Integration of Motion Capture and EMG data
for Classifying the Human Motions

Gaurav N. Pradhan, Navzer Engineer, Mihai Nadin, Balakrishnan Prabhakaran
Department of Computer Science
University of Texas at Dallas, Richardson, TX 75083
{gnp021000, navzer, nadin, praba}@utdallas.edu

Abstract

Three dimensional motion capture facility is a powerful tool for quantitative and qualitative assessment of multi-joint external movements. Electro-myograph (EMG) signals give the physiologic information of muscles while doing motions. In this paper, our objective is to integrate these two different bio-medical data together and to extract precise and accurate feature information for classifying the human motions. When both forms of data are integrated and analyzed together, the information achieved will be immensely useful to quantify the complex human motions for medical reasons or sport performances. These biological quantifications of biomechanical data, are useful for gait analysis and several orthopedic applications, such as joint mechanics, prosthetic designs, and sports medicines.

The different dimensionality reduction approaches such Integral of Absolute value and Weighted Singular Value Decomposition are used to extract the preliminary features from EMG and motion capture data respectively. On combining these feature vectors, fuzzy clustering such as Fuzzy c-means (FCM) is performed on these vectors that are mapped as the points in multi-dimensional feature space. We get the degree of memberships with every cluster for each mapped point. This extracted information is used as the final feature vectors for classifying the human motions.

1. Introduction

Motion Capture is the process of recording a live human motion event and translating it into three-dimensional positional and orientation information of joints in space over time. The EMG signal is a biomedical signal that measures the electric currents generated in muscles during contractions that occur while performing the motions. When both these information are integrated and analyzed together, the information achieved will be immensely useful to quantify the complex human motions for medical reasons or sport performances. These biological measuring of biomechanical data, are useful for gait analysis and several orthopedic applications, such as joint mechanics, prosthetic designs, and sports medicines.
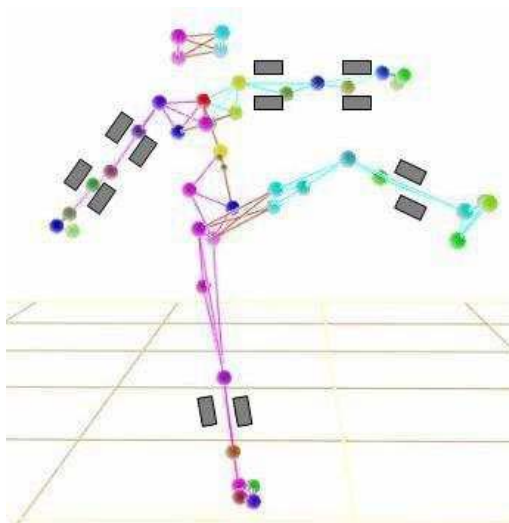

Figure 1. Human motion capture data is captured using the reflectors (round-shaped) on the body and EMG activity is measured simultaneously using EMG electrodes (rectangular-shaped, gray-color).

Figure 1 shows the retro-reflective markers (round-shaped) on the participant's body in the 3D space. As the participant keeps moving, the cameras track the movement

of these markers and give the exact position and orientation of the joints/segments in the 3D space. Hence, every motion is represented by a matrix which contains the 3D positional information for all joints, in the form of 3-column per joint (called as "joint matrix") in whole motion matrix. Also, EMG electrodes (rectangular-shaped and gray in color) are attached to the limbs as shown in Figure 1. Since human motions are mostly natural activities, semantically similar motions such as walking can have large variations in EMG signals. The data acquisitions of the both sensors is triggered at the same time when the participant starts performing action.

The main interest of this paper lies in understanding the relationship between the motion capture and EMG data for the different kind of human motions. Once the nature and characteristics of this relation is understood, this collaboration of information will be useful for clinical diagnosis and biomedical applications. Until now, extensive research has been made on EMG signal analysis, processing, and pattern classification on only EMG etc. Also, in the field of human motion database, similarity matching of motions, indexing, content-based retrieval of human motions are being done. But as far our knowledge goes, integration the motion capture and EMG data and then analyzing both to classify the human motions is never being tried. In this paper, we make an effort to develop a motion classification technique that depends on both kinds of data.

Figure 2 shows the sample synchronous EMG and motion capture data. The participant is performing the raising of arm on instruction. In third figure, we can see the 3D positional trajectory of the wrist motion in the 3D-space while raising arm. While he raises the arm, there are muscle activities in upper-arm and forearm that are captured by biceps and upper forearm EMG electrodes/sensors respectively. Thus, seeing motion capture along with EMG data gives us the better picture internally as well as externally for analyzing the motions. Though both EMG and Motion Capture are synchronous while capturing they have different properties that make them difficult to have unique feature extraction technique. The differences between the two data are following

- The EMG data is more non-stationary in nature. It depends on anatomical and physiological properties of muscles whereas motion capture data depends on physical movements of the joints while performing motions.

- The EMG data measures the electric currents generated in muscles during contraction, while motion capture data measures the 3D positional (and rotational too, but we will be neglecting them in our current work) values of each joints during performing action.

- The resolution of the EMG data is in mV (millivolts), and the resolution of the motion capture data is mm (millimeters).

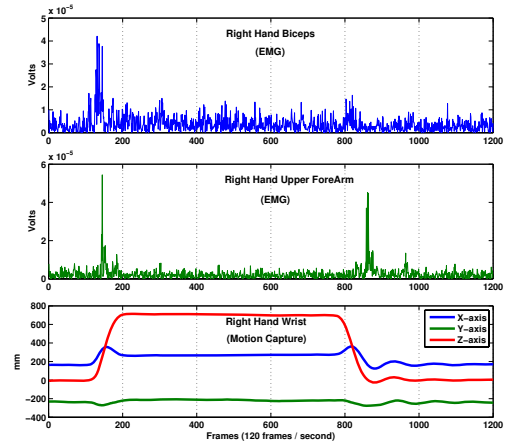- The EMG data is not at all immune to the noise like motion capture data.

Figure 2. The 3D-motion trajectory of the wrist joint and corresponding muscle activity in biceps and upper forearm while raising as arm.

The EMG signal, is more complicated signal, as it is controlled by nervous system. The two similar motions performed, even by assuming with same local speed, it is not necessary that EMG signal will be similar. And of course, vice-versa is equally true. Thus, if we consider two motions equidistant from the mean of the cluster of similar motions in feature space, they may be significantly different form each other. And even, two motions given in a cluster may be similar even though they are far away from each other in feature space. Thus, in biomedical data such as motion capture and especially EMG, the boundaries between the classes of motions are not sharply defined.

To overcome the above differences we use different feature extraction techniques for EMG and motion capture data. We use the sliding window approach to extract the features from motion matrix data. To get a final feature vector corresponding to a window of a motion, we combine these two sets of features and map it as a point in multi-dimensional feature space, which is combination of EMG and motion capture feature space. We perform the fuzzy clustering such as fuzzy c-means (FCM) on these mapped points to generate the degree of memberships with every cluster for each point. Due to non-stationary property of the

EMG signal, fuzzy clustering has an advantage over traditional clustering techniques. For a given motion, highest degree of membership for each cluster among all the divided windows of a motion becomes the final feature vector for the given motion. The separability of these feature vectors among different motions depends on the fuzzy clustering. This extraction technique projects the effect of both motion capture and EMG in a single feature vector for the corresponding motion.

2. Related Work

[10] constructed qualitative features describing geometric relations between specified body points of a pose and uses these features to induce a time segmentation of motion capture data streams for motion indexing. For each query, a user has to select suitable features in order to obtain high-quality retrieval results. The posture features of each motion frame are extracted and mapped into a multidimensional vector in [3] for motion indexing. These methods are more posture specific; and matching in two motions is carried by indexing first and last frames that may not be same for most of the similar motions, affecting similarity results. In [9], the authors use a hierarchical motion description for a posture, and use clustering-based key-frame extraction for retrieving the motions. To extract key-frames they need to find similarity between each consecutive frame, which is time consuming. Similarly Keogh et al. [8] used bounding envelops for similarity search in one attribute time series under uniform scaling. Also, lot of work is proposed on retrieving nearest- neighbors for the queries in multi-attribute data repository. The iDistance [14] is a distance-based index structure; here dataset is partitioned into clusters and transformed into lower dimension using similarity with respect to reference point of cluster. MUSE [13] extends [14] where partitioning of dataset at each level of the index tree is based on the differences between corresponding principal components.

In the past decades, much research has been done on the recognition of EMG signals, most of which is reviewed in [12]. The researchers, has investigated various techniques to extract the feature vectors which include zero-crossing [7], EMG Histogram [15], coefficients of an EMG autoregressive model [5] etc., Another classification techniques used to classify the EMG signals are neural networks [1], fuzzy systems [2], fractal analysis [6] etc.

3. Feature Extraction

Using motion capture facility and Myomonitor EMG facility, the external 3D positional information of the human segments and the electric current flowing through muscles internally is captured/measured synchronously for the performed motions. Both data characterize the motion in different format, but they definitely give more information when they are analyzed together than analyzed separately. Our goal is to extract the desirable features from both form of data and to estimate the human motion precisely. To start with, we extract the corresponding preliminary features from motion capture and EMG separately using the techniques, which are discussed in shortly. Then using our approach, we combine these two different preliminary extracted features to come with a single final feature vector which can then reflect the effect of motion capture and EMG in the combined single feature space.

3.1. EMG Data

The EMG signals are acquired using surface electrodes attached at the skin surfaces. Each electrode measures the electric flow in associated muscles. We follow a traditional measure to extract the feature of the EMG using the Integral of Absolute Value (IAV). We calculate IAV separately for individual channel. Each channel is defined by each EMG sensor. Let x_i be the ith sample of an EMG signal/data and N be the window size for computing the feature components. In a stream of EMG signal let IAV_j be the *Integral of Absolute Value* of jth window of EMG which is calculated as,

$$IAV_j = \frac{1}{N} \sum_{i=(j-1)*N+1}^{j*N} | x_i | \qquad (1)$$

3.2. Motion Capture Data

With the global positions, it becomes difficult to analyze the motions performed at different locations and in different directions. Thus, we do the local transformation of positional data for each body segment by shifting the global origin to the pelvis segment because it is the root of all body segments.

An appropriate mapping function is required to map 3D motion joint matrices into 3D feature points in the feature space. In our implementation, we used the linearly optimal dimensionality reduction technique SVD [4] for this purpose. For any $l \times 3$ joint matrix A and window size N, the SVD for the $j^t h$ window is given as follows,

$$A_{N \times 3}^j = U_{N \times N}^j \cdot S_{N \times 3}^j \cdot V_{3 \times 3}^j \qquad (2)$$

S^j is a diagonal matrix and its diagonal elements are called singular values. And columns of V^j matrix are called right singular vectors. We add up the three right singular vectors weighted by their associated normalized singular values to

construct the features for a joint motion's j^{th} window as follows:

$$f_c^j = \sum_{i=1}^{3}(w_i \cdot v_{ci}) \qquad (3)$$

where $w_i = \frac{\rho_i}{\sum_{j=1}^{3}\rho_j}$, $\sum_{i=1}^{3}w_i = 1$, $c = \{1,2,3\}$, and $[\rho_1, \rho_2, \rho_3]$ is singular value vector and v_{ci} is the cth component of the ith right singular vector and w_i is the normalized weight for the ith right singular vector. The weighted joint feature vector of length 3 represents the contribution of the corresponding joint to the motion data in 3D space for the jth window of length N and also captures the geometric similarity of motion matrices.

3.3. Combining Feature Vectors

Having extracted the feature vectors for each window from motion capture and EMG, the next step is to combine them by appending one to other. Thus, m-length EMG feature vector (i.e. a point in m-dimensional EMG feature space) and n-length motion capture feature vector (i.e. a point in n-dimensional motion capture feature space) form a $(m+n)$-length feature vector represented as a point in $(m+n)$-dimensional feature space. Since the EMG data is non-stationary in nature, it introduces vagueness in the feature vectors. Vagueness is a problem that requires a fuzzy approach/solution to handle. We use *fuzzy c-means* clustering (FCM) algorithm [11] to cluster the points in $(m+n)$-d or k-d(let $k = m + n$) feature space where each point represents the combined feature vector for each window.

The FCM on all points M in k-d feature space is given by,

$$[center_{c \times k}, U_{M \times c}, objFcn] = FCM(M, c) \qquad (4)$$

where c is the pre-determined number of clusters we are interested into, the performance of the classification varies on choice of cluster numbers which will be discussed later in Section 6. $center$ gives the center/median points for all c clusters in k-d space and matrix U gives the degree of membership for each M points(i.e. windows) with respect to each cluster. $objFcn$ contains a history of the objective function across the iterations which is of least interest in our approach.

Each motion of length say l, is divided into l/N windows and thus mapped as l/N points in k-d feature space. Using fuzzy c-means clustering each of these points has degree of membership with the fuzzy clusters. Thus, for each l/N points we choose highest degree of membership and its corresponding cluster. The highest degree of membership indicates that the point is more closer to the corresponding cluster than other clusters. Thus, for the given motion

which is represented in form of l/N feature points in k-d feature space, we have final feature vector corresponding to this motion in form of the maximum and minimum of the highest degree of membership for each cluster.

Consider a motion R, each window of motion R has a corresponding point in k-d space. And we get a degree of membership with all clusters c for each point in matrix U using FCM. Thus, for each window/point i in motion R, we get the highest degree of membership (hd_i) with cluster $closestCluster_i$ as follows,

$$
\begin{aligned}
hd_i &= max(\forall k : U_{ik}) \qquad &(5) \\
closestCluster_i &= argmax(\forall k : U_{ik}) \qquad &(6)
\end{aligned}
$$

The final feature components of motion R corresponding to cluster j i.e. $[ff_{j,max}, ff_{j,min}]$ are given as follows, for all j,

$$\forall i : ((closestCluster_i = c_j) \wedge (ff_{j,max} = max(hd_i))) \quad (7)$$
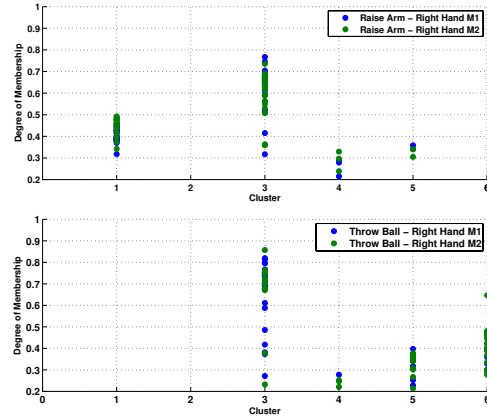$$\forall i : ((closestCluster_i = c_j) \wedge (ff_{j,min} = min(hd_i))) \quad (8)$$

Figure 3. The range of highest degree of information per cluster for two sets of two similar motions "Raise-Arm" and "Throw Ball" with cluster c = 6.

Figure 3 illustrates the previous discussion using two sets of two similar right hand motions. In motion $Raise-Arm$, $M1$ has 0.78 as maximum of the highest degree of membership with Cluster 3 and 0.32 as the minimum of the highest degree of membership. Thus, for $M1$, feature component corresponding to Cluster 3 are $[0.32, 0.78]$. Similarly for all other motions with different clusters. Thus the length of the final feature vector is $2*c$ where c is the number of clusters.

From Figure 3 we see that, all the points corresponding to windows in motion $Raise - Arm$ are near Cluster 1, 3, 4, 5, and for motion $Throw - Ball$ near Cluster 3, 4, 5 and 6. Figure 4 shows the final feature vectors for two sets of similar motions.
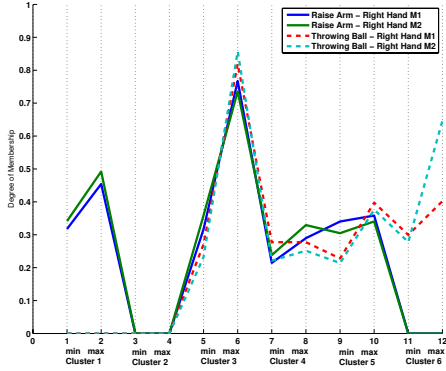

Figure 4. The feature vectors for the two sets of similar motions corresponding to the motions shown in Figure 3.

4. Searching for similar motions (Classification)

We perform content-based retrieval for the given query matrices(EMG + Motion Capture) from our database. For the given query matrix, we transform it into query feature vector so that it becomes compatible to search in our database by just comparing with low-dimensional feature vectors of motions in database.

The query matrix Q which includes EMG and motion capture, we use the same window size N to transform query EMG and motion capture data into feature vectors. To transform the query matrix of EMG and motion capture we use the same techniques discussed in Sub-Section 3.1 and Sub-Section mocap respectively. Let q_i be the query combined feature vector by combining both EMG and motion capture feature vectors for i^{th} window in query matrix Q.

To get final feature vector corresponding to the query motion, we need to know the maximum and minimum of highest degree of membership with each of the c clusters which are formed by applying FCM on the existent motions in the database. For each window i of query, we get the degree of membership u_{ik} with k^{th} cluster as follows,

$$u_{ik} = \frac{1}{\sum_{j=1}^{c} \left(\frac{||q_i \bullet center_k||_2}{||q_i \bullet center_j||_2} \right)^{\frac{2}{m-1}}} \quad \forall i, k \qquad (9)$$

where $center_j$ is the centroid of the cluster k, while $|| \bullet ||_2$ is the euclidean distance expressing the similarity between query feature point and the center. According to [11], parameter m is chosen in range of $[1.1, 5.0]$, with $m = 2$ in most applications of FCM. Hence, we choose $m = 2$ as it is most widely used.

Having know the degree of membership with all clusters, the final feature vector corresponding to the query is determined the same way as seen in Sub-Section 3.3. We can use any searching technique like linear search to get the nearest neighbors and to classify the query motion. The main goal of this paper is to retrieve the correct matching motion and to do classification on the query motion. For fast searching, our extracted feature vectors can be applied to any indexing technique to prune irrelevant motions.

5. Experimental Procedures

The human motion data was generated by capturing human motions in our Motion Capture Laboratory. This laboratory has 16 high-resolution Vicon cameras connected to a data station running Vicon iQ software. There are 18 3D-attributes for the local transformed motion capture data, and each attribute represents the positional values of one joint of a moving subject. EMG Ag electrodes are used to pick the muscle signals of limbs while performing motions. On each hand, four electrodes are placed mainly on biceps, triceps, upper forearm, and lower forearm. On each leg, two electrodes are placed on front side of shin and on backside of shin. The EMG signals are amplified and band-pass filtered (20-450 Hz) by Delsys Myomonitor system. Te sampling rate is 1000 samples second. This, processed signal is full-wave rectified and down-sampled to 120 Hz to make it uniform with the motion capture system which works at 120 samples per second. Motion capture and EMG system has to be synchronized along time axis, i.e. both has to start their acquisitions of data at the same time when participant starts to perform. Figure 5 shows a circuitry which we designed to make both systems work synchronously with each other using Delsys designed "Trigger Module".

MATLAB behaves as a main controller that sends a trigger to EMG and motion capture to start acquisitions through trigger module. The trigger module is attached to the parallel port of workstation. And MATLAB communicates with trigger module via parallel port using Data Acquisition Toolbox.
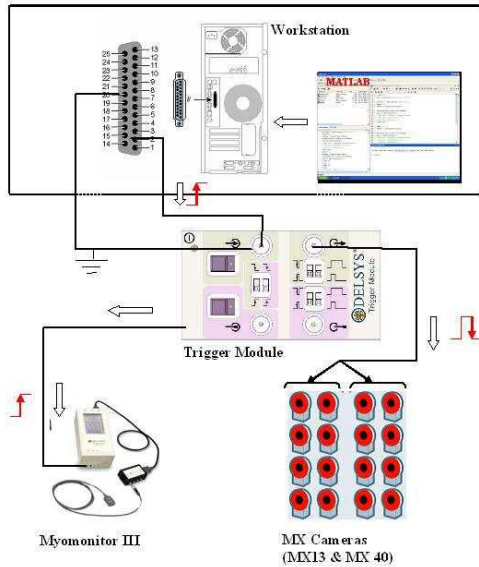
Figure 5. The hardware circuit for synchronizing motion capture and EMG system.

Our test bed consists of different human motions performed by different participants. We analyze differently for upper limbs and lower limbs though our approach is flexible enough to classify the human motions for whole human body. To analyze just one limb makes more sense in prosthetic control and medical rehabilitation of single limb. Thus, when we consider hand, we have four attributes from motion capture in form of human body segments such as clavicle, humerus, radius and hand; and four attributes from EMG such as biceps, triceps, upper forearm and lower forearm. Also, when we consider leg, we have three segments tibia, foot, toe, and two attributes from EMG such as front and back shin. The window sizes were made variable from 50ms to 200ms to study the effect on classification rate, discussed next.

6. Performance Evaluation

To quantify the suitability of the feature vectors mapped in feature space, we evaluate the system performance by two ways. The first way is for certain amount of queries, we check whether the query motion is correctly classified or not. In this case, we measure the average misclassification rate by varying window size and pre-determined clusters used in FCM. The second way to find k-Nearest Neigh-

bors for the given query motion and to check the percentage of returned motions in k which are actually present in the same group of query motion. The other returned motions are false alarms. Thus, we measure the performance of our approach by doing experiments on captured data and varying the parameters like,

- Window Size in range of $50ms - 200ms$,

- Number of Clusters c in range of $4 - 40$.
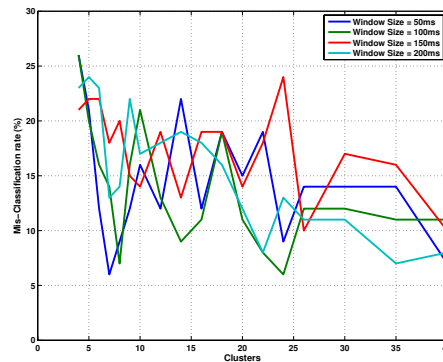
6.1. Mis-classification Rate

Figure 6. Percent of trials Misclassified for the Right Hand.

Figures 6 and 7 shows the effect of window size and change in number of clusters on the average percent of mis-classified trials/queries. Since we analyze the hands and legs separately, Figures 6 and 7 shows the performance of classification on right hand and right leg respectively. The mis-classification is generally between 10-20% for the number of clusters between 10-25, for both feature spaces (hand and leg). The overall mis-classification rate decreases, as number of cluster increases. It is more clearer in right leg trials when number of clusters are large that most of the time mis-classification rate increases by small amount when window size increases.

6.2. k-NN feature space classifier

Retrieving the k-closest motion from the database for the given query motion is a good non-parametric classifier. For all set of experiments k is set to 5. Figures 8 and 9 shows the percentage of correctly classified motions among the k
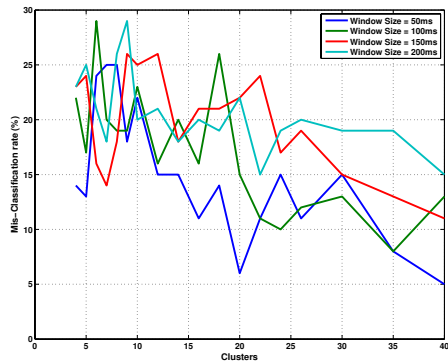
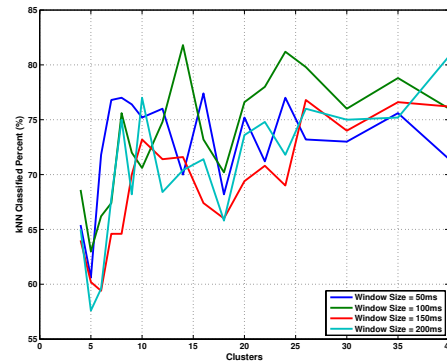**Figure 7. Percent of trials Misclassified for the Right Leg.**

Figure 8. Percent of correctly classified Right Hand motions out of k(= 5) motions retrieved for the right hand query motion.

retrieved motions from the database for the right hand and right leg respectively. Figure 9 clearly shows that as the window size goes on increasing more number of correctly classified motions are retrieved. Also, as the number of clusters goes on increasing, the percentage of kNN classified also goes on increasing. Thus, by analyzing the Figures 6 and 8, we can see that for large number of clusters and larger window size, mis-classification rate decreases and more number of correct closest neighbors are retrieved from the database.

7. Discussion and Conclusions

We have two biomedical signals to measure from the human while he/she is performing; first is the 3D Motion capture data which gives the 3D positional information of all joints and second, the EMG data which records the electric current flowing through muscles due to muscle-contraction while performing motions. In this paper, we discussed a technique to extract the feature vectors, which reflect the characteristic nature of motion capture as well as EMG muscle data on that motion. Since both have different properties, different dimensionality reduction approach is used to extract the preliminary features from them. Integral of Absolute value and Weighted Singular Value Decomposition techniques are used for EMG and motion capture respectively. On combining these feature vectors, fuzzy clustering such as fuzzy c-means (FCM) is performed on these vectors, which are mapped as point in multi-dimensional feature space. We get the degree of memberships for each mapped point with every cluster. This information extracted

becomes the final feature vector for individual motions. The length of the feature vector is in the order of the number of clusters selected for clustering. Fuzzy logic is used because contradictions in the data can be tolerated. Also, it is possible to discover the combined patterns of motion capture and EMG that are not easily detected by other methods. After the extraction of the feature vectors any similarity searching technique can be used to find the nearest neighbors or for classification of the motions.

Our experiments show that mis-classification rate is mostly in between 10-20%, which is understandable due to uncertainty of biomedical data and they are prone to noise. Some other unwanted environment effects such as signal drift, change in electrode characteristics, signal interference may affect the data. Also, other bio-effects such as subject training, fatigue, nervousness etc., can cause the purity of the biomedical signals. We also analyzed, the k-NN feature space classifier to check among the k- most nearest neighbors how many are exact match or how many are there in same class of query. Since, we are considering the raw signal, the average percentage of correct matches among k-NN is about 80%.

The degree of membership with clusters generated by fuzzy c-means clustering is used to extract the feature vectors for the given motion described by motion capture and EMG data. Our approach has shown the satisfactory results in terms of feature classification. Our work represents the new way of integrating the two different bio-medical data which when analyzed together and correctly gives more precise and accurate information.
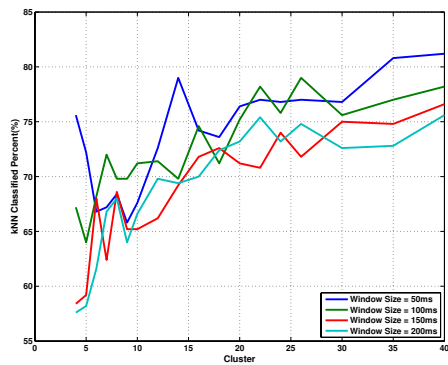
Figure 9. Percent of correctly classified Right Leg motions out of k(= 5) motions retrieved for the right leg query motion.

References

[1] A. D. Boca and D. Park. Myoelectric signal recognition using fuzzy clustering andartificial neural networks in real time. In *IEEE International Conference on Neural Networks and IEEE World Congress on Computational Intelligence*, volume 5, pages 3098–3103, Orlando, FL, USA, 1994.

[2] F. H. Chan, Y.-S. Yang, F.K.Lam, Y.-T. Zhang, and P. Parker. Fuzzy emg classification for prosthesis control. *IEEE Transactions onRehabilitation Engineering*, 8(3):305–311, September 2000.

[3] C.-Y. Chiu, S.-P. Chao, M.-Y. Wu, S.-N. Yang, and H.-C. Lin. Content-based retrieval for human motion data. *Journal of Visual Communication and Image Representation*, pages 446–466, April 2004.

[4] G. H. Golub and C. F. V. Loan. *Matrix Computations*. The Johns Hopkins University Press, Baltimore, Maryland, 1996.

[5] D. Graupe, J. Salahi, and K. Kohn. Multifunction prosthesis and orthosis control via microcomputer identification of temporal patter differences in single-site myoelectric signals. *Jornal of Biomedical Engineering*, 4:17–22, 1982.

[6] V. Gupta, S. Suryanarayanan, and N. Reddy. Fractal analysis of surface emg signals from the biceps. *International Journal of Medical Informatics*, 45(3):185–192, July 1997.

[7] B. Hudgins, P. Parker, and R. Scott. A new strategy for multifunction myoelectric control. *IEEE Transactions on Biomedical Engineering*, 4(1):82–94, 1993.

[8] E. Keogh, T. Palpanas, V. B. Zordan, D. Gunopulos, and M. Cardle. Indexing large human-motion databases. In *Proc. 30th VLDB Conference*, pages 780–791, Toronto, Canada, 2004.

[9] F. Liu, Y. Zhuang, F. Wu, and Y. Pan. 3D motion retrieval with motion index tree. *Computer Vision and Image Understanding*, 92:265–284, June 2003.

[10] M. Muller, T. Roder, and M. Clausen. Efficient content-based retrieval of motion capture data. *ACM Transactions on Graphics (TOG)*, 24:677–685, 2005.

[11] S. Nascimento. *Fuzzy Clustering via Proportional Membership Model*. IOS Press, Amsterdam, Netherlands, 2005.

[12] M. Raez, M. Hussain, and F. Mohd-Yasin. Techniques of emg signal analysis: detection, processing, classification and applications. *Biol Proced Online*, 8:11–35, 2006.

[13] K. Yang and C. Shahabi. Multilevel distance-based index structure for multivariate time series. In *TIME*, Burlington, Vermont, USA, 2005. IEEE Computer Society.

[14] C. Yu, B. C. Ooi, K.-L. Tan, and H. V. Jagadish. Indexing the distance: An efficient method to knn processing. In *Proc. VLDB '01*, pages 421–430, San Francisco, CA, USA, 2001. Morgan Kaufmann Publishers Inc.

[15] M. Zardoshti-Kermani, B. Wheeler, K. Badie, and R. Hashemi. Emg feature evaluation for movement control of upper extremity prostheses. *IEEE Transactions on Rehabilitation Engineering*, 3(4):324–333, December 1995.