

DSCC2019-9150

USER INTENT RECOGNITION FOR TRANSFEMORAL AMPUTEES WITH PROSTHETIC LEGS USING EVOLUTIONARY ALGORITHMS

Hanieh Mohammadi*

Doctoral Student

Dept. of Electrical Eng. and Computer Science
Cleveland State University
Cleveland, OH, 44115
Email: h.mohammadi17@csuohio.edu

Gholamreza Khademi

Doctoral Student

Dept. of Electrical Eng. and Computer Science
Cleveland State University
Cleveland, OH, 44115
Email: g.khademi17@csuohio.edu

Elizabeth C. Hardin

Principal Investigator

Cleveland FES Center, Motion Study Laboratory
Cleveland VA Medical Center
Cleveland, OH, 44106
Email: ehardin@fescenter.org

Dan Simon

Professor

Dept. Electrical Eng. and Computer Science
Cleveland State University
Cleveland, OH, 44115
Email: d.j.simon@csuohio.edu

ABSTRACT

User intent recognition (UIR) enables transfemoral amputees to walk reliably and seamlessly with prosthetic legs. The objective of this paper is to design a UIR system that is optimal in terms of both accuracy and parsimony. We propose the application of two methods to achieve this goal. The first is a filter method, Fisher's linear discriminant score (FLDS); and the second is a wrapper method, linear discriminant analysis (LDA). Both methods are combined with the evolutionary algorithm biogeography-based optimization (BBO) to find optimal feature subsets. The optimal subsets are then compared with a current state-of-the-art feature selection method, in conjunction with several powerful linear and nonlinear classifiers that are used to identify level ground walking at various speeds. Classification performance is enhanced with a majority voting filter. The best performance is achieved with a multi-class support vector machine that is trained with FLDS/BBO feature subset and that reduces the number of required features by up to 70% and attains a mean prediction accuracy of 98.94% for amputee sub-

jects. Results show the capability of advanced subset selection methods to construct a UIR system with simultaneous minimum complexity and maximum performance.

1 INTRODUCTION

Unfortunately, the lifestyle of amputees is significantly influenced due to the loss of their walking ability. Therefore, researchers have recently focused on the development of prosthetic legs to restore walking function to transfemoral amputees. In prosthetic legs, low-level and high-level control systems are the two primary control challenges. The objective of a low-level control system is to allow an amputee to walk similarly to able-bodied persons [1–3]. The objective of a high-level control system is to switch seamlessly among low-level controllers depending on the user's intention [4] (for example, changes in gait speed or ground slope). This paper aims to design a high-level control system, which is referred to as UIR.

Previous UIR research has used surface electromyography (sEMG) signals [5], mechanical sensors [6, 7], optical distance

* Address all correspondence to this author.

sensors [8], or inertial measurement units [9] from transfemoral amputees to design an UIR system with the maximum possible prediction accuracy. For this purpose, several time-domain and frequency-domain features have been extracted from each measurement signal. Then, different linear and nonlinear classifiers such as LDA [10], quadratic discriminant analysis (QDA) [11], artificial neural networks (ANNs) [5], and support vector machines (SVMs) [12] have been used to train a UIR system.

However, among extracted features, there might be irrelevant or redundant predictor variables that do not contribute to the UIR performance. These variables increase system complexity, and the amputee subject could be inconvenienced by wearing additional sensors. Therefore, we are motivated to achieve a UIR with simultaneously high prediction accuracy and high parsimony. UIR has high prediction accuracy if it can accurately classify test walking activities with which it has not been trained, and UIR has high parsimony if it is trained with only significant features extracted from minimal hardware. Accuracy is an attractive objective, because in clinical applications, it is extremely important that UIR can accurately predict activity modes with substantially different characteristics, where misclassification may cause a loss of balance [9,13]. Parsimony is also an attractive objective, because it: (1) requires a minimal number of features and body-worn sensors with significant predictive power; (2) avoids numerical instability and overfitting during training; and (3) requires less computational effort, which facilitates real-time operation with a simple, inexpensive processor. Therefore, in this paper we are motivated to propose a framework that allows for the elimination of sensing redundancies while still achieving high prediction performance.

We propose the application of advanced evolutionary-based feature selection to build an accurate UIR system for transfemoral amputees while, at the same time, eliminating redundancies. In this paper, we investigate filter and wrapper methods, both combined with BBO, to find an optimal feature subset with minimum size and maximum discriminating information for gait mode classification. There are many different heuristic search strategies; we have chosen BBO due to its demonstrated performance in solving high-dimensional real-world problems [14].

We use FLDS as our filter-based selection method [15, 16] and LDA as our wrapper-based selection method [17]. The combination of these two selection methods with BBO are called FLDS/BBO and LDA/BBO. In general, filter methods assess the quality of a subset of features, as measured by information, correlation, or class separability, independently or with respect to the output class. Wrapper methods assess the quality of a subset of features by measuring the prediction accuracy of a classifier trained by that subset [18]. Unlike wrapper methods, filter methods do not train a classifier to assess the quality of a subset, and therefore they are computationally inexpensive. However, unlike filter methods, wrapper methods consider the bias of the classifier in the selection process [18].

To verify the feasibility of our framework, in Section 2.1, we use experimental data collected from able-bodied and transfemoral amputee subjects as illustrative examples. We use a passive marker motion capture system to measure residual limb and user-prostheses interaction signals for UIR. Section 2.2 discusses the generation of multiple data frames from each measurement signal using sliding windows. Time-domain (TD) and frequency-domain (FD) features are extracted from data frames. Section 2.3 discusses FLDS/BBO and LDA/BBO in detail, presents a feature subset quality metric based on mutual information (MI), and combines MI with greedy forward search (MIFS) for subset selection. We compare the performance of our two new methods, FLDS/BBO and LDA/BBO, with MIFS. Section 2.4 uses LDA, QDA, a multi-layer perceptron (MLP) network, SVM, and multinomial logistic regression (MLR) to evaluate the performance of the selected subsets. Section 3 presents results and a discussion. Finally, Section 4 provides a conclusion and suggestions for future work.

2 MATERIALS AND METHODS

2.1 Experimental Setup and Data Collection

We used experimental data obtained from two able-bodied and two transfemoral amputee subjects as illustrative examples to demonstrate the capability of the developed framework to design an accurate and compact UIR. Although we used only four subjects in this study, the results are still significant because the UIR training in this paper: (1) can be individualized and thus may not need to be generalized to a larger number of participants, and (2) is a preliminary study whose generalizability will be explored in future research. To design an accurate UIR system for gait mode recognition, informative and discriminating signals of various human gait modes were collected. Measurement signals that reflect the state of the residual limb and the interactions between the amputee and the prosthesis are widely used in UIR [6,7]. In this paper, we used vertical hip position and thigh angle to reflect the state of the residual limb, and hip flexion moment (often called thigh moment in the literature), computed via inverse dynamics, to reflect the user-prostheses interaction. Measurement signals were collected for two able-bodied subjects (AB01 and AB02) and two transfemoral amputee subjects (AM01 and AM02). All the experiments were approved by the Department of Veterans Affairs Institutional Review Board. AB01 and AB02 were asked to walk on a treadmill at three different speeds, including user-preferred velocity, and slower and faster than user-preferred velocity. AM01 and AM02 were asked to walk at two different speeds, including user-preferred velocity and slower than user-preferred velocity. The UIR system is required to identify four gait modes for able-bodied subjects: ST (standing), SW (slow walking), NW (normal walking), and FW (fast walking). The UIR system is required to identify three gait modes for amputee subjects: ST, SW, and NW.



FIGURE 1. EXPERIMENTAL SETUP – THE LEFT FIGURE SHOWS ABLE-BODIED SUBJECT AB01, AND THE RIGHT FIGURE SHOWS AMPUTEE SUBJECT AM01 WEARING AN OTTO-BOCK PROSTHESIS ON THE RIGHT LEG.

Data collection was performed at the Motion Study Laboratory of the Cleveland Department of Veterans Affairs Medical Center. The treadmill featured two force platforms, one on the right side and one on the left side, to collect ground reaction forces at 1 kHz. Three dimensional coordinates were recorded with 47 reflective markers attached to the subject using a 16-camera passive marker motion capture system (Vicon, Oxford, UK) at 100 Hz. A second-order low-pass filter with cutoff frequency of 6 Hz was used to eliminate noise. A 3D biomechanical rigid body model was constructed from the marker data, and segmental and joint kinematics and kinetics (joint moments) were computed as input for the UIR system. Detailed methods and sample results can be found in [19]. Figure 1 shows the experimental setup for both able-bodied and amputee subjects.

Both amputee subjects used an Ottobock semi-passive prosthesis on their right leg. All four subjects (two able-bodied and two transfemoral amputees) were male with ages in the range 20–64 years, weight in the range 74–99 kg, and height in the range 172–188 cm. Multiple walking trials were collected from each subject, where each trial was a sequence of various gait modes. AB01 and AB02 performed four walking trials, each lasting about 60 seconds. AM01 and AM02 performed six and fifteen walking trials, respectively, each lasting about 30 seconds.

In summary, we note a few important points. Firstly, in this paper, the type of walking activities used for recognition is not our main focus, but rather the assessment of the proposed methodology to eliminate irrelevant/redundant features for UIR is our main goal. Secondly, in human activity mode recognition applications, an entire stride is typically used for non-real-time classification [20]. However, in UIR, we use a small window of measurement signals, mostly within a few milliseconds, to identify user's intent for real-time prosthesis control.

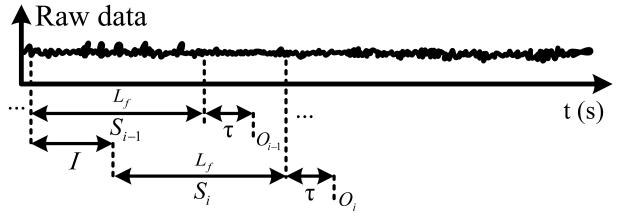


FIGURE 2. DATA SEGMENTATION; S_I INDICATES THE I -TH FRAME WITH LENGTH L_f , INCREMENT LENGTH IS I , PROCESSING TIME IS τ , AND THE IDENTIFIED GAIT MODE OF FRAME S_I IS O_I .

2.2 Feature extraction

To train an UIR system with a rich set of features, several TD and FD features are extracted from each data frame of three measurement signals. We use a sliding window with length L_f and increment I to create data frames, as illustrated in Fig 2. The class decision O_i associated with each frame S_i is output with period I , and τ is the processing time for feature extraction and classification.

Two key properties of human muscle contraction guide us in our selection of appropriate values for the sliding window parameters. The first property indicates that the frame length should be at least 200 ms to obtain informative features (that is, $L_f \geq 200$ ms). However, a large frame length would result in computational burden as well as decision bias due to the presence of multiple activity modes in a single frame [21]. The second property indicates that the UIR system must provide its output decision with a delay of less than 300 ms to provide prosthesis control that is responsive to gait mode changes (that is, $I \leq 300$ ms) [22, 23]. In addition, the sliding window parameters in Fig 2 should satisfy $\tau \leq I \leq L_f$ for real-time implementation. In Section 2.4, we will consider another constraint due to the majority voting filter (MVF) to choose parameter values.

TD features are extracted from each frame of data and include slope sign change (SSC), zero crossing (ZC), waveform length (WL), variance (VAR), mean absolute value (MAV), two modified MAVs, root mean square (RMS), Willison amplitude (WAMP), skewness (SK), kurtosis (KU), and correlation (COR) coefficient and angle (ANG) between two frames of data from different measurement signals. In addition, the following FD features are extracted from each frame: mean frequency (MNF), median frequency (MDF), maximum frequency (MAXF), and four coefficients of a fourth-order auto-regressive model (AR4). These features are used here because they have demonstrated competitive performance in the classification and control of prosthetic limbs [24–26]. Readers are referred to [24] for more details about these features.

In summary, we extract 20 features from each frame of data. Since we have three measurement signals (vertical hip position,

thigh angle, and hip flexion moment), the feature set has $20 \times 3 = 60$ elements. We normalize to find a data set with zero mean and unit variance.

2.3 Optimal subset selection

Optimal subset selection is used to eliminate redundant features from the training set while maintaining prediction accuracy. This will lead to an accurate and compact UIR system. We investigate filter and wrapper methods, combined with BBO, to find optimal feature subsets for gait mode classification.

2.3.1 Fisher's linear discriminant score: filter-based selection

FLDS is a filter-based selection method that provides a measure of separation between distinct classes. We use FLDS to assess the quality of a candidate subset of features. A subset of features with larger FLDS value implies a higher level of class separation and consequently a lower rate of misclassification. For a data set with K distinct classes and m observations, FLDS is defined as [15, 16, 27]

$$\text{FLDS} = \text{Trace}(S_W^{-1} S_B) \quad (1)$$

where between-class S_B and within-class S_W scatter matrices are calculated as

$$\begin{aligned} S_B &= \sum_{i=1}^K S_{B_i} = \sum_{i=1}^K m_i(\mu_i - \mu)(\mu_i - \mu)^T \\ S_W &= \sum_{i=1}^K S_{W_i} = \sum_{i=1}^K \sum_{x \in \pi_i} (x - \mu_i)(x - \mu_i)^T \\ \mu &= \frac{1}{m} \sum_{j=1}^m x_j \end{aligned} \quad (2)$$

where x_j is the j -th training pattern $j \in \{1, 2, \dots, m\}$, and π_i is the i -th population with mean μ_i and $m_i < m$ observations that belong to class $i \in \{1, 2, \dots, K\}$. Fig 3 illustrates between-class and within-class scatter matrices for an example data set with three distinct classes. Fig 3 and Eq 1 show that the larger the sum of square distances from the mean of the populations to the overall mean, the larger the FLDS value. Moreover, the smaller the sum of within-class variances, the larger the FLDS value.

2.3.2 Linear discriminant analysis: wrapper-based selection

Wrapper methods assess the quality of a candidate subset of features using the prediction accuracy of a classifier that is trained and validated by that subset. We may train a classifier many times depending on the search strategy used with the wrapper method. This indicates the need for a

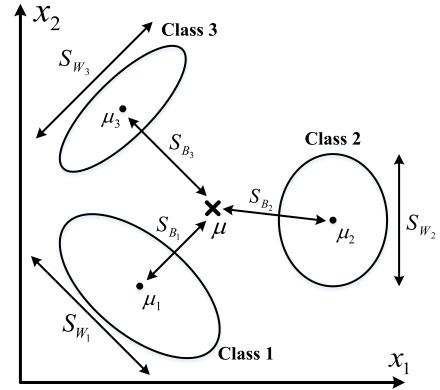


FIGURE 3. ILLUSTRATION OF BETWEEN-CLASS AND WITHIN-CLASS SCATTER MATRICES FOR A DATA SET WITH THREE DISTINCT CLASSES.

computationally fast classifier. Therefore, we use the LDA classifier in this paper due to its demonstrated performance in gait mode classification [9] as well as its efficient training procedure.

We use prediction error as the wrapper-based quality measure rather than prediction accuracy. This is simply because our search strategy preference is to solve a minimization problem rather than a maximization problem. The quality measure is defined as the mean prediction error provided by our cross validation (CV) procedure. In c -fold CV, we randomly partition the training patterns into c separate complementary subsets. At each time, the classifier is trained using $c - 1$ subsets, and is validated with the remaining subset. We repeat this training procedure c times, and the average prediction error is used to estimate the quality of selected subset.

2.3.3 Search strategy We have discussed two powerful tools to evaluate the quality of candidate feature subsets. A search strategy can utilize these quality metrics to find the best set of features. It is clear that the use of exhaustive search to find the optimal subset is not feasible due to the existence of 2^{60} different subsets in a training set with $n = 60$ features. Therefore, we propose an evolutionary-based search strategy to seek for an optimal subset of features.

We use BBO as the search strategy. BBO is chosen because of its demonstrated performance in solving high-dimensional real-world problems [14]. BBO is a stochastic evolutionary search strategy based on the migration, evolution, and extinction of species [28]. BBO is initialized with a population of N random individuals in the search domain. Each individual (candidate subset of features) is a binary sequence with a length equal to the number of features n (independent variables) where the binary value “1” indicates that a given feature is included in the candidate subset, and a “0” indicates otherwise. BBO operators

evolve the population toward an optimal solution. Standard BBO operators are: (1) migration of independent variables among individuals on the basis of their objective function values; (2) random mutation of each individual; and (3) maintenance of elite individuals in the population [28]. In a later section, we will tune the BBO operators for efficient performance of the UIR system (population size N , number of generations Gen , mutation rate μ , and number of elites E).

We define the following objective function for the filter-based selection method:

$$\text{cost}_i = \alpha \left(1 - \frac{\text{FLDS}_i}{\max_i(\text{FLDS}_i)} \right) + (1 - \alpha) \left(\frac{l_i}{D} \right) \quad (3)$$

$$i = 1, \dots, N$$

and we define the following objective function for the wrapper-based selection method:

$$\text{cost}_i = \alpha e_i + (1 - \alpha) \left(\frac{l_i}{D} \right) \quad i = 1, \dots, N \quad (4)$$

cost_i is the cost of the i -th individual, N is the population size, D is the problem dimension (equal to the number of candidate features), l_i is the total number of selected features for the i -th individual, FLDS_i is the filter-based quality measure for the i -th individual, $\alpha \in [0, 1]$ is the importance of performance relative to the number of features, and e_i is the average prediction error of the model trained by the i -th individual. Prediction error is in the range $[0, 1]$, and is defined as the number of incorrectly classified test data divided by the total number of classified test data. The first term in Eqs. 3 and 4 penalize an individual with a poor quality measure, whereas the second term penalizes the size of the selected feature subset. The filter-based method FLDS and the wrapper-based method LDA are combined with BBO to yield the methods denoted as FLDS/BBO and LDA/BBO, respectively.

2.3.4 Sequential forward selection using mutual information In this section, we discuss MI-based sequential forward selection [29], referred to as MIFS, to select the most informative subset of features for comparison with FLDS/BBO and LDA/BBO. MI is a filter method that provides a measure of mutual dependence between two random variables. For instance, MI between an input feature X and an output class Y is given by [29, 30]

$$I(X; Y) = \sum_{y \in Y} \sum_{x \in X} p(x, y) \log \left(\frac{p(x, y)}{p(x)p(y)} \right) \quad (5)$$

where $p(x, y)$ is the joint probability density function (pdf), and $p(x)$ and $p(y)$ are the marginal pdfs. MI in Eq. 5 can be esti-

mated by approximating the joint and marginal pdfs with histograms [30]. MI measures the amount of information provided by an input feature about the output class. Therefore, the larger the value of MI, the better the quality of the input feature for classification.

First, MIFS selects the feature from the n candidate features that has the maximum MI with the output class. The next feature selected by MIFS is the one that provides the maximum information about the output class and the minimum information about the previously selected features. This strategy selects features with maximum information but low correlation with previously selected features. The MIFS feature selection metric is given as

$$ind = \operatorname{argmax}_i \left[I(f_i; Y) - \beta \sum_{s_j \in S} I(f_i; s_j) \right] \quad (6)$$

where ind is the index of the selected feature, f_i is the i -th feature in the candidate set, s_j is the j -th feature in the selected set S , Y is the output class, and $\beta \geq 0$ is a tuning parameter. MIFS repeats the procedure to obtain an informative subset with size $k \leq n$ where k is a user-defined parameter.

2.4 Classification and majority voting filter

We use various classifiers to assess the quality of the feature subsets that are obtained by FLDS/BBO, LDA/BBO, and MIFS. We use LDA, QDA, MLP with one (MLP-1) and two (MLP-2) hidden layers, SVM with RBF kernels, and MLR. Throughout the paper, we train each classifier for each individual subject using c -fold CV, and the prediction accuracy, averaged over all subjects and folds, is reported as the overall classification performance metric.

Finally, classification decisions are input to an MVF for improved prediction performance. MVF uses the current, r previous, and r next classification decisions to output the gait mode with maximum occurrence as the currently identified activity mode [26]. Thus, MVF requires $2r + 1$ classification decisions. As mentioned in Section 2.2, the maximum delay time to provide a classification decision is 300 ms. Therefore, due to the need for r future classification decisions, the following constraint must be satisfied by the sliding window parameters:

$$r \times I \leq 300 \text{ ms} \quad (7)$$

3 RESULTS AND DISCUSSION

This section performs optimal subset selection using FLDS/BBO, LDA/BBO, and MIFS. UIR is then trained and evaluated for each individual subject with selected feature subsets.

3.1 Training data set

Experimental data collected from able-bodied and amputee subjects are used for optimal subset selection. To satisfy the constraints in Section 2.2, we use $L_f = 250$ ms and $I = 50$ ms, so the average value of τ is considerably less than I . It was shown that these values create the best trade-off between UIR accuracy and computational complexity [31]. The increment length was selected to allow $r = 5$ for MVF, which comfortably satisfies the constraint of Eq. 7. We have found that MVF with 11 elements provides a good trade-off between prediction accuracy and real-time computational effort. We extract TD and FD features from all the data frames produced by the sliding window with these parameter values. AB01, AB02, AM01, and AM02 provide training matrices with size 5375×60 , 5744×60 , 2101×60 , and 6998×60 , respectively.

3.2 Preliminary feature selection

We use able-bodied gait data to identify potential features with high discriminating information content for UIR, and then we use a selected subset of features to train and validate the UIR system with amputee gait data. Prior to feature selection using either BBO (FLDS and LDA) or greedy search (MI), we evaluate the performance of each TD and FD feature separately using FLDS, LDA, and MI.

Table 1 shows the average amount of discriminating information available in each feature type. Note that each feature type includes three individual features; one from vertical hip position, one from thigh angle, and one from hip flexion moment. AR4 is an exception because it includes four autoregressive coefficients, and so the AR4 feature type in the table actually includes 12 individual features. The table also shows the processing ratio (PR) for able-bodied subject. PR is the comparative computational time required to calculate each feature. Large FLDS, MI, or LDA score is desirable, and small PR is desirable, and could indicate that the given feature type might provide accurate and fast classification for UIR.

Table 1 indicates that FLDS, MI, and LDA are consistent in their relative assessment of the feature types, and TD features perform better than FD features. For instance, all three methods indicate WL, WAMP, and VAR as the most informative features for gait mode classification with relatively low PR. In addition, FD features such as FMD, FMN, and FMAX have the highest PR, which is undesirable.

To graphically depict the significance of each feature, we plot the MI scores in Fig 4. The figure shows that the three WL features extracted from vertical hip position, thigh angle, and hip flexion moment have high MI scores (their average score was reported in Table 1). More specifically, WL from hip vertical position has the highest MI score. Fig 4 also clearly shows that TD features outperform FD features. To save space, Fig 4 shows MI scores for F_{17} only for the three vertical hip position coefficients;

TABLE 1. AVERAGE AMOUNT OF DISCRIMINATING INFORMATION, AND PROCESSING RATIO (PR), FOR EACH FEATURE TYPE, USING ABLE-BODIED GAIT DATA.

Features	Symbol	FLDS	MI	LDA (%)	PR (%)
SSC	F_1	0.538	0.1668	49.89	0.669
ZC	F_2	0.268	0.0675	41.38	0.519
WL	F_3	10.49	0.6365	85.69	0.236
VAR	F_4	2.597	0.2774	73.70	0.242
MAV	F_5	1.202	0.3899	51.68	0.093
MAV1	F_6	0.531	0.4336	27.99	0.939
MAV2	F_7	0.320	0.4425	24.97	1.110
RMS	F_8	1.381	0.4085	53.55	0.134
WAMP	F_9	32.84	0.5257	56.09	0.544
SK	F_{10}	0.036	0.0318	31.25	2.104
KU	F_{11}	0.054	0.0190	31.56	1.657
FMD	F_{12}	0.108	0.0303	31.92	28.27
FMN	F_{13}	0.214	0.1789	40.46	27.98
FMAX	F_{14}	0.112	0.0264	31.38	28.01
CORR	F_{15}	0.165	0.0599	33.96	2.952
ANG	F_{16}	1.138	0.3086	46.30	0.320
AR4	F_{17}	0.306	0.0043	42.80	4.221

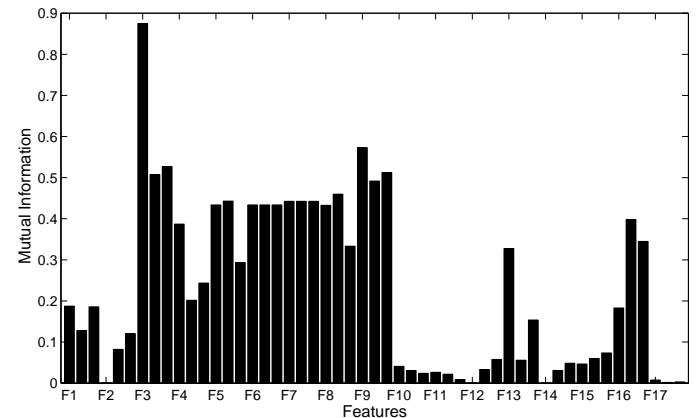


FIGURE 4. SIGNIFICANCE OF EACH FEATURE EXTRACTED FROM EACH MEASUREMENT SIGNAL USING MI SCORE

the other nine coefficients have approximately zero MI with the output class.

At this point, we exclude low-information features to reduce

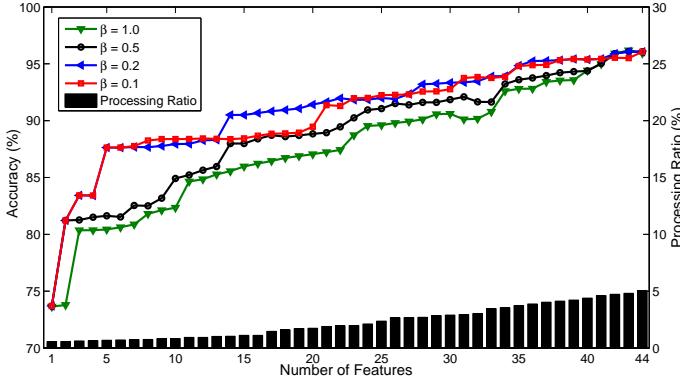


FIGURE 5. PROCESSING RATIO AND ACCURACY OF LDA, TRAINED WITH ABLE-BODIED SUBJECT DATA WITH MIFS FEATURE SUBSETS, AT EACH MIFS ITERATION.

the computational effort for finding the optimal feature subset. Table 1 and Fig 4 recommend the exclusion of: (1) MAV1 and MAV2 due to their poor quality measure, large PR compared to MAV, and inherent similarity with MAV; (2) FMD, FMN, and FMAX due to their poor quality measure and very large PR; and (3) ZC extracted from hip vertical position due to no correlation with the output class. After excluding these 16 features, we seek an optimal feature subset from a set of $60 - 16 = 44$ features using FLDS/BBO, LDA/BBO, and MIFS.

Note that we do not exclude AR4 features. The combination of the AR4 features provides significant information in terms of LDA accuracy (see Table 1) even though each individual feature has near-zero mutual information. It is possible that an AR4 feature that is useless by itself can provide a significant performance improvement when used with other AR4 features.

3.3 Optimal feature selection

We implement MIFS with different values of β , which consequently results in different feature subsets. MIFS initially selects the feature with the highest MI, and then, depending on the β value, includes more features based on their significance levels. At each step, we evaluate the mean classification performance of the selected subset using LDA on able-bodied subjects. In addition, processing ratio (PR), which is defined as the relative processing time required for validating the test data, is computed.

Fig 5 shows the quality of the selected subsets at each MIFS step for some representative values of β . Prediction performance increases as MIFS adds the next informative feature to the subset, and consequently PR increases. Fig 5 indicates that MIFS with $\beta = 0.2$ results in the best subset of features. We note that the feature subset size k is a user-defined parameter. We choose the subset with $k = 14$ and $\beta = 0.2$ because it provides a good trade-off between prediction performance and PR.

Table 2 shows the tuning parameters of FLDS/BBO and

TABLE 2. TUNING PARAMETERS FOR FLDS/BBO AND LDA/BBO. LDA/BBO HAS THE SAME PARAMETER VALUE AS FLDS/BBO, EXCEPT FOR THOSE SHOWN IN PARENTHESES.

	Symbol	Value
FLDS/BBO (LDA/BBO)		
Mutation rate	μ	0.06
Number of elites	E	10 (8)
Population size	N	100 (50)
Number of generations	Gen	4400 (200)
Problem dimension	D	44
Migration model	m_{flag}	sinusoidal
Cost function weight	α	0.8

LDA/BBO. We performed a sensitivity analysis of each selection method with respect to its tuning parameters, one at a time, to find the best parameter values. FLDS/BBO and LDA/BBO are run for 15 independent trials, and the feature subset with the minimum objective function value is taken as the best result.

We use the classifiers discussed in Section 2.4 to compare the performance of the feature subsets obtained by FLDS/BBO, LDA/BBO, and MIFS. The MLP-1 algorithm that we use includes ten hidden nodes, and the MLP-2 algorithm that we use includes ten and eight hidden nodes in the first and second layers, respectively. We use one-against-one binary SVM to implement multi-class SVM, and $\sigma = 1$ is the standard deviation parameter of RBF kernel. We also used steepest descent for training MLP and MLR classifiers with a maximum of 1500 iterations. All of the tunable parameters of the classifiers are tuned to achieve satisfactory performance.

Table 3 shows mean classification accuracy of able-bodied subjects with feature subsets obtained by FLDS/BBO, LDA/BBO, and MIFS. Results show that FLDS/BBO and LDA/BBO are statistically significantly better than MIFS at a 5% significance level. There is no meaningful difference between FLDS/BBO and LDA/BBO in terms of classification performance. LDA/BBO achieved similar results as FLDS/BBO with fewer features. MLP-2 and SVM result in better prediction accuracy than other classifiers. For instance, MLP-2 achieves an accuracy of 96.44% with the FLDS/BBO-selected subset. Results indicate the capability of both filter and wrapper methods, in combination with BBO, to achieve accurate and compact UIR.

3.4 Majority voting filter

The SVM classifier, which demonstrated good performance above, is now chosen to investigate the influence of MVF on prediction accuracy. In this experiment, SVM is trained for each

TABLE 3. MEAN CLASSIFICATION ACCURACY (PERCENT) OF SIX DIFFERENT CLASSIFIERS TRAINED FOR ABLE-BODIED SUBJECTS WITH THREE DIFFERENT FEATURE SUBSETS.

Classifier	Feature Selection Method		
	FLDS/BBO	LDA/BBO	MIFS
LDA	93.94 \pm 2.86	94.08 \pm 2.57	90.58 \pm 2.82
QDA	95.93 \pm 2.29	95.64 \pm 2.34	91.54 \pm 3.24
MLP-1	95.20 \pm 1.99	95.36 \pm 2.06	89.91 \pm 1.69
MLP-2	96.40 \pm 1.53	96.30 \pm 1.43	91.09 \pm 1.55
SVM	96.04 \pm 2.13	96.03 \pm 2.14	91.35 \pm 1.92
MLR	95.54 \pm 2.09	95.64 \pm 2.26	89.59 \pm 1.94
subset size	16	13	14

TABLE 4. CLASSIFICATION PERFORMANCE (PERCENT) OF SVM TRAINED WITH THREE DIFFERENT FEATURE SUBSETS FOR ABLE-BODIED SUBJECTS, WITH AND WITHOUT MVF.

Features	without MVF	with MVF
FLDS/BBO	95.67 \pm 2.23	98.96 \pm 0.90
LDA/BBO	95.75 \pm 2.38	99.04 \pm 0.87
MIFS	90.67 \pm 2.09	98.26 \pm 1.64

subject using four-fold CV where each of the four walking sequences is used as a fold. Table 4 shows the mean classification performance of SVM trained with three different feature subsets, with and without MVF. Table 4 confirms statistically significant improvement when MVF is used. MVF decreases classification error by approximately 77% when using SVM trained with a feature subset that was selected with LDA/BBO.

3.5 Misclassification distribution

To provide details on the misclassification distribution with and without MVF, the average confusion matrices for able-bodied subjects is shown in Table 5. The confusion matrices are shown for SVM trained with the LDA/BBO feature subset. Table 5 shows that SVM without MVF results in considerable misclassification among different walking activities, which may lead to sudden undesirable jumps between output classes, and eventually unsafe walking. However, the same classification method with MVF results in no misclassification of the three walking activities (SW, NW, and FW) with each other, and less than 1% misclassification of these three activity modes with ST. This minor misclassification appears at the transition between classes rather than appearing as a sudden jump in the middle of an activity

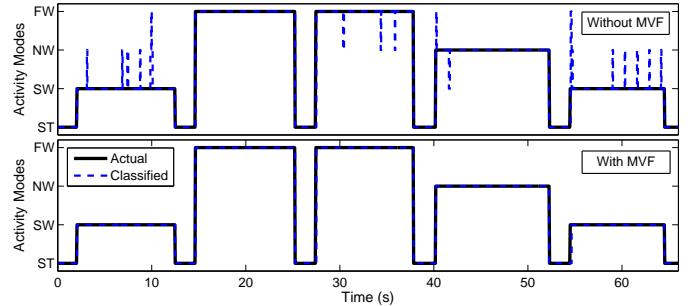


FIGURE 6. PREDICTION RESULTS OF A SAMPLE WALKING SEQUENCE CLASSIFIED WITH SVM, TRAINED WITH THE LDA/BBO FEATURE SUBSET, FOR AB01, WITH AND WITHOUT MVF.

mode. To further illustrate this point, Fig 6 shows the performance of SVM trained with the LDA/BBO feature subset for a test walking sequence for AB01.

Finally, we evaluate the performance of UIR trained individually with amputee subjects AM01 and AM02. We train and test QDA, and SVM with RBF parameter $\sigma = 4$, with and without MVF. Table 6 shows the mean classification accuracy for these classifiers trained with feature subsets obtained by FLDS/BBO, LDA/BBO, and MIFS. Results verify the superiority of BBO-based feature selection over MIFS. The best mean prediction accuracy is obtained with SVM trained with the FLDS/BBO feature subset, using MVF, and is equal to 98.94%. FLDS/BBO subset uses 16 features out of a total of 60 features, which reduces the size of feature set by 73.33%.

We compare our framework with other methods, although we note that the classification algorithm, the experimental setup, the gait modes, and the extracted features might be different. For example, [32] and [9] obtained 98% and 95% accuracy when UIR was trained with 52 and 56 features extracted from 13 and 14 measurement signals, respectively. SVM trained with the full feature set and without MVF results in a mean prediction accuracy of 98.36% for amputee subjects. In comparison, we achieve 97.20% accuracy with FLDS/BBO subset with 16 features. This indicates the satisfactory performance of our framework, which is able to eliminate unneeded features with no significant degradation in overall accuracy.

4 CONCLUSION

We collected experimental gait data from two able-bodied and two transfemoral amputee subjects. Several efficient TD and FD features are extracted from data frames using a sliding window technique. Poor features are excluded prior to applying formal feature selection methods. We combined filter and wrapper methods with BBO to obtain FLDS/BBO and LDA/BBO for optimal feature subset selection. The selected

TABLE 5. AVERAGE CONFUSION MATRICES OF SVM TRAINED WITH THE LDA/BBO FEATURE SUBSET FOR ABLE-BODIED SUBJECTS, WITH AND WITHOUT MVF. NOTE THAT THE OFF-DIAGONAL TERMS IN THE MVF CONFUSION MATRIX ARE ALL LESS THAN 1%.

		Classified Mode							
		without MVF				with MVF			
		ST	SW	NW	FW	ST	SW	NW	FW
Actual Mode	ST	99.71	0.00	0.00	0.29	99.81	0.00	0.00	0.19
	SW	0.43	97.56	1.65	0.36	0.67	99.33	0.00	0.00
	NW	0.23	1.68	96.76	1.33	0.47	0.00	99.53	0.00
	FW	0.37	0.00	1.94	97.69	0.43	0.00	0.00	99.57

TABLE 6. MEAN CLASSIFICATION ACCURACY (PERCENT) OF QDA AND SVM TRAINED INDIVIDUALLY WITH AM01 AND AM02, WITH FEATURE SUBSETS OBTAINED BY FLDS/BBO, LDA/BBO, AND MIFS.

Features	QDA		SVM	
	without MVF	with MVF	without MVF	with MVF
FLDS/BBO	91.81 ± 2.76	96.62 ± 2.50	97.20 ± 1.78	98.94 ± 0.93
LDA/BBO	89.82 ± 2.46	97.17 ± 1.29	93.72 ± 7.02	97.05 ± 4.87
MIFS	72.26 ± 6.08	74.54 ± 6.42	87.04 ± 8.29	94.16 ± 9.39

feature subsets were compared with MIFS feature subsets using several different classifiers enhanced with MVF. For able-bodied subjects, SVM trained with the LDA/BBO feature subset, and using MVF, resulted in the best accuracy of 99.04%. For amputee subjects, SVM trained with the FLDS/BBO feature subset, and using MVF, resulted in the best accuracy of 98.94%. Results indicate the effectiveness of BBO for construction of an optimal UIR system in terms of accuracy and simplicity. We note that the lower-limb amputee demographic at the Veterans Affairs Medical Center, where data collection was performed, is dominated by males. Over 98% of veterans who underwent amputation in 2011 were male [33]. This paper presented preliminary results to test the feasibility of our UIR system. Future work will have to include wider demographics (for example, ages and genders). Future work will also investigate the performance of UIR for other walking activities, and will measure the required input signals directly rather than with cameras (for example, piezo-electric force sensors and inertial measurement units). In addition, it is of great interest to compare BBO with other EA-based search strategies. Finally, combining UIR with prosthesis control switching in human subject tests is the ultimate goal of this research. The Matlab source code used to generate these results is available at <http://embeddedlab.csuohio.edu/prosthetics/research/user-intent-recognition.html>.

ACKNOWLEDGMENT

This research was supported by National Science Foundation Grant 1344954, and a Cleveland State University Graduate Student Research Award.

REFERENCES

- [1] Sup, F., Varol, H. A., and Goldfarb, M., 2011. “Upslope walking with a powered knee and ankle prosthesis: initial results with an amputee subject”. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, **19**(1), pp. 71–78.
- [2] Khademi, G., Richter, H., and Simon, D., 2016. “Multi-objective optimization of tracking/impedance control for a prosthetic leg with energy regeneration”. In IEEE Conference on Decision and Control, pp. 5322–5327.
- [3] Khademi, G., Mohammadi, H., Richter, H., and Simon, D., 2018. “Optimal mixed tracking/impedance control with applications to transfemoral prostheses with energy regeneration”. *IEEE Transactions on Biomedical Engineering*, **65**(4), pp. 894–910.
- [4] Tucker, M. R., Olivier, J., Pagel, A., Bleuler, H., Bouri, M., Lamberg, O., del R Millán, J., Riener, R., Vallery, H., and Gassert, R., 2015. “Control strategies for active lower extremity prosthetics and orthotics: a review”. *Journal of Neuroengineering and Rehabilitation*, **12**(1), p. 1.
- [5] Huang, H., Kuiken, T. A., and Lipschutz, R. D., 2009.

- “A strategy for identifying locomotion modes using surface electromyography”. *IEEE Transactions on Biomedical Engineering*, **56**(1), pp. 65–73.
- [6] Varol, H. A., Sup, F., and Goldfarb, M., 2010. “Multi-class real-time intent recognition of a powered lower limb prosthesis”. *IEEE Transactions on Biomedical Engineering*, **57**(3), pp. 542–551.
- [7] Khademi, G., Mohammadi, H., Simon, D., and Hardin, E. C., 2015. “Evolutionary optimization of user intent recognition for transfemoral amputees”. In IEEE Biomedical Circuits and Systems Conference, pp. 1–4.
- [8] Liu, M., Wang, D., and Huang, H. H., 2016. “Development of an environment-aware locomotion mode recognition system for powered lower limb prostheses”. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, **24**(4), pp. 434–443.
- [9] Stolyarov, R. M., Burnett, G., and Herr, H., 2017. “Translational motion tracking of leg joints for enhanced prediction of walking tasks”. *IEEE Transactions on Biomedical Engineering*, **in print**.
- [10] Young, A. J., Simon, A. M., Fey, N. P., and Hargrove, L. J., 2013. “Classifying the intent of novel users during human locomotion using powered lower limb prostheses”. In IEEE International Conference on Neural Engineering, pp. 311–314.
- [11] Ha, K. H., Varol, H. A., and Goldfarb, M., 2011. “Volitional control of a prosthetic knee using surface electromyography”. *IEEE Transactions on Biomedical Engineering*, **58**(1), pp. 144–151.
- [12] Huang, H., Zhang, F., Hargrove, L. J., Dou, Z., Rogers, D. R., and Englehart, K. B., 2011. “Continuous locomotion-mode identification for prosthetic legs based on neuromuscular–mechanical fusion”. *IEEE Transactions on Biomedical Engineering*, **58**(10), pp. 2867–2875.
- [13] Young, A., Kuiken, T., and Hargrove, L., 2014. “Analysis of using EMG and mechanical sensors to enhance intent recognition in powered lower limb prostheses”. *Journal of Neural Engineering*, **11**(5), p. 056021.
- [14] Simon, D., Omran, M. G., and Clerc, M., 2014. “Linearized biogeography-based optimization with re-initialization and local search”. *Information Sciences*, **267**, pp. 140–157.
- [15] Duda, R. O., and Hart, P. E., 1973. *Pattern classification and scene analysis*, Vol. 3. Wiley New York.
- [16] Gu, Q., Li, Z., and Han, J., 2012. “Generalized Fisher score for feature selection”. *arXiv preprint arXiv:1202.3725*.
- [17] Tang, J., Aleyani, S., and Liu, H., 2014. “Feature selection for classification: A review”. *Data Classification: Algorithms and Applications*, p. 37.
- [18] Kohavi, R., and John, G. H., 1997. “Wrappers for feature subset selection”. *Artificial Intelligence*, **97**(1-2), pp. 273–324.
- [19] van den Bogert, A. J., Geijtenbeek, T., Even-Zohar, O., Steenbrink, F., and Hardin, E. C., 2013. “A real-time system for biomechanical analysis of human movement and muscle function”. *Medical and Biological Engineering and Computing*, **51**(10), pp. 1069–1077.
- [20] Lara, O. D., and Labrador, M. A., 2013. “A survey on human activity recognition using wearable sensors.”. *IEEE Communications Surveys and Tutorials*, **15**(3), pp. 1192–1209.
- [21] Hudgins, B., Parker, P., and Scott, R. N., 1993. “A new strategy for multifunction myoelectric control”. *IEEE Transactions on Biomedical Engineering*, **40**(1), pp. 82–94.
- [22] Valls-Solé, J., Rothwell, J. C., Goulart, F., Cossu, G., and Munoz, E., 1999. “Patterned ballistic movements triggered by a startle in healthy humans”. *The Journal of Physiology*, **516**(3), pp. 931–938.
- [23] Stark, L., 2012. *Neurological control systems: Studies in bioengineering*. Springer Science & Business Media.
- [24] Phinyomark, A., Limsakul, C., and Phukpattaranont, P., 2009. “A novel feature extraction for robust EMG pattern recognition”. *arXiv preprint arXiv:0912.3973*.
- [25] Oskoei, M. A., and Hu, H., 2007. “Myoelectric control systems – a survey”. *Biomedical Signal Processing and Control*, **2**(4), pp. 275–294.
- [26] Englehart, K., and Hudgins, B., 2003. “A robust, real-time control scheme for multifunction myoelectric control”. *IEEE Transactions on Biomedical Engineering*, **50**(7), pp. 848–854.
- [27] Welling, M., 2005. Fisher linear discriminant analysis. Tech. rep., Department of Computer Science, University of Toronto.
- [28] Simon, D., 2013. *Evolutionary optimization algorithms*. John Wiley & Sons.
- [29] Battiti, R., 1994. “Using mutual information for selecting features in supervised neural net learning”. *IEEE Transactions on Neural Networks*, **5**(4), pp. 537–550.
- [30] Tourassi, G. D., Frederick, E. D., Markey, M. K., and Floyd Jr, C. E., 2001. “Application of the mutual information criterion for feature selection in computer-aided diagnosis”. *Medical Physics*, **28**(12), pp. 2394–2402.
- [31] Khademi, G., Mohammadi, H., and Simon, D., 2019. “Gradient-based multi-objective feature selection for gait mode recognition of transfemoral amputees”. *Sensors*, **19**(2), p. 253.
- [32] Young, A. J., Simon, A. M., Fey, N. P., and Hargrove, L. J., 2014. “Intent recognition in a powered lower limb prosthesis using time history information”. *Annals of Biomedical Engineering*, **42**(3), pp. 631–641.
- [33] Healthcare inspection prosthetic limb care in va facilities. Tech. rep., Department of Veterans Affairs Office of Inspector General.