

Topology Preserving Domain Adaptation for Addressing Subject Based Variability in SEMG Signal

Rita Chattopadhyay#, Narayanan C Krishnan+ and Sethuraman Panchanathan#

#Center for Cognitive Ubiquitous Computing,
School of Computing, Informatics, and Decision Systems Engineering,
Arizona State University, Tempe, Arizona, 85287, USA.

+ Center for Advanced Studies in Adaptive Systems,
School of Electrical Engineering and Computer Science,
Washington State University, Pullman, Washington, 99163, USA.

Abstract

A subject independent computational framework is one which does not require to be calibrated by the specific subject data to be ready to be used on the subject. The greatest challenge in developing such a framework is the variation in parameters across subjects which is termed as *subject based variability*. Spectral and amplitude variations in surface myoelectric signals (SEMG) are analyzed to determine the fatigue state of a muscle. But variations in the spectrum and magnitude of myoelectric signals across subjects cause variations in both marginal and conditional probability distributions in the features extracted across subjects, making it difficult to model the signal for any automated signal classification. However we observe that the manifold of the multidimensional SEMG data have an inherent similarity as the physiological state moves from no fatigue to fatigue state. In this paper we exploit this specific feature of the SEMG data and propose a domain adaptation technique that is based on intrinsic manifold of the data preserved in a low dimensional space, thus reducing the marginal probability differences between the subjects, followed by an instance selection methodology, based on similar conditional probabilities in the mapped domain. The proposed method provides significant improvement in subject independent accuracies compared to cases without any domain adaptation methods and also compared to other state-of-the-art domain adaptation methodologies.

Introduction

Characterization and quantification of muscle fatigue through noninvasive sensing mechanism such as surface electromyography (SEMG), is a challenging problem. While changes in the properties (shift in the power spectral density, root mean square, instantaneous frequency, etc) of SEMG signals with respect to muscle fatigue have been reported in the literature (Kumar, Pah, and Bradley 2003), (Georgakis, Stergioulas, and Giakas 2003), (Gerdle, Larsson, and Karlsson 2000), the large variation in these measures across different individuals makes the task of modeling SEMG difficult, and automating the process of signal classification as a generalized tool, complex. The variation

in SEMG parameters from subject to subject creates a difference in the data distribution (both in marginal and conditional probabilities) making it difficult to base any generalized framework directly on the traditional machine learning algorithms. In the quest to address this challenge and develop a generalized framework for detecting different stages of fatigue from SEMG signals, we propose a transfer learning framework based on a domain adaptation methodology, that addresses both marginal and conditional probability differences between the distributions.

We observe that the SEMG data collected continuously over time from a non-fatigue to a fatigue state forms a manifold with a well-defined topology that is consistent across different subjects. Traditional dimensionality reduction based domain adaptation techniques (PCA, or KPCA/KDA) ignore the topology of the data distribution, while adapting data from one domain (subject) to another. In the proposed approach we learn a new low dimensional feature space using a nonlinear dimensionality reduction technique (ISOMAP) which preserves the distribution topology (Joshua B. Tenenbaum 2000). In our approach, the low dimensional feature space maps marginal distribution of data corresponding to source and target subjects into a common kernel space and a sample selection strategy in the mapped domain draws conditional probabilities between the different subjects closer. Thus the proposed approach addresses both marginal and conditional probability differences.

We have validated our framework on SEMG signals collected from 10 people during a repetitive gripping activity. We extracted 12 amplitude and frequency domain features from the SEMG signal. Comprehensive experiments on the SEMG data set demonstrate that the proposed method improves the subject independent classification accuracy by 20% over the techniques that do not employ transfer learning methodologies and by 16% over other transfer learning methodologies.

Background and Related work

There are number of techniques that can be used to objectively determine the level of fatigue in a subject. The most reliable of these is the direct measurement of chemical properties in the muscle of the subject. Since this is an invasive technique it is inappropriate for routine utilization, away from the clinical environment. The electromyography

(EMG) is a biosignal recording of the skeletal muscle activity of the body. It is routinely used by clinicians for analysis of the skeletal muscle activity. EMG may be recorded from the surface of the skin without any invasion of the body known as surface EMG (SEMG). SEMG provides a non-invasive way to identify fatigue. Indication of localized muscle fatigue has been frequently based on the observed shift of the power spectral density of the SEMG (Kumar, Pah, and Bradley 2003), (Contessa, Adam, and Luca 2009), (Georgakis, Stergioulas, and Giakas 2003), (Lowery et al. 2000), (Koumantakis et al. 2001), (Sung, Zurcher, and Kaufman 2008). Several parametric measures of SEMG signal have been used as a relative indicator of the muscle fatigue phenomenon for an individual subject. These include the root mean square (rms), instantaneous frequency, zero crossing rate, mean-frequency, and median-frequency. In general there is a large variation in these measures due to variance in SEMG power spectrum and its shift for different subjects. Hence most of the work done in past towards quantification of fatigue from SEMG has been very subject specific. Contessa et al (Contessa, Adam, and Luca 2009) collected data from 4 subjects - once each at the beginning and conclusion of a fatiguing exercise and observed a significant difference between the data patterns collected from different subjects. Gerdle et al (Gerdle, Larsson, and Karlsson 2000) observed variations in the root mean square (RMS) of the EMG signal across different subjects even when the subjects performed the same activity under similar experimental conditions. Investigations into the use of physiological data for recognizing the different emotional states of a person have reported a subject independent classification accuracies of the order of 70% (Leon et al. 2007), (Kim and Andre 2008). To the best of our knowledge we propose the first systematic approach to address subject based variability in SEMG signals considering the sample distribution differences arising due to subject based variability.

The variation in SEMG parameters from subject to subject creates differences in the data distribution. This difference negatively affects the performance of traditional machine learning algorithms as it leads to differences in train and test data which are drawn from different subjects. In such cases domain adaptation methods have to be applied to reduce the differences in both conditional and marginal probabilities between the training and test data distributions. Many existing methods perform domain adaptation based on only marginal probability differences between the two data distributions. Shimodaira et al (Shimodaira 2000) biased the training samples by a test-to-training ratio to match the marginal distribution of the test data. Sugiyama et al (Sugiyama et al. 2008) tried to reduce the gap in marginal probabilities by minimizing the KL-divergence between test and weighted training data and Bickel et al (Bickel, Brückner, and Scheffer 2009) by discriminating training against test data with a probabilistic classifier. Huang et al (Huang et al. 2007) re-weight the instances in source domain so as to minimize the marginal probability difference, referred as *Kernel Mean Matching (KMM)*, using Maximum Mean Discrepancy (MMD) (Borgwardt et al. 2006) as the distance measure. Method suggested by Pan et al (Pan et al. 2009) is

based on feature mapping for reducing the marginal probability differences between the source and target distribution based on minimizing MMD, referred as *Transfer Component Analysis (TCA)*. The details of some of these techniques that have been implemented in this work are discussed under *Comparison with related work*.

There are also methods in the literature that address the differences in the conditional probability distributions. Gao et al (Gao et al. 2008) reduce this difference by proposing an approach which compares the clustering manifold of the test or target domain data around an instance with the manifold formed by the labels generated by the training or source domain data. This is restricted by the assumption that the test data follows a ‘clustering’ manifold. Zhong et al (Zhong et al. 2009) propose an approach that addresses both marginal and conditional probability differences (*KMapEnsemble (KE)*), based on domain mapping using Kernel Discriminant Analysis(KDA), followed by clustering based instance selection. KDA removes any topological properties of the data distribution that might aid in the classification process. Thus this technique is not suited for SEMG data that displays some topological properties that is explained earlier in *Introduction*.

This paper presents a methodology of domain adaptation that preserves the topology of the original data distributions while mapping both the source and target data into a common domain. This is followed by an effective methodology to select the training data having the same conditional probability without any assumption about any particular manifold of either the original or the mapped data distributions. To the best of our knowledge, manifold learning has been used in applications with large number of features such as image analysis, text analysis etc and this is the first application of this methodology for addressing differences across distributions for purposes of domain adaptation.

Proposed Approach

It is observed that SEMG data collected over a fatiguing exercise from different individuals shows variations in both conditional and marginal distributions. In spite of these differences, the SEMG data shows specific topological patterns that is consistent across different individuals. This pattern is illustrated in Figure 1.

In our proposed approach we exploit this similarity across subjects and use a domain adaptation methodology that preserves the topology of the input data distribution, to map the data from multiple subjects into a common domain. This process minimizes the differences in the marginal probability distributions across the subjects. We also observe that the SEMG data as shown in Figure 1 has significant differences in conditional probabilities across subjects, having conflicting conditional probabilities in several cases. The second part of the framework reduces this difference through an instance selection technique. To summarize the proposed framework is divided into two parts. The first part learns a new low dimensional feature space using a nonlinear dimensionality reduction technique called ISOMAP (Joshua B. Tenenbaum 2000), which preserves the topology of the input data distribution and the second part performs instance

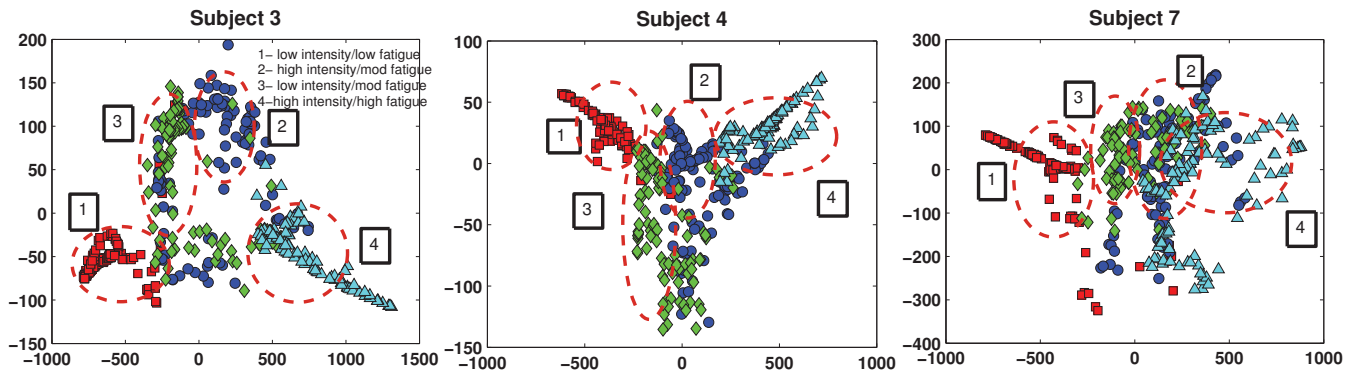


Figure 1: Three sample subjects (subjects 3, 4, 7) with four classes (four physiological stages) in our SEMG data set: Projected using Isomap based topology preserving methodology

selection in the mapped domain, based on conditional probability similarities between the distributions. Thus the proposed approach tries to reduce both marginal and conditional probability differences between the distributions.

Isomap based Domain Adaptation

Isomap (Joshua B. Tenenbaum 2000) is a manifold learning technique that extends the traditional multi-dimensional scaling by incorporating the geodesic distances imposed by a weighted graph. The Isomap algorithm takes as input the distances $d(i, j)$ between all pairs x_i, x_j from N data points in the high-dimensional input space X , measured in the standard Euclidean metrics and outputs coordinate vectors in a D -dimensional Euclidean space Y that best represents the intrinsic geometry of the data. The algorithm has three main steps. The first step constructs the neighborhood graph G over all the N data points by connecting points x_i and x_j . The weights associated with the edges connecting two points x_i and x_j is set to the Euclidean distance between these points $d_x(i, j)$. The second step computes the shortest paths between any two points by initializing $d_G(i, j) = d_x(i, j)$ if x_i, x_j are linked by an edge, $d_G(i, j) = \inf$ otherwise. Then for each value of $k = 1, 2, \dots, N$ in turn, replace all entries $d_G(i, j)$ by $\min\{d_G(i, j), d_G(i, k) + d_G(k, j)\}$. The final matrix $D_G = \{d_G(i, j)\}$ would contain all the shortest path distances between all pairs of points in G . Finally, the third step applies classical MDS (multi dimensional scaling) to the graph matrix D_G and constructs an embedding of the data in a D -dimensional Euclidean space Y that best preserves the manifold's estimated intrinsic geometry. The coordinate vectors y_i for points in Y or the mapped features in the new domain are obtained by minimizing the cost function:

$$\phi(Y) = \sum_{ij} (\|x_i - x_j\| - \|y_i - y_j\|)^2 \quad (1)$$

where $\|x_i - x_j\|$ is the Euclidean distance between the high-dimensional datapoints x_i and x_j and $\|y_i - y_j\|$ is the Euclidean distance between the corresponding low dimensional datapoints y_i and y_j . Thus the low dimensional representation of the data in isomap preserves the geodesic distances in the input data. Other classical dimensionality reduction algorithms such as PCA or MDS do not detect

the inherent manifold in the data. PCA finds a low dimensional representation that best preserves the variance of the data in the high dimensional input space and MDS finds an embedding that preserves the interpoint Euclidean distances. Isomap algorithm uses the MDS framework to preserve the geodesic distances as per the manifold of the high dimensional input data.

The proposed algorithm uses isomap to project the training and test data from multiple subjects to a common domain and then utilizes a K nearest neighbor based methodology to select instances from the training data which have similar conditional probability as the test data. Since in the proposed domain adaptation algorithm we address conditional probability as well, we require a few labeled samples from the target domain data D^T . The training data D^S consists of data from multiple subjects and the test or target domain data (data from a subject under test) consists of some labeled data D_l^T and lots of unlabeled data D_u^T such that $D^T = D_l^T + D_u^T$. The main steps of the proposed approach are as follows:

- Step 1: Compute the low dimensional projection of the available labeled test data D_l^T using Isomap methodology.
- Step 2: Use the mapping to project the training data D^S and unseen unlabeled test data D_u^T as well, into the same mapped space.
- Step 3: In the mapped domain, compute the Euclidean distance from each labeled test data in D_l^T to each of the data points in the training data D^S belonging to the same class.
- Step 4: In the mapped domain, sort the distances in increasing order of the value and select k nearest points from the training data D^S for each of the data samples in D_l^T to form $D_{selected}^S$.
- Step 5: Learn a classifier on the mapped selected training data $D_{selected}^S$ and mapped labeled target domain data D_l^T and compute the labels of the mapped test data D_u^T .
- Step 6: Compute new low dimensional projection again, using the selected data from training domain i.e. $D_{selected}^S$ and labeled test data D_l^T in the original space, and obtain the new mapping.

- Step 7: Go back to Step 2, for N number of iterations.
- Step 8: Compute the class declared majority number of times in N iterations for each of the D_u^T and assign the same to the data instance.
- Compute classification accuracy of the D_u^T .

Comparison with related work

We compare the proposed methodology with three recently published domain adaptation methods. One of the methods we compared our work with is based on marginal probability differences only (Pan et al. 2009), the second method with which we compare our work is based on conditional probability differences between the distributions (Gao et al. 2008) and the third domain adaptation method with which we compared our method addresses both marginal and conditional probability differences (Zhong et al. 2009) but does not use a topology preserving feature mapping methodology. We present here a short overview of each of these methods.

Transfer Component Analysis (TCA)

This method learns a kernel in the mapped domain such that the marginal distribution differences between the source (training) and target (test) domains is reduced using Maximum Mean Discrepancy defined as follows: As per this measure the difference between the two distributions $X_S = \{x_{S_1} \cdots x_{S_{n_1}}\}$ and $X_T = \{x_{T_1} \cdots x_{T_{n_2}}\}$ with distributions P and Q is given by

$$Dist(X_S, X_T) = \left\| \frac{1}{n_1} \sum_{i=1}^{n_1} \phi(x_{S_i}) - \frac{1}{n_2} \sum_{i=1}^{n_2} \phi(x_{T_i}) \right\|_H \quad (2)$$

where x_{S_i}, x_{T_i} are data from X_S and X_T , and H is a universal reproducible kernel Hilbert space (RKHS) (Steinwart 2002) and ϕ is the feature map induced by a universal kernel such that $\phi : X \rightarrow H$ and $\phi(x_{S_i})$ and $\phi(x_{T_i})$ are the corresponding mapped features into the RKHS. Pan et al proposed learning a feature map such that the $Dist(X_S, X_T)$ is minimized.

Locally Weighted Ensemble (LWE)

The LWE framework computes the label y of an unlabeled target domain data x , having k source models using the ensemble local weighing scheme as follows:

$$P(y|x) = \sum_{i=1}^k w_{M_i, x} P(y|M_i, x) \quad (3)$$

where $P(y|M_i, x)$ is the prediction made by one of the k models M_i for target data point x and $w_{M_i, x}$ is the weight of the model M_i at point x obtained as follows:

$$w_{M_i, x} = \frac{s(G_{M_i}, G_{M_T} : x)}{\sum_{i=1}^k s(G_{M_i}, G_{M_T} : x)} \quad (4)$$

where G_{M_i} and G_T are the graphs around point x . G_T is built by connecting points belonging to the same cluster manifold as x is, in the test data and G_{M_i} is built by connecting points belonging to the same class as x is assigned to, by the M_i model. $s(G_{M_i}, G_{M_T} : x)$ is the measure of similarity between these two graphs at point x in test data.

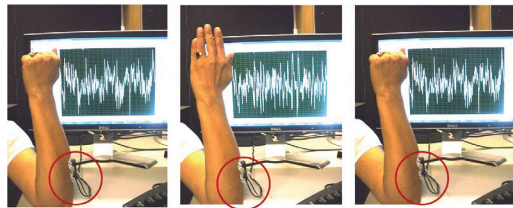


Figure 2: SEMG data collection during a repetitive gripping activity.

Kernel Ensemble (KE)

This method suggested by Zhong et al (Zhong et al. 2009) obtains the common domain between the training and test data by performing Kernel discriminant analysis (KDA) on the few sampled test data available and uses this feature mapping to map the training (source) and unseen test data (target) into the common domain, thus addressing the marginal probability difference between the source and target domain. It then applies a cluster based instance selection method to select training or source instances with similar conditional probability. There are two main draw backs with this method, one is it does not preserve the topology of the input high dimensional data in the new mapped domain and second it expects the mapped data to have a clustering manifold, where as in the proposed approach we preserve the original topology of the data, we do not assume the mapped data to be of any specific manifold.

Experiments

SEMG data

The SEMG data was collected during a repetitive gripping action performed by the forearm. Figure 2 shows the subject with surface EMG differential electrodes on the extensor carpi radialis muscle to record the SEMG signal. The subject performs a cycle of flexion-extension of forearm as shown in Figure 2 at two different speeds i.e. low speed (1 cycles/sec) and high speed (2 cycles/sec) repetitively for about 4 minutes. The cycles of low and high speed are alternated after every minute to form four phases or classes i.e. (1) *low intensity of activity and low fatigue*, (2) *high intensity of activity and moderate fatigue*, (3) *low intensity of activity and moderate fatigue* and (4) *high intensity of activity and high fatigue*. Figure 3 shows the raw SEMG signal obtained during the repetitive gripping for a representative subject.

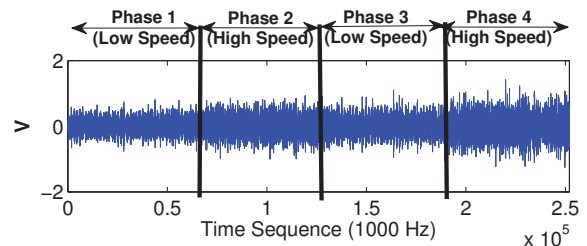


Figure 3: The raw continuous SEMG data of subject 4 performing repetitive movement for 4 minutes.

The raw SEMG activity was recorded by Grass Model 8-16C at 1000Hz and passed through a band pass filter of 20Hz to 500Hz. The data was collected and saved by the LabView software(from National Instruments)running on a PC. Data of the order of 1.92 Million samples ($1000*4*60*8$), was collected from 8 subjects including male and female of the age group of 25 years to 45 years. A set of twelve amplitude and frequency domain features are extracted from every burst localized by identifying the intersection between a linear envelope of 3 Hz and the mean line ($mean + 2 \times standard_deviation$) to identify the four phases of intensity and fatigue levels as mentioned by the class/phase definition. Each subject data consists of around 450 to 350 samples of 12 dimensional feature vectors, belonging to four classes with around 80 to 100 feature vectors per class.

Experimental Procedure

In order to evaluate the effectiveness of the proposed Topology Preserving Domain Adaptation (TPDA) methodology we compared the results with four baseline method SVM-C, SVM-M, SMA, TSVM (Transductive SVM) and with three recently published domain adaptation methods; namely Transfer Component Analysis (TCA), Locally Weighted Ensemble (LWE), and Kernel Ensemble (KE). The details of each of these techniques are provided under section heading *Comparison with related work*. We also compared our results with PCA based domain mapping followed by instance selection methodology as performed in TPDA, referred as PCA-DA. The definition of other methodologies are as follows: SVM-C refers to *all but one* method where in the training data comprises of data from all seven subjects and the test data is the data from the eighth subject. SVM-M, refers to *majority voting* based ensemble framework. The class y assigned to each unlabeled test data x is $\max_y NV(y|x)$ where $NV(y|x)$ is the number of votes given for class y for a particular test sample x by the seven auxiliary sources. SMA refers to *Simple Model Averaging*, which provides equal weight to all the classifiers learned on each auxiliary source domain in an weighted ensemble framework used to generate the label for the target domain data. SVM from LibSVM package was used as basic classifier for all these methods. TSVM is semisupervised method called Transductive SVM implemented using svmllight package.

For TCA, KE and the proposed method TPDA, 10% of the test data which is 6 to 7 samples per class, is made available for training purpose. Rest of the 90% of the test data was unseen test data. All the methods are tested on the same pool of unseen unlabeled test data. The accuracies are computed in a subject independent manner. In a particular experiment a subject data is considered as test data and the training data comprises of the data from the rest of the seven subjects. TCA was performed using linear kernel. The neighbourhood parameter used in K-Isomap and also in the KNN based instance selection algorithm was set to 12. The reduced dimension was kept as 6 for TCA, KE, TPDA and PCA-DA. These parameter values were selected on the basis of best performance as a result of 5 fold cross validation over the test training data. All the results presented are average over 10 rounds of execution with different sets of random data

selected as labeled test data for each test subject.

Results and Discussion

We use subject independent classification accuracy as metrics for performance evaluation. The results shown in Table 1 are obtained by implementing the methods SVM-C, SVM-M, SMA, TSVM, TCA, LWE, KE and the proposed TPDA methodologies. The first column of the Table 1 indicates the subject data under test. The training data consists of the data from rest of the seven subjects. We observe that the different weighting schemes for SVM, namely SVM-C, SVM-M and SMA result in poor performance. All these schemes pertain to the case where brute force transfer of knowledge is done causing negative transfer (Rosenstein, Marx, and Kaelbling 2005). We also observe that results obtained by implementing TCA which addresses only marginal probability differences are comparable to TSVM, which does not perform any domain adaptation, instead employs a label propagation technique to determine the labels of the target domain data. We also observe that LWE which addresses conditional probability differences only between the domains gave better results than TCA and KE for test subjects 2,3 and 4. LWE gave better results than TPDA for subject 2 and comparable for subject 4. These results show that considering conditional probability difference is paramount for addressing the distribution difference across subjects in SEMG data. We also observe that the overall or average classification accuracies obtained when both marginal and conditional probabilities are addressed by methodologies KE and TPDA are better than any of the other methodologies implemented. However TPDA outperforms KE for all cases with an average performance improvement of around 7.6%. We also implemented PCA based domain adaptation followed by instance selection as per proposed methodology, known as PCA-DA, for comparison with the proposed methodology, as PCA is by far the most popular linear technique. The results show that for SEMG data, preserving input data topology, significantly improves the average subject independent classification accuracies.

Thus we observe that the proposed Topology Preserving Domain Adaptation method (TPDA) provides a gain of 21% to 32% over the methods with out any domain adaptation (SVM-C, SVM-M, SMA and TSVM) and 8% to 22% over the state-of-the-art domain adaptation methods (TCA, LWE and KE). We also observe that standard deviation of the results across the subjects is minimum for TPDA. This also shows that the proposed method is able to address better the variations in distributions across the subjects and present a less variant common representation for the SEMG data.

Conclusion and Future work

We propose a new domain adaptation method to addresses subject based variability in myoelectric signals so as to develop a generalized framework for measurement of physiological status across subjects. To the best of our knowledge, this is the first systematic analysis of the SEMG data based on distribution differences across subjects, also this is the first application of topology preserving non linear dimensionality reduction technique, Isomap, for domain adapta-

Table 1: Comparative Performance of Proposed Method (TPDA) on SEMG data - Accuracy (%)

| Test Sub | SVM-C | SVM-M | SMA | TSVM | TCA | LWE | KE | PCA-DA | TPDA |
|----------|-------|-------|-------|-------|-------|-------|-------|--------|-------|
| 1 | 70.76 | 33.9 | 44.96 | 49.09 | 45.15 | 67.44 | 71.85 | 66.81 | 82.12 |
| 2 | 43.69 | 50.76 | 44.61 | 55.68 | 68.93 | 77.54 | 74.62 | 59.46 | 75.76 |
| 3 | 50.11 | 56.85 | 56.84 | 65.09 | 56.78 | 75.55 | 74.79 | 65.79 | 83.96 |
| 4 | 59.65 | 47.93 | 49.67 | 56.98 | 52.68 | 81.22 | 69.35 | 78.02 | 81.75 |
| 5 | 40.37 | 44.79 | 50.15 | 62.5 | 60.15 | 52.48 | 73.44 | 67.29 | 83.20 |
| 6 | 59.21 | 61.45 | 60.33 | 71.32 | 76.92 | 65.77 | 73.92 | 66.08 | 82.32 |
| 7 | 47.13 | 46.91 | 45.76 | 60.73 | 55.64 | 60.32 | 77.97 | 72.46 | 83.33 |
| 8 | 69.85 | 64.53 | 74.46 | 68.55 | 67.24 | 72.81 | 79.48 | 68.34 | 84.13 |
| Average | 55.09 | 50.85 | 53.34 | 61.24 | 60.43 | 69.14 | 74.42 | 68.03 | 82.07 |
| Std_Dev | 11.53 | 9.88 | 10.24 | 7.25 | 10.14 | 9.56 | 3.20 | 5.39 | 2.69 |

tion. Here we apply Isomap to obtain the mappings of the data in a common domain which preserves the topology of the input data. This feature being reasonably similar across SEMG data from different subjects, provide the best mapping for domain adaptation. The proposed technique is scalable to number of subjects and number of classes without affecting the performance. Experimental results show that the proposed method is very effective in addressing subject independent classification accuracy of the generalized framework. Success in developing the generalized model for measuring fatigue based on SEMG signals would open doors to many broader applications of other physiological data such as EKG, EEG, pulse rate besides SEMG itself in applications related to health monitoring in every day movement, industrial work and geriatric care.

References

- Bickel, S.; Brückner, M.; and Scheffer, T. 2009. Discriminative learning under covariate shift. In *JMLR*, volume 10, 2137–2155.
- Borgwardt, K. M.; Gretton, A.; Rasch, M. J.; Kriegel, H. P.; Scholkopf, B.; and Smola, A. J. 2006. Integrating structured biological data by kernel maximum mean discrepancy. *Bioinformatics* 22(14):49–57.
- Contessa, P.; Adam, A.; and Luca, C. J. D. 2009. Motor unit control and force fluctuation during fatigue. *Journal of Applied Physiology*.
- Gao, J.; Fan, W.; Jiang, J.; and Han, J. 2008. Knowledge transfer via multiple model local structure mapping. In *KDD*, 283–291.
- Georgakis, A.; Stergioulas, L.; and Giakas, G. 2003. Fatigue analysis of the surface EMG signal in isometric constant force contractions using the averaged instantaneous frequency. *IEEE Transactions on Biomedical Engineering* 50(2):262–265.
- Gerdle, B.; Larsson, B.; and Karlsson, S. 2000. Criterion validation of surface EMG variables as fatigue indicators using peak torque: a study of repetitive maximum isokinetic knee extensions. *Journal of Electromyography and Kinesiology* 10(4):225–232.
- Huang, J.; Smola, A. J.; Gretton, A.; Borgwardt, K. M.; and Scholkopf, B. 2007. Correcting sample selection bias by unlabeled data. In *NIPS*, 601.
- Joshua B. Tenenbaum, Vin de Silva, J. C. L. 2000. A global geometric framework for nonlinear dimensionality reduction. *SCIENCE*.
- Kim, J., and Andre, E. 2008. Emotion recognition based on physiological changes in music listening. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 30(12):2067–2083.
- Koumantakis, G. A.; Arnall, F.; Cooper, R. G.; and Oldham, J. A. 2001. Paraspinal muscle EMG fatigue testing with two methods in healthy volunteers. reliability in the context of clinical applications. *Clinical Biomechanics* 16.
- Kumar, D.; Pah, N.; and Bradley, A. 2003. Wavelet analysis of surface electromyography. *IEEE Transactions on Neural Systems and Rehabilitation Engineering* 11(4):400–406.
- Leon, E.; Clarke, G.; Callaghan, V.; and Sepulveda, F. 2007. A user independent real time emotion recognition system for software agents in domestic environment. *Engineering Application of Artificial Intelligence* 20(3):337–345.
- Lowery, M.; Vaughan, C.; Nolan, P.; and O’Malley, M. 2000. Spectral compression of the electromyographic signal due to decreasing muscle fiber conduction velocity. *Rehabilitation Engineering, IEEE Transactions on* 8(3):353–361.
- Pan, S. J.; Tsang, I. W.; Kwok, J. T.; and Yang, Q. 2009. Domain adaptation via transfer component analysis. In *IJCAI*.
- Rosenstein, M.; Marx, Z.; and Kaelbling, L. 2005. To transfer or not to transfer. In *NIPS 2005 Workshop on Inductive Transfer: 10 Years Later*.
- Shimodaira, H. 2000. Improving predictive inference under covariate shift by weighting the log-likelihood function. In *JSPI*.
- Steinwart, I. 2002. On the influence of the kernel on the consistency of support vector machines. In *JMLR*, volume 2, 93.
- Sugiyama, M.; Nakajima, S.; Kashima, H.; Buenau, P. V.; and Kawanabe, M. 2008. Direct importance estimation with model selection and its application to covariate shift adaptation. In *NIPS*.
- Sung, P. S.; Zurcher, U.; and Kaufman, M. 2008. Reliability difference between spectral and entropic measures of erector spinae muscle fatigability. *Journal of Electromyography and Kinesiology* In Press, Corrected Proof.
- Zhong, E.; Fan, W.; Peng, J.; Zhang, K.; Ren, J.; Turaga, D.; and Verscheure, O. 2009. Cross domain distribution adaptation via kernel mapping. In *KDD*. ACM.