

RESEARCH ARTICLE

Extracting time-frequency feature of single-channel vastus medialis EMG signals for knee exercise pattern recognition

Yi Zhang^{1,2,3}*, Peiyang Li^{2,3}, Xuyang Zhu^{2,3}, Steven W. Su⁴*, Qing Guo¹, Peng Xu^{2,3}*, Dezhong Yao^{2,3}

1 School of Aeronautics and Astronautics, University of Electronic Science and Technology of China, 611731, Chengdu, China, **2** Key Laboratory for NeuroInformation of Ministry of Education, School of Life Science and Technology, University of Electronic Science and Technology of China, 610054, Chengdu, China, **3** Center for Information in BioMedicine, University of Electronic Science and Technology of China, 610054, Chengdu, China, **4** Centre for Health Technologies, Faculty of Engineering and Information Technology, University of Technology, Sydney, NSW 2007, Australia

* These authors contributed equally to this work.

* yi.zhang@uestc.edu.cn (YZ); steven.su@uts.edu.au (SS); xupeng@uestc.edu.cn (PX)



OPEN ACCESS

Citation: Zhang Y, Li P, Zhu X, Su SW, Guo Q, Xu P, et al. (2017) Extracting time-frequency feature of single-channel vastus medialis EMG signals for knee exercise pattern recognition. PLoS ONE 12(7): e0180526. <https://doi.org/10.1371/journal.pone.0180526>

Editor: Yih-Kuen Jan, University of Illinois at Urbana-Champaign, UNITED STATES

Received: January 4, 2017

Accepted: June 17, 2017

Published: July 10, 2017

Copyright: © 2017 Zhang et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: All relevant data are within the paper and its Supporting Information files.

Funding: This study was supported by the following fundings: Fundamental Research Funds for the Central Universities, ZYGX2015J118, <http://www.uestc.edu.cn/>; National Natural Science Foundation, 61522105, <http://www.nsf.gov.cn/>; National Natural Science Foundation, 51675087, <http://www.nsf.gov.cn/>; and the China Postdoctoral Science Foundation funded project,

Abstract

The EMG signal indicates the electrophysiological response to daily living of activities, particularly to lower-limb knee exercises. Literature reports have shown numerous benefits of the Wavelet analysis in EMG feature extraction for pattern recognition. However, its application to typical knee exercises when using only a single EMG channel is limited. In this study, three types of knee exercises, i.e., flexion of the leg up (standing), hip extension from a sitting position (sitting) and gait (walking) are investigated from 14 healthy untrained subjects, while EMG signals from the muscle group of vastus medialis and the goniometer on the knee joint of the detected leg are synchronously monitored and recorded. Four types of lower-limb motions including standing, sitting, stance phase of walking, and swing phase of walking, are segmented. The Wavelet Transform (WT) based Singular Value Decomposition (SVD) approach is proposed for the classification of four lower-limb motions using a single-channel EMG signal from the muscle group of vastus medialis. Based on lower-limb motions from all subjects, the combination of five-level wavelet decomposition and SVD is used to comprise the feature vector. The Support Vector Machine (SVM) is then configured to build a multiple-subject classifier for which the subject independent accuracy will be given across all subjects for the classification of four types of lower-limb motions. In order to effectively indicate the classification performance, EMG features from time-domain (e.g., Mean Absolute Value (MAV), Root-Mean-Square (RMS), integrated EMG (iEMG), Zero Crossing (ZC)) and frequency-domain (e.g., Mean Frequency (MNF) and Median Frequency (MDF)) are also used to classify lower-limb motions. The five-fold cross validation is performed and it repeats fifty times in order to acquire the robust subject independent accuracy. Results show that the proposed WT-based SVD approach has the classification accuracy of 91.85% ±0.88% which outperforms other feature models.

2017M612950. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing interests: The authors have declared that no competing interests exist.

1 Introduction

The lower-limb motion is of vital importance in human daily living activities. It is frequently difficult for physically weak persons (i.e., elderly, disabled, and injured persons) to perform daily lower-limb activities. Recently, power-assisted robotic systems have been developed to target these problems and assist people who need help in terms of their daily living [1–6]. Most of these systems are originally activated by EMG signals as they can directly indicate the electrophysiological responses to daily living activities. Therefore, the study of EMG-based biofeedback and its relationship to lower-limb motions is significantly concerned with the control of power-assisted robotic systems [7–10].

A motor unit usually includes a motor neuron and the skeletal muscle fibres innervated by that motor neuron. The force of a muscle contraction is regulated by the number of activated motor units. The measured single-channel EMG signal is a superimposed electrical activity with the sum of activations from multiple motor units and a wide range of noise [11] [12]. For an exercise motion, the physiological behaviours from the brain activity are intrinsic [13]. In addition, the electric indications of the activated motor units for one muscle contraction should be unique, and only vary due to brain activity. The single-channel EMG signal relative to one particular muscle contraction is used to interpret the electrophysiological response from one perspective [14, 15]. In this study, a single-channel-based EMG pattern classification is developed in order to provide an easy-to-use condition for detections of lower-limb motions.

As described in detail previously [10], Agonist-antagonist muscles exist in human joints e.g., elbow, hip, wrist, knee, and ankle. These human joints are usually activated by either the biarticular or uniaxial muscles. For instance, a flexion/extension motion of human knee joint is actuated by the muscles of biceps femoris, semitendinosus, gastrocnemius, rectus femoris, vastus lateralis and vastus medialis. Most of these muscles are biarticular muscles except the uniaxial muscles of vastus lateralis and vastus medialis. In this study, the muscle of vastus medialis is used since it is primarily associated with muscle fatigue and has the significance for the human-machine interface [16].

The EMG signal is contaminated with various artifacts, and its time series data is not practical for classification purposes [17]. The conventional methods are to apply the time-dependent and the frequency-dependent properties to detect the EMG features across different knee exercise patterns [10, 18, 19], such as Mean Absolute Value (MAV), Root-Mean-Square (RMS), integrated EMG (iEMG), Zero Crossing (ZC), Mean Frequency (MNF) and Median Frequency (MDF). Since EMG signals vary in frequency content over time, the conventional analysis methods cannot accurately describe its time-dependent statistical properties. In order to alleviate this problem, the following EMG feature extraction models of upper-limb activities (specific to wrist motions) have been developed including Fourier Transform (FT), Wavelet Transform (WT), Autoregressive (AR), Power Spectral Estimation (PSE), and Smooth Localised Complex Exponential (SLEX) [17, 20–23]. The recent review work in [17], where the classification accuracy rate of these feature vectors was evaluated by an Artificial Neural Network (ANN) classifier, clearly show that the time-frequency analysis methods have high accuracy rates, while the WT is much more efficient than the FT, as well as being better on the localisation of dominant subfrequency bands. In addition to Electroencephalogram (EEG) signals, it has been reported in previous study [24] that WT was employed to handle the EEGs and form a matrix, and then the singular value of the matrix is extracted. However, this method was limited to EMG-based classifications for lower limb activities.

The objective of this study is to propose the WT-based Singular Value Decomposition (SVD) approach for classification of EMG-based lower limb activities using a single-channel

EMG signal from the muscle group of vastus medialis. Although WT-based approaches have been widely used for EMG signals, the combination of WT and SVD to comprise the feature vector for classifications of lower limb activities, are not well documented. In addition, the classification of those motions using only one single EMG channel is also nontrivial.

The experiment for data collections was approved by the University of Technology Sydney (UTS) Human Research Ethics Committee. An informed consent was obtained prior to data collection. The University of Electronic Science and Technology of China (UESTC) also approved to use the databases obtained from UTS for publications. The databases of 14 healthy untrained subjects are taken into the study. The subjects undergo three exercise programs associated with the knee joint, leg extension from a sitting position (sitting), flexion of the leg up (standing), and gait (walking), while EMG signals from the vastus medialis muscle of the detected leg are monitored and recorded. Based on the EMG datasets, four types of lower-limb motions including standing, sitting, stance phase of walking, and swing phase of walking, are segmented.

The Support Vector Machine (SVM) is then configured to build a multiple-subject classifier for which the subject independent accuracy will be given across all subjects for the classification of four types of lower-limb motions. In order to effectively indicate the classification performance, EMG features from time-domain and frequency-domain are also used to classify lower-limb motions. The five-fold cross validation is performed and it repeats fifty times in order to acquire the robust subject independent accuracy. By comparing outcomes from typical feature models with the one from the proposed WT-based SVD approach, the WT-based SVD approach provide best subject independent accuracy for standing, sitting, stance phase of walking, and swing phase of walking.

The remainder of the paper is organised as follows. Section ‘Method’ describes WT, SVD, EMG features, classification and evaluation procedure. Section ‘Experiment’ introduces experimental equipment, exercise procedures and protocols. Section ‘Results’ provides the outcomes from time-domain features, frequency-domain features, as well as the SVD-WT approach. Section ‘Discussions and Limitations’ clarifies the limitations and future directions of this study. Section ‘Conclusion’ concludes this study.

2 Methods

The University of Technology Sydney (UTS) Human Research Ethics Committee (UTS HREC 2009000227) approved this experiment and collected the EMG datasets. The University of Electronic Science and Technology of China (UESTC) also approved to use the databases obtained from UTS for publications. A written informed consent was obtained from every subject prior to the experiment. The University of Technology Sydney (UTS) Human Research Ethics Committee (UTS HREC 2009000227) approved this consent procedure.

2.1 Time-domain features

The Mean Absolute Value (MAV) is used to calculate the average of the absolute value of all time samples, which is essential for determining the baseline of single EMG channel during daily lower-limb motions. The equation is given by Eq (1)

$$\bar{X} = \frac{1}{N} \sum_{k=1}^N |x_k|, \quad (1)$$

where x_k is the potential at the k^{th} sampling and the parameter N is the number of samples [25].

The Root-Mean-Square (RMS) is also employed as a feature extraction method of the EMG signals in this study. The equation of RMS is written as Eq (2)

$$RMS_i = \sqrt{\frac{1}{N} \sum_{k=1}^N x_k^2}, \text{ for } i = 1, \dots, I \tag{2}$$

where I is the number of segments.

The integrated EMG (iEMG) [26] is the mathematical integral of the absolute value of EMG normalized candidate segments, where can be expressed as Eq (3)

$$iEMG_i = \sum_{k=1}^N |x_k|, \text{ for } i = 1, \dots, I. \tag{3}$$

Zero Crossing (ZC) is a feature that tracks the number of times the waveform crosses zero, switching from a positive signal to a negative one, and vice versa [27], which was combined with the tone labels to detect the onset of movement during the procedure of data segmentation. Eq (4) is used for zero crossings

$$\begin{cases} \text{sgn}(-x_k \times x_{k+1}) \\ |x_k - x_{k+1}| \geq \text{threshold} \end{cases} \tag{4}$$

where the sgn function returns a positive number if the threshold is exceeded, and the zero crossing count is incremented [27, 28].

2.2 Frequency-domain features

Mean Frequency (MNF) [29] is the average frequency of an EMG segment representing the lower-limb motion. MNF is calculated as the sum of the product of the amplitude spectrum and the frequency, divided by the total sum of spectrum intensity. The mathematic expression can be written as Eq (5)

$$MNF_i = \frac{\sum_{j=1}^M P_{ij} f_{ij}}{\sum_{j=1}^M P_{ij}}, \text{ for } i = 1, \dots, I \tag{5}$$

where P_{ij} is the EMG power spectrum at the frequency bin j for the i th segment, f_{ij} is the frequency value of EMG power spectrum at the frequency bin j for the i th segment, and M is the length of frequency window.

Median Frequency (MDF) is the frequency at which the spectrum is divided into two regions with equal amplitude. It can be calculated in the following two steps: [29]

a). The signal intensities in the whole spectrum are summed and then divided by two (see Eq (6)),

$$MDF_i = \frac{1}{2} \sum_{j=1}^M P_{ij}, \text{ for } i = 1, \dots, I. \tag{6}$$

b). The MDF frequency is selected at which the cumulative intensity first exceeds MDF_i .

2.3 Five-level wavelet decomposition

The frequency of human EMG signal is usually between 10–500 Hz [30, 31]. WT is a time-frequency tool to analyse EMG signals. It can decompose signals into different scales and provide more information from time and frequency domains. In this paper, a five-level wavelet

decomposition is employed by using Daubechies4 algorithm [32]. Based on the decomposition of EMG signals, the restricted frequency bandwidth for the components from cD1, cD2, cD3, cD4, cA5 are 256 ~ 512 Hz, 128 ~ 256 Hz, 64 ~ 128 Hz, 32 ~ 64 Hz, and 16 ~ 32 Hz, respectively. cAn ($n = 1, 2, \dots, 5$) is the coefficients for low-frequency components of the signals. cDn is the coefficients for the high-frequency components. At each level, only one parameter from cDn can be retained according to the SVD method which can compress cDn to one parameter.

2.3.1 Singular Value Decomposition (SVD). According to Section 2.3, the EMG data of four lower-limb motions above are decomposed by WT, and feature matrices ($M_{ij} = [cD_{1i}; cD_{2i}; cD_{3i}; cD_{4i}; cD_{5i}; cA_{5i}]$, $i = 1, 2, \dots, I$) will be built. In order to pursue the simple and effective feature vector for the following classification by SVM, the singular value of M_{ij} will be obtained by SVD.

Let $A, B \in C^{m \times n}$, if there are m order unitary matrix U and n order unitary matrix V , $U^T AV = B$, then A and B are unitary equivalent. If $A \in C_r^{m \times n}$ ($r > 0$), the eigenvalues of $A^T A$ should have the following relation equation as Eq (7)

$$\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_r \geq \lambda_{r+1} = \dots = \lambda_n = 0, \tag{7}$$

then $\sigma_i = \sqrt{\lambda_i}$ ($i = 1, 2, \dots, n$) are the singular values of A .

If $A \in C_r^{m \times n}$ ($r > 0$), then there must be m order unitary matrix U and n order unitary matrix V , they satisfy Eq (8)

$$U^T AV = \begin{bmatrix} \Sigma & 0 \\ 0 & 0 \end{bmatrix} \tag{8}$$

where $\Sigma = \text{diag}(\sigma_1, \sigma_2, \dots, \sigma_r)$, and σ_i ($i = 1, 2, \dots, r$) are the nonzero singular values of A . Eq (8) can be transformed to Eqs (9), (10) and (11)

$$A = U \begin{bmatrix} \Sigma & 0 \\ 0 & 0 \end{bmatrix} V^T \tag{9}$$

where

$$U = AA^T \tag{10}$$

$$V = A^T A \tag{11}$$

Eq (9) is called the SVD of A .

2.4 Support vector regression

Let $\{u_i, y_i\}_{i=1}^N$ be a set of inputs and outputs data points ($u_i \in U \subseteq R^d$, $y_i \in Y \subseteq R$, N is the number of points). The goal of the support vector regression is to find a function $f(u)$ which has the following form

$$f(u) = w \cdot \phi(u) + b, \tag{12}$$

where $\phi(u)$ is the high-dimensional feature spaces which are nonlinearly transformed from u . The weight vector w and bias b are defined as the hyperplane by the equation $\langle w \cdot \phi(u) \rangle + b = 0$.

The hyperplane is estimated by minimizing the regularized risk function

$$\frac{1}{2} \|w\|^2 + C \frac{1}{N} \sum_{i=1}^N L_\epsilon(y_i, f(u_i)), \tag{13}$$

The first term is called the regularized term. The second term is the empirical error measured by ϵ -insensitivity loss function which is defined as

$$L_\epsilon(y_i, f(u_i)) = \begin{cases} |y_i - f(u_i)| - \epsilon, & |y_i - f(u_i)| > \epsilon \\ 0, & |y_i - f(u_i)| \leq \epsilon \end{cases} \tag{14}$$

This defines an ϵ tube. The radius ϵ of the tube and the regularization constant C are both determined by user.

The selection of parameter C depends on application knowledge of the domain. Theoretically, a small value of C will under-fit the training data because the weight placed on the training data is too small, thus resulting in large values of MSE (mean square error) on the test sets.

However, when C is too large, SVR will over-fit the training set so that $\frac{1}{2} \|w\|^2$ will lose its meaning and the objective goes back to minimize the empirical risk only. Parameter ϵ controls the width of the ϵ -insensitive zone. Generally, the larger the ϵ the fewer number of support vectors and thus the sparser the representation of the solution. However, if the ϵ is too large, it can deteriorate the accuracy on the training data.

By solving the above constrained optimization problem, we have

$$f(u) = \sum_{i=1}^N \beta_i \phi(u_i) \cdot \phi(u_i) + b, \tag{15}$$

As mentioned above, by the use of kernels, all necessary computations can be performed directly in the input space, without having to compute the map $\phi(u)$ explicitly. After introducing kernel function $k(u_i, u_j)$, the above equation can be rewritten as follows

$$f(u) = \sum_{i=1}^N \beta_i k(u_i, u) + b, \tag{16}$$

where the coefficients β_i corresponding to each (u_i, y_i) . The support vectors are the input vectors u_j whose corresponding coefficients $\beta_j \neq 0$. For linear support regression, the kernel function is thus the inner product in the input space

$$f(u) = \sum_{i=1}^N \beta_i \langle u_i, u \rangle + b, \tag{17}$$

For nonlinear SVR, there are a number of kernel functions which have been found to provide good generalization capabilities, such as polynomials, radial basis function (RBF), sigmoid.

Here we present the polynomials and RBF kernel functions as follows:

Polynomial kernel: $k(u, u') = ((u \cdot u') + h)^p$.

RBF Kernel: $k(u, u') = \exp\left(-\frac{\|u - u'\|^2}{2\sigma^2}\right)$.

Details about SVR, such as the selection of radius ϵ of the tube, kernel function, and the regularization constant C , can be found in [33–36]. The origin of the SVM code applied in this study was obtained from [37].

2.5 Classification

For the classification of four knee exercises, the singular values of these matrices decomposed by WF, M_{ij} , are calculated as the time-frequency features of EMG signals in the vastus medialis muscle. For the selected candidate segments, the formula is

$$M_{ij} = \begin{bmatrix} \sigma_{11} & \sigma_{12} & \dots & \sigma_{1j} \\ \sigma_{21} & \sigma_{22} & \dots & \sigma_{2j} \\ \vdots & \vdots & \vdots & \vdots \\ \sigma_{i1} & \sigma_{i2} & \dots & \sigma_{ij} \end{bmatrix} \quad (18)$$

where i is the number of singular values for each trial data, and j is the number of EMG candidate segments.

The values of RMS and iEMG (see Eqs (2) and (3), respectively) of the time-domain based EMG signals are chosen as the time-domain features. In addition, the frequency based characteristics inclusive of the values of MNF and MDF (see Eqs (5) and (6), respectively) across all trials are calculated as the frequency-domain features.

For the testing sample, x , the discriminant function is

$$f(x) = \sum_{i=1}^{N_s} y_i \alpha_i G(x, x_i) + b \quad (19)$$

where N_s is the number of the resulting support vectors, α_i is the positive Lagrangian multipliers and $G(x, x_i)$ is a kernel function. The vital regularization parameter in SVM is also determined by the five-fold cross-validation procedure that is repeated fifty times in order to avoid the random errors.

2.6 Evaluation procedure

In order to indicate the classification performance in terms of standing, sitting, stance phase of walking, and swing phase of walking, features from the time domain, the frequency domain, and the WT-based SVD approach are tested separately through the predefined evaluation procedure shown in Fig 1. In the procedure, the raw EMG data is first filtered by a high-pass filter with 20 Hz. The signals then are trial-by-trial segmented following the synchronized goniometry data. After preprocessing and segmentation of the raw EMG signals, the features represented by the analyses of the time domain, the frequency domain, and the WT-based SVD approach are extracted to build the feature vectors. A SVM-based multiple subject classifier was trained by the trials across all subjects and the trained classifier was then used for predicting and accumulating of the discrimination evidence over features. The datasets of four types

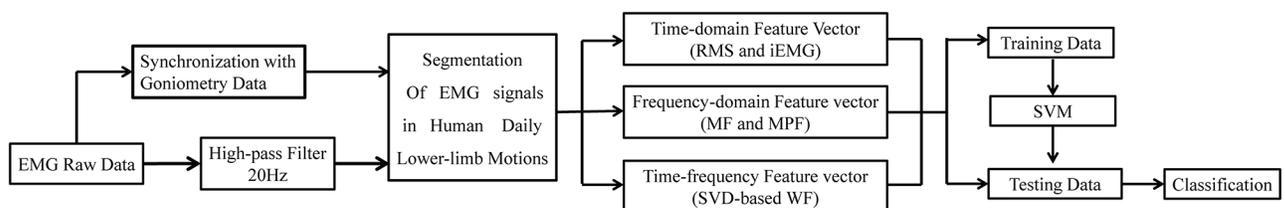


Fig 1. The evaluation procedure for classification of four types of lower-limb motions.

<https://doi.org/10.1371/journal.pone.0180526.g001>

of lower-limb motions from all subjects were randomly shuffled and then partitioned totally five times. It then was divided into a training set (80% of dataset) or a testing set (20% of dataset). The five-fold cross validation is used for the method evaluations and it is repetitively run fifty times to avoid the random errors. The average of the total accuracy results for fifty-time five-fold cross-validations is computed as the final accuracy result.

3 Experiment

The University of Technology Sydney (UTS) Human Research Ethics Committee (UTS HREC 2009000227) approved this experiment and collected the EMG datasets. A poster was used to introduce this experiment in order to recruit the participants from UTS campus, which was publicly open during the summer holiday in 2012. An informed consent was obtained for each participant prior to data collection. After the experiment, the remuneration was paid to the participants. The University of Electronic Science and Technology of China (UESTC) also approved to use the databases obtained from UTS for publications.

The EMG data collection process was carried out with one electrode placed on the muscle group of vastus medialis and the goniometer on the knee joint. The Datalog device MWX8 (<http://www.biometricsltd.com/datalog.htm>) is used for EMG data acquisition including 8 digital channels and 4 analog channels. One of the digital channels is used for monitoring and recording EMG signals and another for goniometry data of the knee joint of the detected foot. The data are first acquired and memorized in the MWX8 internal storage with microSD card, and then transmitted through the real-time Datalog software via Bluetooth adapter for the off-line analyses for this study. The sampling frequency is 1 kHz [38].

Since the EMG signals repetitively obtained from one subject can vary from the emotional state, food and caffeine intake, previous activity, fatigue, only healthy and untrained subjects were invited in this experiment. Moreover, the participants were asked to have a light meal at least two hours before the experiment and not to engage in intense or prolonged exercise for 24 hours prior to each experiment. Environmental conditions were the same for all participants.

In order to minimize the impedance of the electrode-skin contact in-between, the hair and dead skin cells on the target electrode placement location were shaved from the skin surface before prior to experiment. Once shaved, the skin was cleaned with alcohol and the electrode is then placed on the target skin after it comes to dry. The bipolar electrode (SX230) is used for the experiment. A fixed electrode distance of 20 mm is offered. It offers an integral electrode with fixed inter-electrode distance which gives consistent high quality results and effectively limits the risk of inter-electrode cross-talk. The R506 compatible with Datalog device MWX8 has a standard 4mm snap connector, which could minimize the risk of simultaneously recording similar signals. The input impedance of the EMG amplifier is more than 10,000,000 M Ohms [38].

The databases of 14 healthy untrained subjects (age: 28 ± 7.82 years old, height: 174 ± 5.88 cm, mass: 67.64 ± 11.35 kg) are taken into the study. The subjects undergo three exercise programs associated with the knee joint, leg extension from a sitting position (sitting), flexion of the leg up (standing), and gait (walking), while EMG signals from the vastus medialis muscle of the detected leg are monitored and recorded. Two experimental sessions are arranged, i.e., standing-sitting and walking, in which subjects are asked to randomly perform these exercise programs. For both sessions, subjects are asked to start the onset of motions with a one-second tone-paced timing with one-second interval, and follow a rest with five-second interval. Each motion is repeated five times in order to avoid intra-subject variability. For the standing-sitting motion, the order of performing both motions is randomly chosen. In addition, a fifteen-

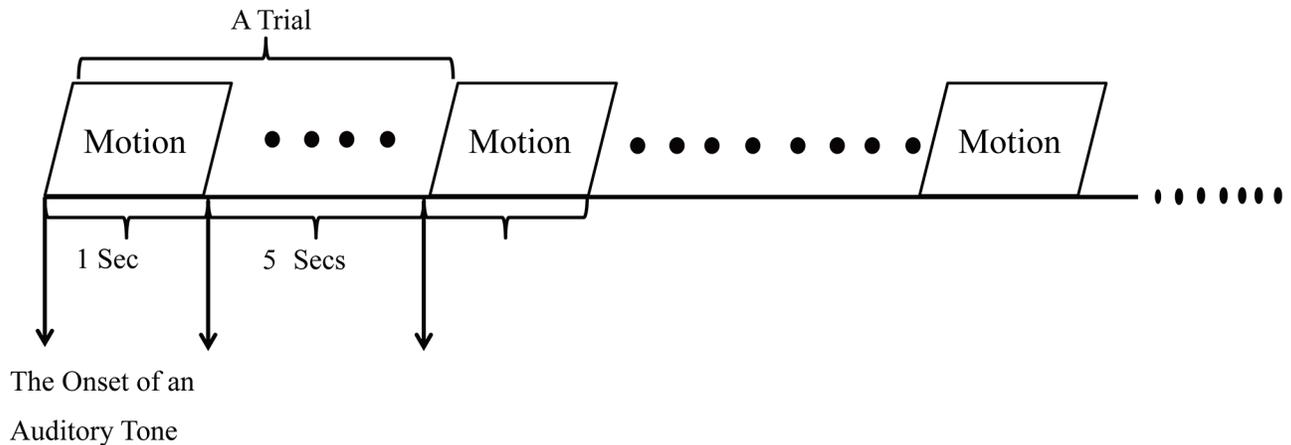


Fig 2. The experiment protocol for EMG data collection in terms of standing-sitting and walking sessions.

<https://doi.org/10.1371/journal.pone.0180526.g002>

minute break is offered between two sessions. The details of exercise protocol are as follows in Fig 2. During the experiments, a tone-paced timing is presented which follows an automatic tone, pacing interval where the onset of each motion is matched with the onset of an auditory tone in order to remind subject to start and stop their motions. Based on the EMG datasets obtained from both sessions, four types of lower-limb motions including standing, sitting, stance phase of walking, and swing phase of walking, are segmented, which will be used for the classification evaluation of the proposed WT-based SVD approach. For the details of original data, see S1 File.

4 Results

4.1 Time-domain analyses

In the preprocessing step, the raw EMG data is filtered by using a 20 Hz butterworth high-pass filter. The filtered data is then segmented following the synchronized goniometry data. The segmented trials include lower-limb motions for standing, sitting, stance phase of walking, and swing phase of walking. The EMG time series of such types of trails are presented in Fig 3. It can be found that for both standing and sitting motions the EMG strength increases (decreases) along with the process of flexion (extension) of the detected leg; for swing and stance phases the variations of the amplitudes of the swing phase are less affected compared with those of the stance phase.

The time-domain features for each trial across all subjects including MAV, RMS, iEMG, and ZC, are calculated. Table 1 shows the mean and STD values of those time-domain features for four lower-limb motions. Based on the segmented EMG data for all trials, the statistical analysis for these four time-domain features is evaluated by one-way analysis of variance analysis, which shows that MAV, RMS, iEMGP, and ZC have means significantly different for lower-limb motions (standing, sitting, stance phase of walking, and swing phase of walking) ($p < 0.05$).

4.2 Frequency-domain analyses

In this study, the frequency-domain features of EMG segmentations are also investigated. The raw data is filtered by 20 Hz high-pass filter and then segmented following the synchronized goniometry data in order to obtain the specified EMG signals for lower-limb motions. The

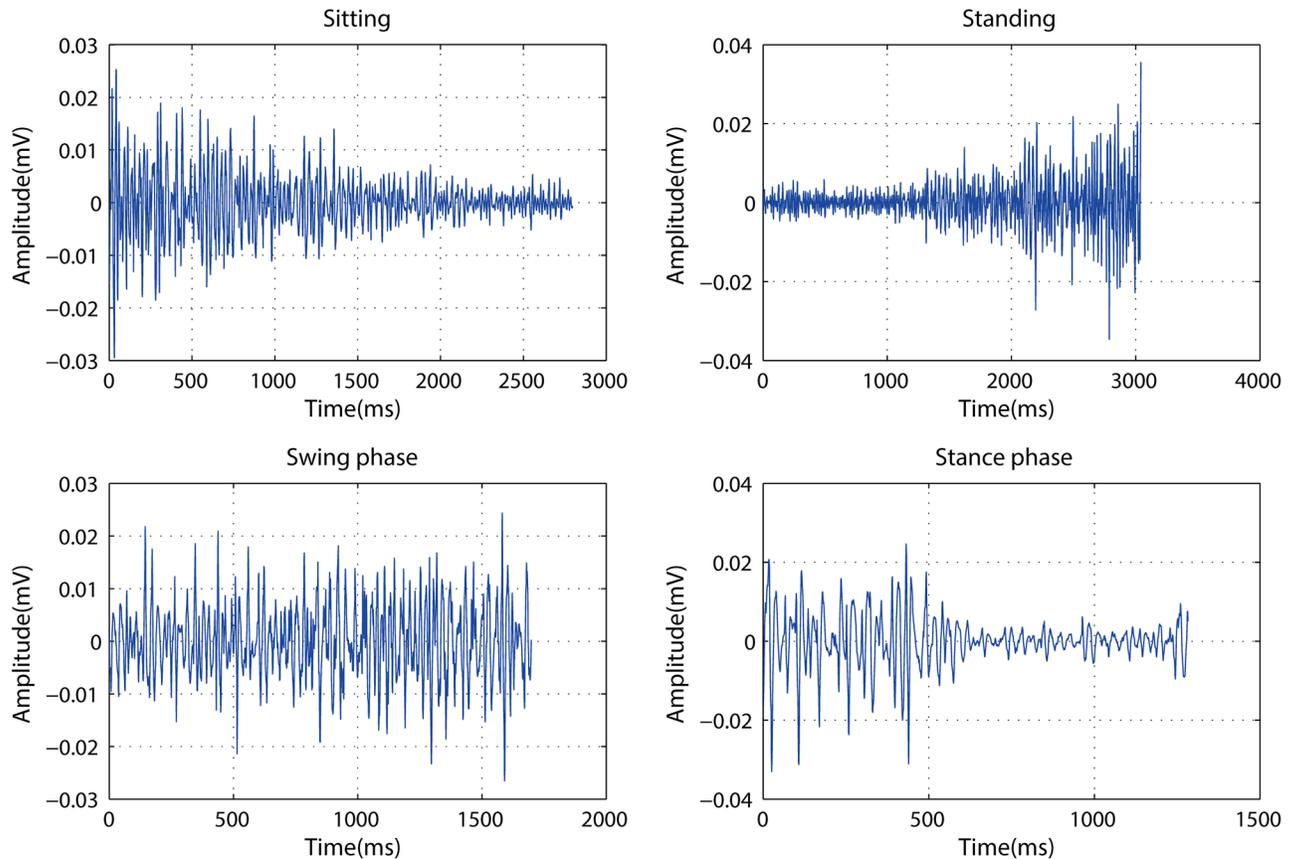


Fig 3. The EMG time series for the segmented lower-limb motions in terms of the swing phase, the stance phase, sitting, and standing.

<https://doi.org/10.1371/journal.pone.0180526.g003>

Welch’s averaged periodogram method [39] is utilized to demonstrate the Power Spectrum Density (PSD) of the EMG segmented data. The PSD curves based on the different types of lower-limb motions from one subject are shown in Fig 4.

The frequency-domain features such as MNF and MDF are calculated based on each trail data. The averaged MNF and MDF results for all trials of each motions across all subjects are provided in Table 2. Based on the calculated MNF and MDF values for all trials, the statistical analysis for the frequency-domain features is evaluated by one-way analysis of variance analysis, which shows that MNF and MDF have means significantly different for lower-limb motions (standing, sitting, stance phase of walking, and swing phase of walking) ($p < 0.05$).

Table 1. The mean and STD for temporal features based on experimental trails from all participants on the vastus medialis muscle.

	MAV Mean±STD%	RMS Mean±STD%	iEMG Mean±STD%	ZC Mean±STD%
Swing	0.02±50	0.03±33.3	25.42±45.25	137.03±35.88
Stance	0.01±100	0.02±50	27.28±78.85	260.82±22.11
Standing	0.01±200	0.01±400	5.44±143.38	97.84±73.95
Sitting	0.01±300	0.02±300	11.65±188.24	132.82±56.99

STD% represents the percentage of the standard deviation over the corresponding mean.

<https://doi.org/10.1371/journal.pone.0180526.t001>

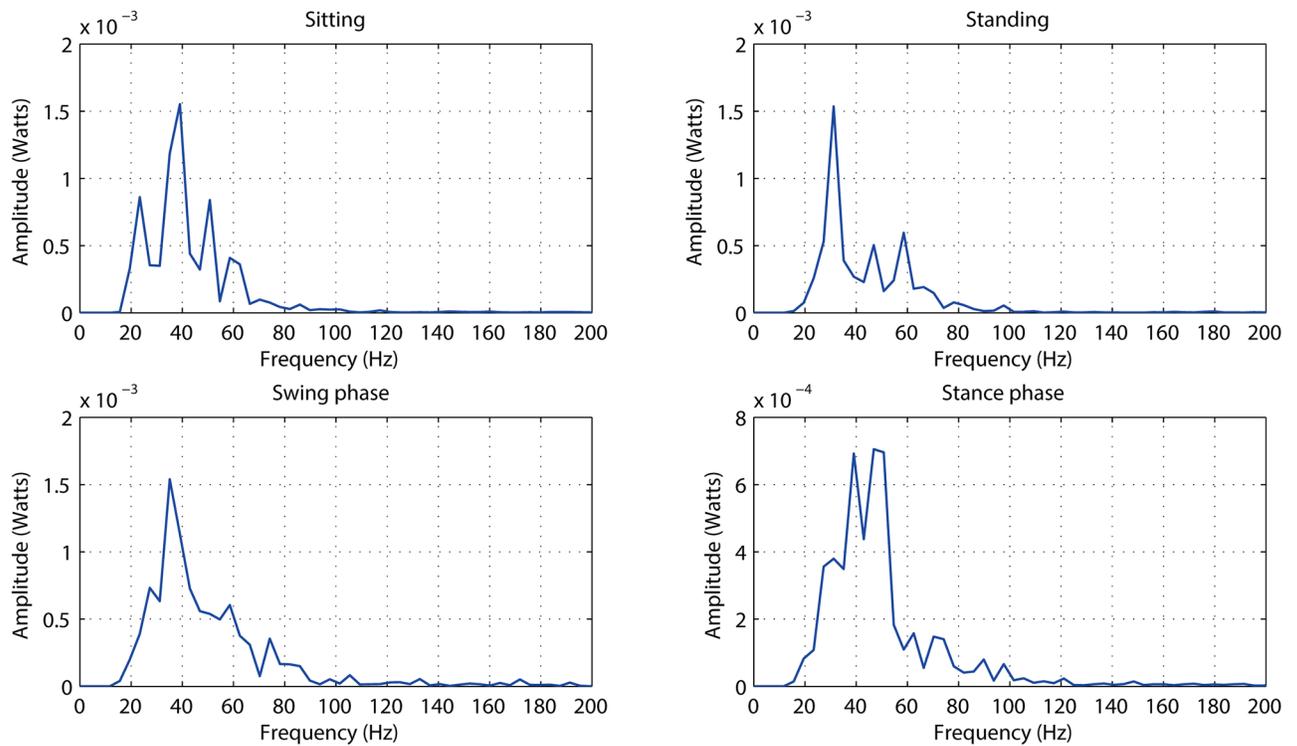


Fig 4. The PSD curves of the segmented EMG time series in terms of the swing phase, the stance phase, sitting, and standing.

<https://doi.org/10.1371/journal.pone.0180526.g004>

Furthermore, paired t-test results indicate that there are significant differences between any two of the four lower-limb motions for the calculated MNF and MDF values. Therefore, both MDF and MNF are used to represent the frequency-domain feature.

4.3 Time-frequency features extracted by WT-based SVD analyses

In order to obtain the time-frequency features by using the proposed WT-based SVD approach, the raw EMG data are also preprocessed through a 20 Hz high-pass filter. The filtered EMG data are then segmented to establish the candidate datasets for classifications. As mentioned in Section ‘Methods’, the five-level wavelet decomposition is applied to decompose the candidate EMG segmentations into five different scales time series. The coefficients for the high-frequency components (cD1, cD2, cD3, cD4, and cD5) and the coefficient for the

Table 2. The mean and STD for frequency-domain features based on experimental trails from all participants on the vastus medialis muscle.

	MDF Mean±STD%	MNF Mean±STD%
Swing phase	45.43±11.64	53.07±9.37
Stance phase	48.48±13.51	54.17±7.72
Standing	36.52±22.78	43.82±19.17
Sitting	38.38±22.64	45.25±19.51

STD% represents the percentage of the standard deviation over the corresponding mean. The unit for mean values of MDF and MNF is Hz.

<https://doi.org/10.1371/journal.pone.0180526.t002>

Table 3. The mean and STD results of time-frequency features by WT-based SVD approach based on experimental trails from all participants on the vastus medialis muscle.

	Stance Mean±STD%	Swing Mean±STD%	Standing Mean±STD%	Sitting Mean±STD%
cD1	0.22±68.18	0.25±48	0.26±430.77	0.26±315.38
cD2	0.44±40.91	0.35±71.43	0.23±213.64	0.48±274.47
cD3	0.85±40	0.65±78.46	0.24±216.67	0.4±231.71
cD4	0.54±62.96	0.42±66.67	0.1±150	0.16±168.75
cD5	0.21±55	0.16±68.75	0.04±125	0.06±150
cA5	0.11±70	0.08±75	0.02±100	0.03±100

STD% represents the percentage of the standard deviation over the corresponding mean.

<https://doi.org/10.1371/journal.pone.0180526.t003>

low-frequency component at the fifth level (cA5) are given. The singular value of each component is then computed by SVD. Table 3 shows the mean and STD values of the singular values for each decomposed components. Based on these results, an one-way analysis of variance analysis is also used, by which the statistical analysis results indicate that the six components have means significant difference for lower-limb motions (standing, sitting, stance phase of walking, and swing phase of walking) ($p < 0.05$).

4.4 Classification results

Based on the proposed evaluation procedure, the SVM is used to build a multiple-subject classifier for which the subject independent accuracy will be given across all subjects for the classification of four types of lower-limb motions. The five-fold cross validation is performed and it repeats fifty times in order to acquire the robust subject independent accuracy.

There were 260 EMG trials collected from three subjects participating in the lower-limb motion experiment, and for all subjects they were set into four sets of 40 segments for standing, 40 for sitting, 90 for the swing phase and 90 for the stance phase. The datasets are then randomly partitioned into five equal sized subsets. Of the five subsets, a single subset is retained as the validation data for testing, and the remaining four subsets are used for training. The evaluation procedure is then repeated fifty times, with each of the five subsets used exactly once as the validation data. The averaged results from the fold are then used to produce the subject independent accuracy across all subjects' trails.

Three different features are used as the feature vectors to train the SVM-based multiple-subject classifier, i.e., time-domain features (MAV, RMS, iEMG, and ZC), frequency-domain feature (MNF+MDF), and WT-based SVD features (cD1, cD2, cD3, cD4, cD5, and cA5). In order to evaluate the classification performance, there are five different feature vectors established for training and testing the lower-limb motions, including time-domain features (MAV+RMS+iEMG+ZC), frequency-domain feature (MNF+MDF), combined time and frequency features (MAV+RMS+iEMG+ZC+MDF+MNF), WT-based SVD features (cD1+cD2+cD3+cD4+cD5+cA5), and combined time, frequency, and WT-based SVD features (MAV+RMS+iEMG+ZC+MDF+MNF+cD1+cD2+cD3+cD4+cD5+cA5). The classification results are demonstrated in Table 4

The WT-based SVD features are demonstrated in Fig 5, where the WF-based SVD coefficients for all trials with four lower-limb motions are plotted. Based on this feature vector, the classification results show that the proposed WT-based SVD approach has the classification accuracy of 91.85%±0.88% which outperforms the time-domain features (66.96%±3.12%), the frequency-domain features (77.1%±0.01%), the combined time and frequency features

Table 4. The subject independent accuracies for standing, sitting, stance phase of walking, and swing phase of walking.

Feature vectors	Accuracy(%)	STD(%)
MAV+RMS+iEMG+ZC	66.96	3.12
MNF+MDF	77.1	0.01
MAV+RMS+iEMG+ZC+MDF+MNF	55.11	0.02
cD1+cD2+cD3+cD4+cD5+cA5	91.85	0.88
MAV+RMS+iEMG+ZC+MDF+MNF+cD1+cD2+cD3+cD4+cD5+cA5	53.36	0.02

STD represents the percentage of the standard deviation over the corresponding accuracy.

<https://doi.org/10.1371/journal.pone.0180526.t004>

(55.11%±0.02%), and the combined time, frequency, and WT-based SVD features (53.36%±0.02%).

Discussions and limitations

It has been clearly shown that either time-domain features or frequency-domain features cannot completely depict the differences among those motions. Results also confirmed that the combined features (combined time and frequency features and combined time, frequency, and WT-based SVD features) also cannot give a better classification accuracy compared to the WT-based SVD features. It is the fact that the combined features cannot maximize the discriminative power of EMG features for classifying lower-limb motions with a single EMG channel placed on the vastus medialis.

The contribution of this work is to introduce the SVD-based WT approach for the single-channel-based classification of lower limb activities. In the first step, it was natural to first explore the recognition accuracy of such proposal. In order to find discriminative features, the statistical analysis was employed in this study. The databases of 14 healthy untrained subjects are taken into the study. The subjects undergo three exercise programs associated with the knee joint, leg extension from a sitting position (sitting), flexion of the leg up (standing), and gait (walking), while EMG signals from the vastus medialis muscle of the detected leg are monitored and recorded. Based on the EMG datasets obtained from both sessions, four types of lower-limb motions including standing, sitting, stance phase of walking, and swing phase of walking, are segmented, which will be used for the classification evaluation of the proposed WT-based SVD approach.

Fig 3 explicitly shows that the amplitudes between sitting, standing stance phase of walking, and swing phase of walking, are distinctively different. The multivariate analysis is utilized by the one-way analysis of variance where the time-domain features, MAV, RMS, iEMG and ZC, have means significantly different. Thus, those are used for the analysis of time features in relation to knee exercise pattern recognition. Secondly, Mean Frequency (MNF) and Median Frequency (MDF) are used to indicate the frequency-domain characteristics of the EMG trails (see Fig 4). Based on these statistical results, MNF is selected as frequency-domain features due to the significant difference between any two of lower-limb motions.

Based on the classification results, it has been clearly shown that either time-domain features or frequency-domain features cannot completely depict the differences among those motions. Following this reason, the time-frequency features extracted by using the proposed SVD-WT approach are used in this study. The results present time-frequency features by using SVD-WT outcomes is better than ones by using either time-domain features or frequency-domain features (see Fig 5 and Table 4). However, this study has not addressed an

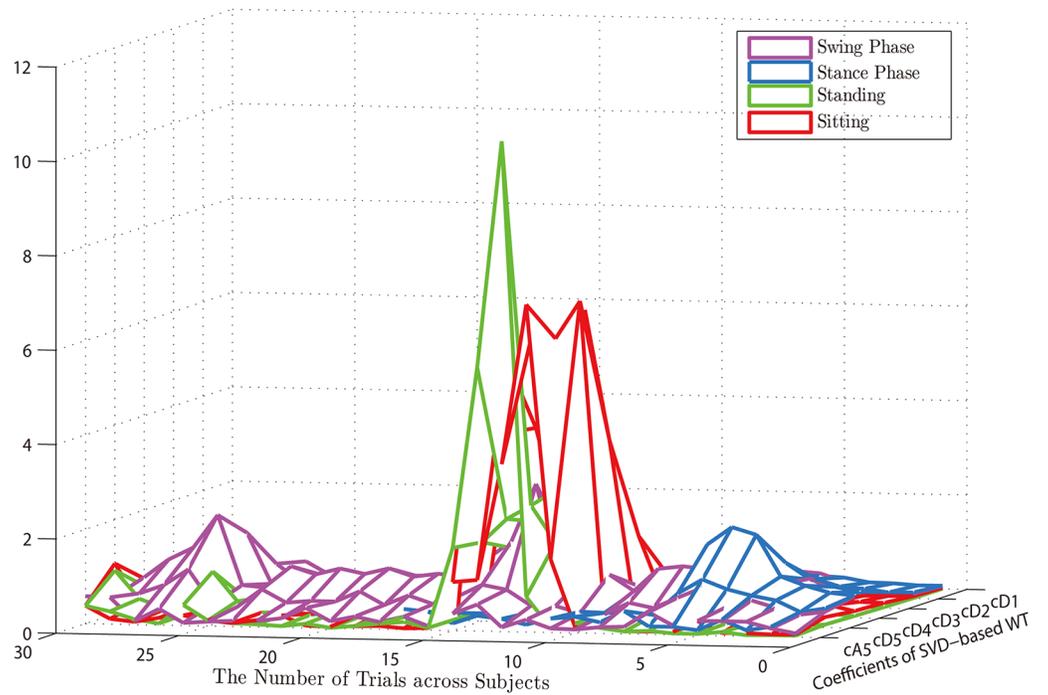
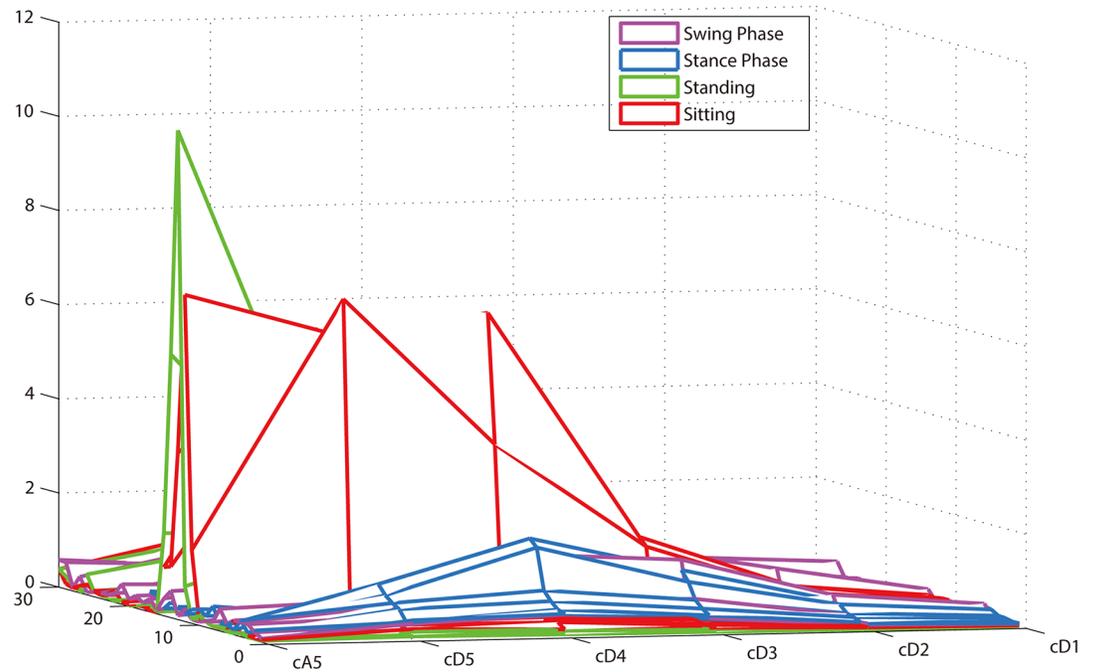


Fig 5. The WT-based SVD features for all trials with four lower-limb motions.

<https://doi.org/10.1371/journal.pone.0180526.g005>

optimal feature model described by the specified wavelet components that might achieve a better accuracy performance compared to the feature vectors with all wavelet components. In the next step of this study, Fisher-Markov selector will be applied in this study and quantitatively optimize the proposed WT-based SVD feature model [40] [41].

Conclusion

In this study, the single-channel EMG signals from the muscle group of vastus medialis were recorded. 14 subjects participated in the three exercise programs, while EMG signals from the muscle group of vastus medialis and the goniometer on the knee joint of the detected leg were synchronously monitored and recorded. Four types of lower-limb motions including standing, sitting, stance phase of walking, and swing phase of walking, were segmented. Based on the experimental data, the SVM is then configured to build a multiple-subject classifier for which the subject independent accuracy will be given across all subjects for the classification of four types of lower-limb motions. In order to effectively indicate the classification performance, EMG features from time-domain (e.g., Mean Absolute Value (MAV), Root-Mean-Square (RMS), integrated EMG (iEMG), Zero Crossing (ZC)) and frequency-domain (e.g., Mean Frequency (MNF) and Median Frequency (MDF)) were also used to classify lower-limb motions. The five-fold cross validation was performed and it repeats fifty times in order to acquire the robust subject independent accuracy. Three different features were used as the feature vectors to train the SVM-based multiple-subject classifier, i.e., time-domain features, frequency-domain feature, and WT-based SVD features. In order to evaluate the classification performance, there are five different feature vectors established for training and testing the lower-limb motions, including time-domain features (MAV+RMS+iEMG+ZC), frequency-domain feature (MNF+MDF), combined time and frequency features (MAV+RMS+iEMG+ZC+MDF+MNF), WT-based SVD features (cD1+cD2+cD3+cD4+cD5+cA5), and combined time, frequency, and WT-based SVD features (MAV+RMS+iEMG+ZC+MDF+MNF+cD1+cD2+cD3+cD4+cD5+cA5). Based on this feature vector, the classification results show that the proposed WT-based SVD approach has the classification accuracy of $91.85\% \pm 0.88\%$ which outperforms other feature models.

Supporting information

S1 File. Original EMG datasets. The EMG original data of the vastus medialis muscle of the detected leg from 14 healthy subjects were collected and recorded in the file while performing three exercise programs, i.e., leg extension from a sitting position (sitting), flexion of the leg up (standing), and gait (walking).
(ZIP)

Acknowledgments

This work was supported by the grant from the Fundamental Research Funds for the Central Universities (ZYGX2015J118), the National Natural Science Foundation (61522105, 51675087), and the China Postdoctoral Science Foundation funded project (2017M612950). Yi Zhang found the WT-based SVD approach, applied it to the classification of lower-limb motions, and drafted the paper; Yi Zhang, Peiyang Li, and Xuyang Zhu completed the results and revised the paper; Steven W Su provided the experimental data and proofreading of this paper; Qing Quo, Peng Xu, and Dezhong Yao supervised the project and revised the paper.

Author Contributions

Conceptualization: Yi Zhang.

Data curation: Yi Zhang, Peiyang Li, Steven W. Su.

Formal analysis: Yi Zhang, Peiyang Li, Qing Guo.

Funding acquisition: Yi Zhang.

Investigation: Peiyang Li.

Methodology: Yi Zhang, Peiyang Li, Peng Xu.

Resources: Yi Zhang, Xuyang Zhu.

Software: Yi Zhang, Xuyang Zhu.

Supervision: Steven W. Su, Peng Xu, Dezhong Yao.

Validation: Yi Zhang, Peiyang Li, Xuyang Zhu, Qing Guo, Peng Xu.

Visualization: Peiyang Li, Xuyang Zhu, Qing Guo.

Writing – original draft: Yi Zhang.

Writing – review & editing: Yi Zhang, Steven W. Su, Peng Xu.

References

1. Zoss AB, Kazerooni H. Biomechanical design of the Berkeley lower extremity exoskeleton (BLEEX). *IEEE/ASME Transactions on Mechatronics*. 2006 Apr; 11(2):128–138. <https://doi.org/10.1109/TMECH.2006.871087>
2. Zhang Y, Zhao H, Zhang X, Qin K. Support Vector Machine(SVM)-based Feature Extraction and Recognition of EMG signals during Lower Extremity Movement. *Manned Spaceflight*. 2015 Apr; 21(2): 179–186.
3. Zeilig G, Weingarden H, Zwecker M, Dudkiewicz I, Bloch A, Esquenazi A. Safety and tolerance of the ReWalk exoskeleton suit for ambulation by people with complete spinal cord injury: A pilot study. *Journal of spinal cord medicine*. 2012 Apr; 35(2):96–101. <https://doi.org/10.1179/2045772312Y.0000000003> PMID: 22333043
4. Au SK, Dilworth P, Herr H. An ankle-foot emulation system for the study of human walking biomechanics. *Proceedings 2006 IEEE International Conference on Robotics and Automation*. 2006 May: 2939–2945.
5. Ferris DP, Sawicki GS, Domingo A. Powered lower limb orthoses for gait rehabilitation. *Topics in spinal cord injury rehabilitation*. 2005 Feb; 11(2):34–49. <https://doi.org/10.1310/6GL4-UM7X-519H-9JYD> PMID: 16568153
6. Kiguchi K, Esaki R, Tsuruta T, Watanabe K. An exoskeleton for human elbow and forearm motion assist. *Proceedings 2003 IEEE/RSJ International Conference on Intelligent Robots and Systems*. 2003 Oct:3600–3605.
7. Oonishi Y, Oh S, Hori Y. A new control method for power-assisted wheelchair based on the surface myoelectric signal. *IEEE Transactions on Industrial Electronics*. 2010 Sep; 57(9): 3191–3196. <https://doi.org/10.1109/TIE.2010.2051931>
8. Simpson RC. Smart wheelchairs: A literature review. *Journal of rehabilitation research and development*. 2005 Aug; 42(4): 423–435. <https://doi.org/10.1682/JRRD.2004.08.0101> PMID: 16320139
9. Lee S, Sankai Y. Virtual impedance adjustment in unconstrained motion for an exoskeletal robot assisting the lower limb. *Advanced Robotics*. 2005 Apr; 19(7): 773–795. <https://doi.org/10.1163/1568553054455095>
10. He H, Kiguchi K, Horikawa E. A study on lower-limb muscle activities during daily lower-limb motions. *International Journal of Bioelectromagnetism*. 2007 Jun; 9(2): 79–84.
11. Mohr M, Nann M, von Tscherner V, Eskofier B, Nigg BM. Task-Dependent Intermuscular Motor Unit Synchronization between Medial and Lateral Vastii Muscles during Dynamic and Isometric Squats. *PLoS ONE* 10;11: e0142048. <https://doi.org/10.1371/journal.pone.0142048> PMID: 26529604

12. Gibson A S C, Lambert M I, Noakes T D. Neural control of force output during maximal and submaximal exercise. *Sports Medicine*. 2001; 31(9): 637–650. <https://doi.org/10.2165/00007256-200131090-00001>
13. Cotman C W, Berchtold N C. Exercise: a behavioral intervention to enhance brain health and plasticity. *Trends in neurosciences*. 2002; 25(6): 295–301. [https://doi.org/10.1016/S0166-2236\(02\)02143-4](https://doi.org/10.1016/S0166-2236(02)02143-4) PMID: 12086747
14. Lloyd DG, Besier TF. An EMG-driven musculoskeletal model to estimate muscle forces and knee joint moments in vivo. *Journal of biomechanics*. 2003 Jun; 36(6): 765–776. [https://doi.org/10.1016/S0021-9290\(03\)00010-1](https://doi.org/10.1016/S0021-9290(03)00010-1) PMID: 12742444
15. Han J, Ding Q, Xiong A, Zhao X. A State-Space EMG Model for the Estimation of Continuous Joint Movements. *IEEE Transactions on Industrial Electronics*. 2015 Jul; 62(7): 4267–4275. <https://doi.org/10.1109/TIE.2014.2387337>
16. Enoka RM, Duchateau J. Muscle fatigue: what, why, and how it influences muscle function. *The Journal of Physiology*. 2008 Jan; 586(1): 11–23. <https://doi.org/10.1113/jphysiol.2007.139477> PMID: 17702815
17. Oladazimi M, Vaneghi FM, Safari MJ, Asadi H, Kaboli SHA. A Review for Feature Extraction of EMG Signal Processing. *International Conference on Computer and Automation Engineering*. 2012 May: 142–148.
18. Chen X, Zhu X. A novel spectral analysis method of EMG feature extraction for wrist motions recognition. 2nd IEEE RAS & EMBS International Conference on Biomedical Robotics and Biomechatronics. 2008 Oct: 73–78.
19. Zhang X, Wang Y, Han RP. Wavelet transform theory and its application in EMG signal processing. 2010 Seventh International Conference on Fuzzy Systems and Knowledge Discovery. 2010 Aug: 2234–2238.
20. Veer K, Sharma T. A novel feature extraction for robust EMG pattern recognition. *Eprint Arxiv*. 2009 Jan; 1(1): 71–80.
21. Katsis CD, Exarchos TP, Papaloukas C, Goletsis Y, Fotiadis DI, Sarmas I. A two-stage method for MUAP classification based on EMG decomposition. *Computers in Biology and Medicine*. 2007 Sep; 37(9): 1232–1240. <https://doi.org/10.1016/j.combiomed.2006.11.010> PMID: 17208215
22. Attenberger A, Buchenrieder K. Wavelet-based detrending for EMG noise removal. 2013 20th IEEE International Conference and Workshops on Engineering of Computer Based Systems (ECBS). 2013 Apr: 196–202.
23. Ombao H, Raz J, Von SR, Guo W. The SLEX model of a non-stationary random process. *Annals of the Institute of Statistical Mathematics*. 2012 Jan; 54(1), 171–200. <https://doi.org/10.1023/A:1016130108440>
24. Zhang X, Diao W, Cheng Z. Wavelet Transform and Singular Value Decomposition of EEG Signal for Pattern Recognition of Complicated Hand Activities. *First International Conference on Digital Human Modeling*. 2007;4561: 294–303.
25. Phinyomark A, Limsakul C, Phukpattaranont P. Application of Wavelet Analysis in EMG Feature Extraction for Pattern Classification. *Measurement Science Review*. 2012 Nov; 11(2):45–52.
26. Merletti R, Torino P. Standards for reporting EMG data. *Journal of Electromyography & Kinesiology*. 1996 Apr; 6(2): 3–4.
27. John A, Vijayan AE, Sudheer AP. Electromyography based control of robotic arm using entropy and zero crossing rate. 2015 Conference on Advances In Robotics. 2015 Jul: 69.
28. Hagg G. Electromyographic fatigue analysis based on the number of zero crossings. *Pflügers Archiv—European Journal of Physiology*. 1981 Jan; 391(1): 78–80. <https://doi.org/10.1007/BF00580699> PMID: 7279604
29. Bilodeau M, Arsenault AB, Gravel D, Bourbonnais D. Time and frequency analysis of EMG signals of homologous elbow flexors and extensors. *Medical and Biological Engineering and Computing*, 1992 Jun; 30(6): 640–644. <https://doi.org/10.1007/BF02446797> PMID: 1297020
30. Stulen FB, De Luca CJ. Frequency parameters of the myoelectric signal as a measure of muscle conduction velocity. *IEEE Transactions on Biomedical Engineering*. 1981 Jul; 28(7), 515–523. <https://doi.org/10.1109/TBME.1981.324738> PMID: 7275132
31. Hamedi M, Salleh SH, Ismail K, Noor AM, Astaraki M., Aslian H. Time-frequency facial gestures EMG analysis using bilinear distribution. 2015 IEEE International Conference on Signal and Image Processing Applications (ICSIPA). 2015 Oct; 169–173.
32. MathWorks. Daubechies wavelet filter computation. 2016;2. Available from: <https://cn.mathworks.com/help/wavelet/ref/dbaux.html>
33. Vapnik VN. *The nature of statistical learning theory*. Springer Science & Business Media, Heidelberg and New York: Springer-Verlag.; 2013.

34. Vapnik VN, Lerner A. Pattern recognition using generalized portrait method. *Automation and remote control*, 1963 Dec; 24(24): 774–780.
35. Vapnik VN. *Statistical learning theory*. Wiley & Interscience, New York: Wiley.; 1998.
36. Gunn SR, Brown M, Bossley KM. *Network performance assessment for neurofuzzy data modelling*. Springer Berlin Heidelberg: Springer-Verlag.; 1998.
37. MathWorks. Support Vector Machine Classification. 2016;1. Available from: <https://cn.mathworks.com/help/stats/support-vector-machine-classification.html>
38. Biometrics. The DataLOG MWX8. 2017;1. Available from: <http://www.biometricsltd.com/datalog.htm>
39. Welch P. D. The use of fast Fourier transform for the estimation of power spectra: A method based on time averaging over short, modified periodograms. *IEEE Trans. Audio Electroacoust.* 1967 Jun; AU-15(2):70–73. <https://doi.org/10.1109/TAU.1967.1161901>
40. Cheng Q, Zhou Hongbo, Cheng J. The Fisher-Markov selector: Fast selecting maximally separable feature subset for multi-class classification with applications to high-dimensional data. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 2011 June; 33(6). 1217–1233. <https://doi.org/10.1109/TPAMI.2010.195>
41. Cheng CC, Bai Y, Cheng J, Soltanian-Zadeh H, Cheng Q. Discriminative features for interictal epileptic discharges in intracerebral EEG signals *Proc. the 5th IEEE Int. Congress on Image and Signal Processing (CISP 2012)*, 2012;. 1791–1795.