. 5a, and Fig. 5c we notice that calculating the distance directly between these two observations (Eq. 7) would lead to high dissimilarity as shown in Fig. 5b, while distance between GDOST in Fig. 5a and Fig. 5d will be the minimum value of distance (Eq. 9) between these observations and actually reflects the fact that they both represent the same movement. Fig. 5b shows how the distance between these two observations differs while performing circle-shift on the time resolution. We can notice that the minimum distance happens with shift value equal to 210 time-sample, as shown in Fig. 5b as a red point, and the corresponding GDOST shifted transform could be seen in Fig. 5d.

We choose the number of feature q (see Eq. 8) that yields the maximum classification accuracy on our database over the range $[10, 400]$, which is $q = 191$. The classification accuracy will start to drop as adding more features will just increases the classifier's over-fitting.

After we finish calculating the synchronized paired-distances between each pair of observations, we get the kernel matrix of MDS, which contains the new features of each observation as a vector of distances from other observations.

In order to evaluate our features extraction methods and dimension reduction approach, we used kNN classifier with $k = 3$, with a 5-Fold cross-validation.

Methods	**Accuracy**	**Reference**
TD features set, kNN	85%	[5]
RMS, Median Frequency, Para-consistent artificial neural network	76% $\pm$ 9.1%	[6]
STFT, SVM, kNN	92%	[21]
ST, PCA	90.96%	[10]
DOST, MDS	87.13%	[10]
DOST, Enhanced MDS	96.73%	this study
GDOST, Enhanced MDS	97.56%	this study

TABLE I: The classification accuracy of different feature extraction and dimension reduction combinations done on the same database and same movements. The table shows the significant improvement of using GDOST and enhanced MDS.

D. Results and Discussions

We applied our methods (Fig. 2) on 5 different subjects. The average accuracy achieved was 97.56%. We chose the same test subjects as in study [10] with the same classification methods in order to perform an accurate comparison and evaluation for our improvements. In [10], DOST with MDS gave 87.13% accuracy, which means by using enhanced MDS kernel and GDOST, we were able to increase the accuracy by 10%. Comparing with the best results obtained on the same database in [10] which was achieved by using ST with PCA with accuracy 90.96%, we notice that our methods improved the accuracy by 6.6%. By comparing DOST and GDOST when both are applied with enhanced MDS, we see that using the GDOST led to improving the accuracy from 96.73% to 97.56%. The significant improvement obtained by these combination (GDOST and enhanced MDS) came from the MDS kernel optimization, so the distances were calculated in a way that actually measure the dissimilarity between observations, besides to using Gaussian window instead of rectangle window in DOST, which improved the TF features of the observations. Table I summarize these results, compared with other studies made on the same database and with the same number of movements.

IV. Conclusion

In the framework of sEMG classification, we improved the accuracy of DOST for feature extraction, and of MDS as a non-linear dimension reduction method. The importance of this comes from the fact that DOST is a time-efficient TF transform, besides to improving MDS kernel calculations to be more suitable to the TF features. Further works will generalize these improvements to other non-linear dimension reduction methods, promoting their use as they perform better over smaller training sets, and they also better presents the cross-subject features [13]. Also, the parameter of the Gaussian window of the GDOST can be optimized based on a criterion equivalent to the energy concentration measurement [22]. This might further improve the quality of the TF features.

References

[1] McDonald, Cody L1,; Westcott-McCoy, Sarah1; Weaver, Marcia R2; Haagsma, Juanita3; Kartin, Deborah1 Global prevalence of traumatic non-fatal limb amputation, Prosthetics and Orthotics International: April 2021 - Volume 45 - Issue 2 - p 105-114 doi: 10.1177/0309364620972258.

[2] Farina, D., Ning Jiang, Rehbaum, H., Holobar, A., Graimann, B., Dietl, H., & Aszmann, O. C. (2014). The Extraction of Neural Information from the Surface EMG for the Control of Upper-Limb Prostheses: Emerging Avenues and Challenges. IEEE Transactions on Neural Systems and Rehabilitation Engineering, 22(4), 797–809. doi:10.1109/tnsre.2014.2305111

[3] Parajuli N, Sreenivasan N, Bifulco P, et al. Real-Time EMG Based Pattern Recognition Control for Hand Prostheses: A Review on Existing Methods, Challenges and Future Implementation. Sensors (Basel). 2019;19(20):4596. Published 2019 Oct 22. doi:10.3390/s19204596

[4] N. Meselmani et al. "Pattern recognition of EMG signals: Towards adaptive control of robotic arms". In: 2016 IEEE International Multidisciplinary Conference on Engineering Technology (IMCET). 2016, pp. 5257. doi: 10.1109/IMCET.2016.7777426.

[5] C. P. Robinson, B. Li, Q. Meng, M. T. Pain, Pattern classification of hand movements using time domain features of electromyography, in: Proceedings of the 4th International Conference on Movement Computing, 2017, pp. 1–6.

[6] G. W. Favieiro, K. O. A. Moura, A. Balbinot, Novel method to characterize upper-limb movements based on paraconsistent logic and myoelectric signals, in: 2016 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), 2016, pp. 395–398. doi:10.1109/EMBC.2016.7590723.

[7] E. Sejdic, I. Djurovic, J. Jiang, Time–frequency feature representation using energy concentration: An overview of recent advances, Digital Signal Processing 19 (2009) 153–183. doi:10.1016/j.dsp.2007.12.004.

[8] Too, J. et al. "Application of gabor transform in the classification of myoelectric signal." TELKOMNIKA Telecommunication Computing Electronics and Control 17 (2019): 873-881.

[9] Z. Zhu, Tian, Wang, Yokoi, C.-W. Huang, Semg feature extraction based on stockwell transform improves hand movement recognition accuracy, Sensors 19 (2019) 4457. doi:10.3390/s19204457.

[10] Somar Karheily, Ali Moukadem, Jean-Baptiste Courbot, Djaffar Ould Abdeslam. sEMG Time-Frequency Features For Hand Movements Classification. 2021. ⟨hal-03223176⟩

[11] I. Jolliffe, Springer-Verlag, Principal Component Analysis, Springer Series in Statistics, Springer, 2002.

[12] J. . Chu, I. Moon, M. . Mun, A real-time EMG pattern recognition system based on linear-nonlinear feature projection for a multifunction myoelectric hand, IEEE Transactions on Engineering 53 (11) (2006) 2232–2239. doi:10.1109/TBME.2006.883695.

[13] N. Rabin, M. Kahlon, S. Malayev, A. Ratnovsky, Classification of human hand movements based on EMG signals using nonlinear dimensionality reduction and data fusion techniques, Expert Syst. Appl. 149 (2020)113281.

[14] J. J. A. Mendes Junior, M. L. Freitas, H. V. Siqueira, A. E. Lazzaretti, S. F. Pichorim, S. L. Stevan, Feature selection and dimensionality reduction: An extensive comparison in hand gesture classification by sEMG in eight channels armband approach, Signal Processing and Control 59 (2020) 101920. doi:https://doi.org/10.1016/j.bspc.2020.101920. URL http://www.sciencedirect.com science/article/pii/ S1746809420300768

[15] M. Jaya Bharata Reddy, R. K. Raghupathy, K. Venkatesh, D. Mohanta, Power quality analysis using discrete orthogonal s-transform (dost), Digital Signal Processing 23 (2) (2013) 616626. doi:https://doi.org/10.1016/j.dsp.2012.09.013. URL https://www.sciencedirect.com/science/article/pii/S1051200412002321

[16] R. Stockwell, A basis for efficient representation of the s-transform, Digital Signal Processing 17 (1) (2007) 371 – 393. doi:https://doi.org/10.1016/j.dsp.2006.04.006.

[17] J. Friedman, T. Hastie, R. Tibshirani, et al., The elements of statistical learning, Vol. 1, Springer series in statistics New York, 2001.

[18] M. A. Cox, T. F. Cox, Multidimensional scaling, in: Handbook of data visualization, Springer, 2008, pp. 315–347.

[19] M. Atzori, A. Gijsberts, C. Castellini, B. Caputo, A.-G. M. Hager, S. Elsig, G. Giatsidis, F. Bassetto, H. Müller, Electromyography data for noninvasive naturally-controlled robotic hand prostheses, Scientific data 1 (1)(2014) 1–13.

[20] Battisti, Ubertino, and Luigi Riba. "Window-dependent bases for efficient representations of the Stockwell transform." Applied and Computational Harmonic Analysis 40.2 (2016): 292-320.

[21] Too, J., Abdullah, A. R., Saad, N. M., Ali, N. M., & Zawawi, T. T. (2019). Application of gabor transform in the classification of myoelectric signal. Telkomnika, 17(2), 873-881.

[22] A. Moukadem, Z. Bouguila, D.O Abdeslam and A. Dieterlen, A new optimized Stockwell transform applied on synthetic and real non-stationary signals, Digital Signal Processing, 2015, 226-238.

[23] M. A. Ozdemir, D. H. Kisa, O. Guren, A. Onan and A. Akan, "EMG based Hand Gesture Recognition using Deep Learning," 2020 Medical Technologies Congress (TIPTEKNO), 2020, pp. 1-4, doi: 10.1109/TIPTEKNO50054.2020.9299264.