A k-Nearest Neighbor Based Algorithm for Human Arm Movements Recognition Using EMG Signals

Mohammed Z. Al-Faiz ,MIEEE
Computer Engineering Dept.
Nahrain University
Baghdad, Iraq
Email: mzalfaiz@ieee.org

Abduladhem A.Ali
Computer Engineering Dept
Basrah University
Basrah, Iraq
Email: Abduladem1@yahoo.com

Abbas H.Miry
Electrical Engineering Dept.
AL-Mustansiriyah University
Baghdad, Iraq
Email: abbasmiry83@yahoo.com

***Abstract*—In a human–robot interface, the prediction of motion, which is based on context information of a task, has the potential to improve the robustness and reliability of motion classification to control human-assisting manipulators. The electromyography (EMG) signals can be used as a control source of artificial arm after it has been processed. The objective of this work is to achieve better classification with multiple parameters using K-Nearest Neighbor for different movements of a prosthetic arm. A K- Nearest Neighbor (K-NN) rule is one of the simplest and the most important methods in pattern recognition. The proposed structure is simulated using MATLAB Ver. R2009a, and satisfied results are obtained by comparing with conventional method of recognition using Artificial Neural Network(ANN), that explains the ability of the proposed structure to recognize the movements of human arm based EMG signals. Results show the proposed technique achieved a uniformly good performance with respect to ANN in term of time which is important in recognition systems, better accuracy in recognition when applied to lower SNR signal .**

I. INTRODUCTION

Robot arms are versatile tools found in a wide range of applications. In recent years, applications where humans and robot arms interact, have received increased attention. The case where the interaction entails the human controlling a robot is called robot teleportation. The latter case requires a user interface to translate the operator commands to the robot. A large number of interfaces have been proposed on this issue in previous works, while the user moves his/her the arm, electromyographic (EMG) activity is recorded from selected muscles, using surface EMG electrodes. Through a decoding procedure the muscular activity is transformed to kinematic variables that are used to control the robot arm EMG signals have been used as control signals for robotics devices in the past. Fukuda [1] proposed a human-assisting manipulator teleported by EMG signals and arm motions. In this case wrist motion of the robot arm was controlled using the muscular activity from the muscles of the forearm [2]. (EMG) signals, which are measured at the skin surface, are the electrical manifestations of the activity of muscles. It provides an important access to the human neuromuscular system. EMG has been well recognized as an effective tool to generate control commands for prosthetic devices and human-assisting manipulators. Up to the present, a number of EMG-based human interfaces have been proposed as a means for elderly people and the disabled to control powered prosthetic limbs, wheelchairs, teleported robots, and so on The core part of these human–robot interfaces is a pattern classification process, where motions or intentions of motions are classified according to features extracted from EMG signals. Commands for device control are then generated from the classified motions [3].

Information extracted from EMG signals, represented in a feature vector, is chosen to minimize the control error. In order to achieve this, a feature set must be chosen which maximally separates the desired output classes. The extraction of accurate features from the EMG signals is the main kernel of classification systems and is essential to the motion command identification [4].

EMG classification is one of the most difficult pattern recognition problems because there exist large variations in EMG features. Especially, it is difficult to extract useful features from the residual muscle of an amputee .So far, many researches proposed many kinds of EMG feature to classify posture and they showed good performance However, how to select a feature subset with the best discrimination ability from those feature is still an issue for classifying EMG signals [5].

Several approaches to solve the motion command identification problem using EMG signals have been suggested, e.g., amplitude of EMG signals by [6] as feature extraction with Artificial Neural Network. Although conventional method using ANN has resulted in some theoretical achievements for prosthetic arms, further advancement such as accurate identification of motion is required to achieve an ultimate goal. This paper presents an EMG pattern recognition method for more accurate identification of a motion command. The proposed method based on K-NN, it requires little computing time in the pattern recognition with accurate identification.

II EMG Signal Fundamentals

The electromyogram (EMG) is the recording of the electrical activity produced within the muscle fibers. The relation of surface EMG to torque makes EMG an attractive alternative to direct muscle tension measurements, necessary in many physical assessments. However, the complexity of the EMG signal origin has been a barrier for developing a quantitative description of this relation. The EMG signal origin and character is necessary background to understand the difficulty of establishing a relationship between surface EMG and torque [7].

The nervous system controls the voluntary movement of various body parts in humans by contracting and relaxing various skeletal muscles. To instantiate a contraction, a neuron generates a small electrical potential on the surface of the muscle fiber. This electrical potential causes depolarization of the muscle fiber tissue and a following depolarization waveform. This waveform travels the length of the muscle fiber and is known as the action potential (AP). Fig. 1 depicts the generation of electric fields in muscle fibers.

Muscle fibers are excited by nerve branches from one motoneuron in groups known as motor units. These motor units are defined as the fundamental unit of contraction and can range from a few muscle fibers for small muscles such as those in the hand and fingers, to thousands of muscle fibers in large muscles. Because each motor unit contains a number of muscle fibers that are attached to the motor neuron at various points, the electrical signal of a motor unit is the summation of the action potential of each muscle fiber, which may be

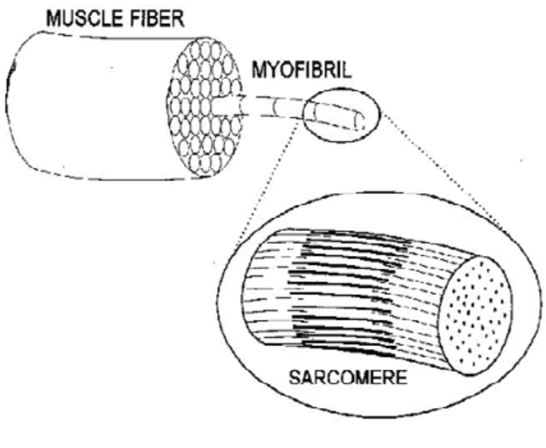

Fig. 1 Muscle Fibers Composition

Phase shifted from the other muscle fibers in that unit. This notion is reinforced in Fig. 2. The electrical potential due to contraction of all fibers in a motor unit during a single activation is referred to as the motor unit action potential (MUAP). This MUAP can be recorded by using electrodes placed on the surface of the skin above the muscle .Also, a muscle is not typically excited by only one action potential. In order to hold a contraction for any length of time, the motor units must be repeatedly activated. This repeated activation gives rise to a series of MUAPs that can be modeled as a pulse train in classical signal processing terms. This series of MUAPs that is produced is referred to as a motor unit action potential train (MUAPT). When measured using a surface electrode, the electromyogram can be defined as the superposition of numerous MUAPTs firing asynchronously. Fig 3 reinforces the notion that the superposition of motor unit action potentials gives rise to surface EMG The surface electromyogram signal typically does not exceed 5-10 mV in amplitude with the majority of signal information being contained between the frequencies of 15 and 400 Hz. The signal is an interference pattern which grows with muscle effort. As a result, the amplitude of the EMG contains a great deal of the signal information which can be modeled as a Gaussian random process. The EMG amplitude can thus be defined as the time-varying standard deviation of the EMG signal and is a measure of the activity level of the muscle under observation. [7].

.

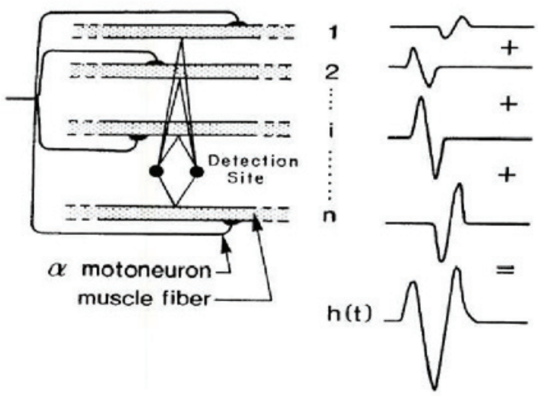

Fig. 2 Motor Unit Action Potential

III. MUSCLE ANATOMY

Agonist-antagonist muscles exist in many human joint. Such human joint is usually activated by many muscles .The following is a summary of the muscles that are responsible for the movement of the arm, wrist, and hand. Abduction of the arm is performed by the deltoid. Human elbow is mainly actuated by two antagonist muscles: biceps and triceps, although it consists of more muscles. Consequently, biceps and a part triceps are bi-

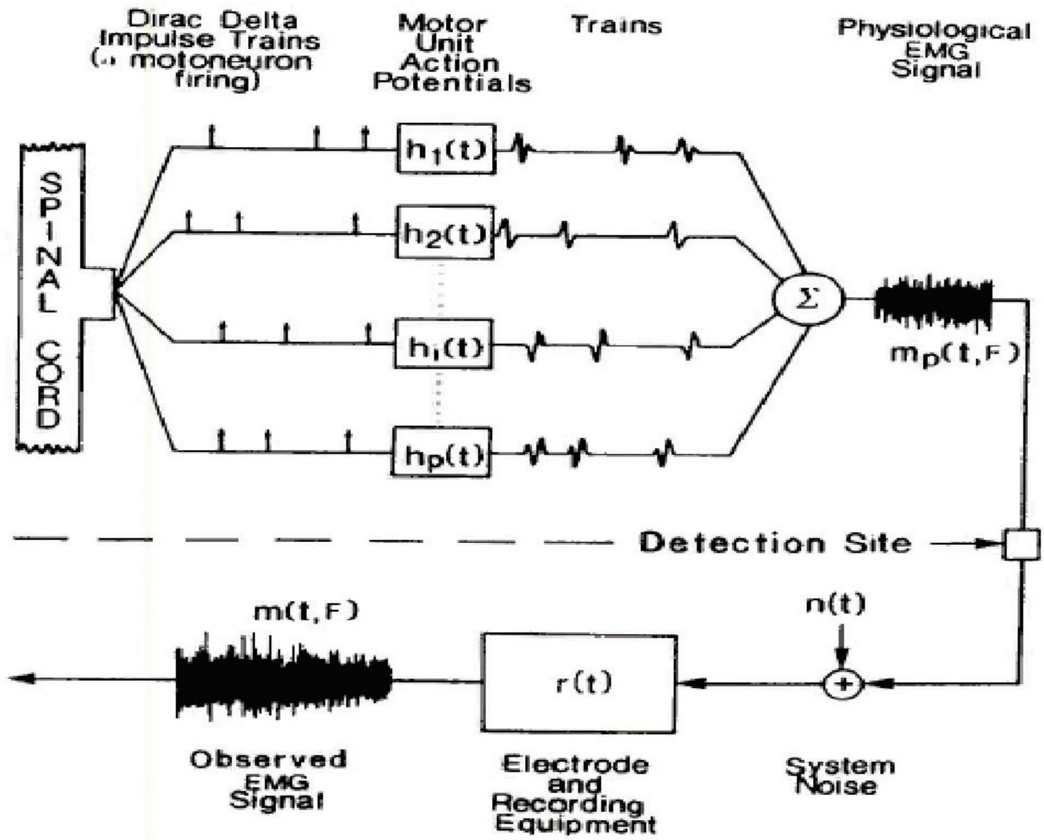

Fig. 3 Superposition of Motor Unit Action Potential Give Rise to Surface EMG

particular muscles. Many studies have been performed to investigate the effects of bi-articular muscles. By adjusting the amount of force generated by these muscles, the elbow angle and impedance can be arbitrary controlled. [8] The biceps brachii and triceps brachii however, which insert from the arm. Contraction of the biceps brachii flexes the elbow and supinates the forearm. Contraction of triceps brachii extends the elbow. Most of the muscles that move the forearm and hand originate within the forearm. The flexor carpi radialis flexes and abducts the wrist, while the flexor carpi ulnaris flexes and adducts. The extensor carpi radialis produces extension and abduction of the wrist, while the extensor carpi ulnaris produces extension and abduction [9].

IV. FEATURE PARAMETERS

The success of any pattern classification system depends almost entirely on the choice of features used to represent the raw signals. It is desirable to use multiple feature parameters for EMG pattern classification since it is very difficult to extract a feature parameter which reflects the unique feature of the measured signals to a motion command perfectly. But the inclusion of an additional feature parameter with a small separability may degrade overall pattern recognition performance

A. Integrated EMG

Integrated EMG (IEMG) is calculated as the summation of the absolute values of the EMG signal amplitude. Generally, IEMG is used as an onset index to detect the muscle activity that used to oncoming the control command of assistive control device. It is related to the EMG signal sequence firing point.

$$IEMG = \sum_{n=1}^{N} |x_n| \qquad (1)$$

Which can be expressed as where N denotes the length of the signal and x_n represents the EMG signal in a segment.

B.Mean Absolute Value

Mean Absolute Value (MAV) is similar to average rectified value (ARV). It can be calculated using the moving average of full-wave rectified EMG. In other words, it is calculated by taking the average of the absolute value of EMG signal. It is an easy way for detection of muscle contraction levels and it is a popular feature used in myoelectric control application. It is defined as

$$MAV = \frac{1}{N}\sum_{n=1}^{N}|x_n| \quad (2)$$

C. Modified Mean Absolute Value

Modified Mean Absolute Value (MMAV) is an extension of MAV using weighting window function w_x. It is shown as

$$MMAV = \frac{1}{N}\sum_{n=1}^{N} w_x|x_n| \quad (3)$$

$$w_x = \begin{cases} 1 & if\ 0.25N \le n \le 0.75N \\ 0.5 & otherwise \end{cases} \quad (4)$$

D. Variance of EMG

Variance of EMG (VAR) uses the power of the EMG signal as a feature. Generally, the variance is the mean value of the square of the deviation of that variable. However, mean of EMG signal is close to zero. In consequence, variance of EMG can be calculated by

$$VAR = \frac{1}{N-1}\sum_{n=1}^{N} x_n^2 \quad (5)$$

E. Waveform Length

Waveform length (WL) is the cumulative length of the waveform over the time segment. WL is related to the waveform amplitude, frequency and time. It is given by [10].

$$WL = \sum_{n=1}^{N-1}|x_{n+1} - x_n| \quad (6)$$

F. Wilson amplitude (WAMP).

This is the number of times that the difference between two consecutive amplitudes in a time segment becomes more than threshold.

$$WAMP = \sum_{i=1}^{N} f(|x_{n+1} - x_n|) \quad (7)$$

Where

$$f(x) = \begin{cases} 1 & if\ x > threshold \\ 0 & otherwise \end{cases} \quad (8)$$

It can be formulated as this feature is an indicator of firing motor unit action potentials (MUAP) and therefore an indicator of the muscle contraction level[11].

V. ARTIFICAL NEURAL NETWORK (ANN)

Artificial neural network is an information processing paradigm that is inspired by the way biological nervous systems, such as brain, process information. It is composed of a large number of highly interconnected processing elements (neurons). Each neuron unit i receives input from some other units, or from an external source. Each input x_{ij} , j = 1, 2, ... has an associated weight w_{ij} . The output y_i is computed by some function f of the weighted sum of its inputs, so called “activation function”. The mathematical representation of unit i is

$$y_i = f(net) = f(\textstyle\sum_j w_{ij} x_{ij}) \quad (9)$$

Generally, in ANN architectures, neurons are sorted in a number of layers and neuron outputs in each layer are interconnected to the other layer neuron inputs. The ANNs are distinguished by their learning and recall mechanisms, the activation functions, the number of layers and neurons, and the distribution of the connections [12]. The back propagation neural network (BP-NN) is used in this work the structure used for (BP-NN) is: six neuron in input layer, twenty neuron in hidden layer and seven neuron in output layer.

VI.. K-NEAREST NEIGHBOR (KNN) ALGORITHM

The K-nearest neighbor (KNN) classification rule is one of the most well-known and widely used nonparametric pattern classification methods. It was originally suggested by Cover and Hart [13], its simplicity and effectiveness have led it to be widely used in a large number of classification problems, [14]. When there is little or no prior knowledge about the distribution of the data, the KNN method should be one of the first choices for classification. It is a powerful non-parametric classification system which bypasses the problem of probability densities completely.NN is a “lazy” learning method because training data is not preprocessed in any way. The class assigned to a pattern is the class of the nearest pattern known to the system, measured in terms of a distance defined on the feature (attribute) space. On this space, each pattern defines a region (called its Voronoi region). When distance is the classical Euclidean distance, Voronoi regions are delimited by linear borders. To improve over 1-NN classification, more than one neighbor may be used to determine the class of a pattern (K-NN) or distances other than the Euclidean may be used [15]. The KNN rule classifies x by assigning it the label most frequently represented among the K nearest samples; this means that, a decision is made by examining the labels on the K-nearest neighbors and taking a vote. KNN classification was developed from the need to perform discriminant analysis when reliable parametric estimates of probability densities are unknown or difficult to determine.

Nowadays, it is the most usable classification algorithm . It has less usability and is labor intensive when the training dataset is large . This algorithm operation is based on comparing a given new record with training records and finding training records that are similar to it. It searches the space for the k training records that are nearest to the new record as the new record neighbors. In this algorithm nearest is defined in terms of a distance metric such as Euclidean distance . Euclidean distance between two records (or two points in n-dimensional space) is defined by forrmula1.

$x_1 = (x_{11}, x_{12}, \ldots., x_{1n}), x_2 = (x_{21}, x_{22}, \ldots., x_{2n})$

$$dist(x_1, x_2) = \sqrt{\sum_{i=1}^{n}(x_{1i} - x_{2i})^2} \qquad (10)$$

Where x_1 and x_2 are two records with n attributes. This Formula measures the distance between x_1 and the point x_2 , in terms of take the difference between the corresponding values of that attribute in record x_1 and in record x_2 . The K-nearest neighbor classifier is a supervised learning algorithm where the result of a new instance query is classified based on majority of the K-nearest neighbor category. The training samples are described by n-dimensional numeric attributes. Each sample represents a point in an n-dimensional pattern space. In this way, all of the training samples are stored in an n-dimensional pattern space. The following discussion introduces an example demonstrating the general concept of this algorithm in detail. The K nearest neighbor algorithm is very simple. It works based on minimum distance from the query instance to the training samples to determine the nearest neighbors. After K nearest neighbors is gathered, take simple majority of these K-nearest neighbors to be the prediction of the query instance. The data for KNN algorithm consists of several multivariate attributes names X_i that will be used to classify the object Y. Suppose that the K factor is set to be equal to 8 (there are 8 nearest neighbors) as a parameter of this algorithm. Then the distance between the query instance and all the training samples is computed. Because there are only quantitative Xi, the next step is to find the K-nearest neighbors. All training samples are included as nearest neighbors if the distance of this training sample to the query is less than or equal to the K^{th} smallest distance. In other words, the distances are sorted of all training samples to the query and determine the K^{th} minimum distance. The unknown sample is assigned the most common class among its k nearest neighbors. As illustrated above, it is necessary to find the distances between the query and all training samples. These K training samples are the closest k nearest neighbors for the unknown sample. Closeness is defined in terms of Euclidean distance.

Let us consider a set of patterns $X = \{x_1, \ldots, x_N\} \subseteq R^P$ of known classification where each pattern belongs to one of the classes $W = \{w_1, w_2, \ldots, w_s\}$. The nearest neighbor (NN) classification rule assigns a pattern z of unknown classification to the class of its nearest neighbor, where $x_i \in X$ is the nearest neighbor to z if

$$D(x_i, z) = min\{D(x_l, z) \quad l = 1,2, \ldots, N\} \qquad (11)$$

D is the Euclidean distance between two patterns in R^p. This scheme is called the 1-NN rule since it classifies a pattern based on only one neighbor of z. The k-NN rule considers the k-nearest neighbors of z and uses the majority rule. Let $t_l, l = 1,2, \ldots, s$ be the number of neighbors from class l in the k-nearest neighbors of z [16],

$$\sum_{l=1}^{s} t_l = k \qquad (12)$$

Then z is assigned to class j if

$$t_j = \underbrace{Max\ \{t_l\}}_{l} \qquad (13)$$

Here is step by step on how to compute K-nearest neighbors KNN algorithm

1-Determine parameter K = number of nearest neighbors.
2-Calculate the distance between the query-instance and all the training samples.
3-Sort the distance and determine nearest neighbors based on the K-th minimum distance.
4-Gather the category Y of the nearest neighbors.
5-Use simple majority of the category of nearest neighbors as the prediction value of the query instance

VII. CASE STUDY

Generally, the methods of Recognition are divided into two states:

Training state: In this state data base is constrained with seven muscles(biceps,triceps, deltoid, flexor carpi radialis, extensor carpi radialis, flexor carpi ulnaris and extensor carpi ulnari). This state has following steps:

Step One: Take six frames from each muscle as shown in Fig.4 to produce forty two frames.

Step Two: Six features are extracted from each frame which introduced in section III .From this step forty two vectors are build ,each vector has six elements.

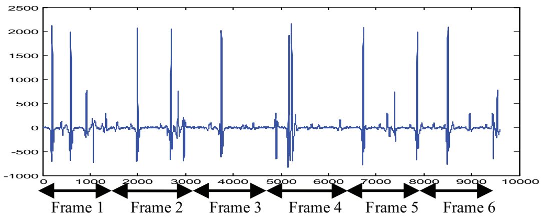

Fig. 4. EMG Signal

$X_i = [\,IEMG_i \;\; MAV_i \;\; MMAV_i \;\; VAR_i \;\; WL_i \;\; WAMP_i\,]$, $i = 1,..,42$ (14)

Classification state: this state has following steps as shown in Fig.5 and Fig. 6

Step one: In this step take the feature extraction(that describe in section III) of input signal to produce vector of six elements Z.

Step Two: Taking Euclidean distance between input vector Z and the data base X.

$$D(x_i, z) = \{D(x_i, z) \quad i = 1,2,\ldots,42\} \tag{15}$$

$$D(x_i, z) = \sqrt[2]{\begin{matrix}(x_{i1} - z_1)^2 + (x_{i2} - z_2)^2 + (x_{i3} - z_3)^2 + \\ (x_{i4} - z_4)^2 + (x_{i5} - z_5)^2 + (x_{i6} - z_6)^2\end{matrix}} \tag{16}$$

Step Three: Sort the distance and take first k[th] element.

Step Four: Calculate which class has more elements in this group which represents this class.

VIII. MOVEMENT RECOGNATION EXPERIMENT

Movement Recognition experiments were carried out in order to evaluate classification performance of the proposed method. To collect data the simulated data were generated from an EMG signal simulator based on a Physiologically and morphologically accurate muscle model. The simulator enables us to generate EMG signals of different complexities with knowledge of the signal intensity represented by the average number of MUP patterns per second (pps), the numbers of MUPT classes, and which MU created each MUP pattern. Several motions are recognized based on classified of seven inputs EMG signals .The performance index for the recognition was given by :

$$R = \frac{C}{T} X100\% \tag{17}$$

Where C the number of times EMG signals was correctly classified is, T is the total number of EMG signals which input to the system, in the present study . The accuracy for each participant was simply calculated by averaging the performance indices over all movements, to simulate real noise, different noises are considered, where a random noise is. The amplitude of noise is one tenth of the amplitude of the peak-to-peak range of the EMG signals which produce new EMG signal with lower SNR . Calculation complexity is an important factor in online applications, particularly in artificial arm control. The results illustrate that the classification using K-NN present the best results and fast in term CPU time from the classification using neural network as shown Table.I,Table.II shows the result of Authentication with noisy signal has lower SNR for the neural network and different value of k .Different value of k mean different number of group that taken. It can be found that the value of *k* doesn't significantly affect on the time of recognition while it achieves best performance on the other method with *k* = 15. The possible reason for the poor results of ANN may be due to the simple decision function realized by this method. EMG signal has variation with the time therefore necessary to check the input signal with more frame for each muscle as shown in K-NN which provides more accuracy in the recognition.

TABLE I CPU Times of Neural Network and K-NN Method

Method	CPU Time(sec)
ANN	1.2824
K-NN,K=7	0.0741
K-NN,K=9	0.0749
K-NN,K=13	0.0752
K-NN,K=15	0.0771
K-NN,K=17	0.0771

IX.CONCLUSION

In this paper, a new is considered about classification method which uses the multiple feature parameters with KNN algorithm in order to increase the classification accuracy. From the results reported in the last section, it can conclude that, the proposed techniques achieved a uniformly good performance in term of time which is important in recognition systems, better accuracy in classification when applied to lower SNR signal and the value of k can enormously affect the accuracy of classification . EMG signals recorded from muscles of the user's and activations from these muscles were used in order to control a human arm. Human arm movements recognition system based on EMG signals using K-NN is very good and easy to implement method from point of its simplicity , accuracy and it can be implemented in many medical applications such as using of artificial limps and high accuracy robot applications (interactive between human and robot).

TABLE II Recognition Accuracy of Neural Network and K-NN Method with Noisy Signal

SNR (dB)	ANN	K-NN K=7	K-NN K=9	K-NN K=13	K-NN K=15	K-NN K=17
25	83%	100%	100%	100%	100%	83%
11	66%	100%	100%	83%	100%	83%
9	66%	100%	100%	83%	100%	66%
7	66%	66%	83%	83%	83%	66%
5	50%	50%	66%	83 %	83%	66 %

REFERENCE

[1] O. Fukuda, T. Tsuji, M. Kaneko, and A. Otsuka, "A human-assisting manipulator teleported by emg signals and arm motions," IEEE Trans. on Robotics and Automation, Vol. 19, No. 2, pp. 210–222, 2000.

[2] P. K. Artemiadis and K. J. Kyriakopoulos." EMG-based Teleoperation of a Robot Arm Using Low-Dimensional Representation", Proceedings of IEEE/RSJ International Conference on Intelligent Robots and Systems, San Diego, Oct. 2007 , pp. 489 – 495.

[3] N.Bu, M. Okamoto, and T, Tsuji, "A Hybrid Motion Classification Approach for EMG-Based Human–Robot Interfaces Using Bayesian and Neural Networks "IEEE Trans. on Robot. Vol. 23, No. 3, pp. 502–511, June 2009.

[4] S.Park and S.Lee," EMG Pattern Recognition Based on Artificial Intelligence Techniques", IEEE Trans on Rehabilitation Engineering , Vol. 6, No. 4, Dec 1998 ,pp:400-405.

[5] H.Huang,Y.Liu and C.Wong,"Automatic EMG Feature Evaluation for Controlling a Prosthetic Hand Using a Supervised Feature Mining Method :An Intelligent Approach ",Proceeding of the IEEE,International Conference on Robotics & Automation ,Taiwan ,September 14-19,2003 pp:220-225

[6] M. Z. Al-Faiz, Y. I. Al-Mashhadany, "Human Arm Movements Recognition Based on EMG Signal "MASAUM Journal Of Basic and Applied Sciences (MJBAS) Volume 1 Issue 2, PP 164-171, (September 2009)

[7] O. Bida, " Influence of Electromyogram (EMG) Amplitude Processing in EMG-Torque Estimation ", M.Sc Thesis , WORCESTER POLYTECHNIC INSTITUTE ,Electrical Engineering , January 2005

[8] K. Kiguchi, . S. Watanabe, K. Izumi and T.Fukuda ," An Exoskeletal Robot for Human Elbow Motion Support—Sensor Fusion, Adaptation, and Control ", IEEE Trans on System , Man and Cybernetics , Vol. 31, No. 3, June 2001 pp:353-361.

[9] R. Boostani and M. Moradi," Evaluation of the forearm EMG signal features for the control of a prosthetic hand", Institute of Physics Publishing , Physiological Measurement Physiol. Meas. 24 ,2003 pp: 309–319 .

[10] A. Phinyomark, C. Limsakul, and P. Phukpattaranont," A Novel Feature Extraction for Robust EMG Pattern Recognition", Journal of Computing , ISSN: 2151-9617,Vol.1, No. 1, Dec 2009, pp:71-80.

[11] R. Elliott ,"Feature Extraction Techniques for Grasp Classification ",Master Thesis ,University of Canterbury,August 1998.

[12] L. Wang and T. S. Buchanan, "Prediction of joint moments using a neural network model of muscle activations from emg signals," IEEE Transactions on Neural System and Rehabilitation Engineering, vol. 10, pp. 30–37, 2002.

[13] T. M. Cover and P. E. Hart, "Nearest neighbor pattern classification,"IEEE Trans. Inform. Theory, vol. IT-13, pp. 21–26, 1967.

[14] H. Fayed and A. Atiya, "A Novel Template Reduction Approach for the -Nearest Neighbor Method",IEEE Trans on Neural Network , Vol. 20, No. 5, May 2009.pp:890-896.

[15] A. Cervantes, I. Galván, and P. Isasi", AMPSO: A New Particle Swarm Method for Nearest Neighborhood Classification", IEEE Trans on System , Man and Cybernetics , Vol. 39, No. 5, Oct 2009 , pp:1082-1091.

[16] N. Pal and S. Ghosh ,"Some Classification Algorithms Integrating Dempster–Shafer Theory of Evidence with the Rank Nearest Neighbor Rules",IEEE Trans on Systems , Man,and Cybernetics—Part A, Vol. 31, No. 1, pp: 59-66 Jan 2001.

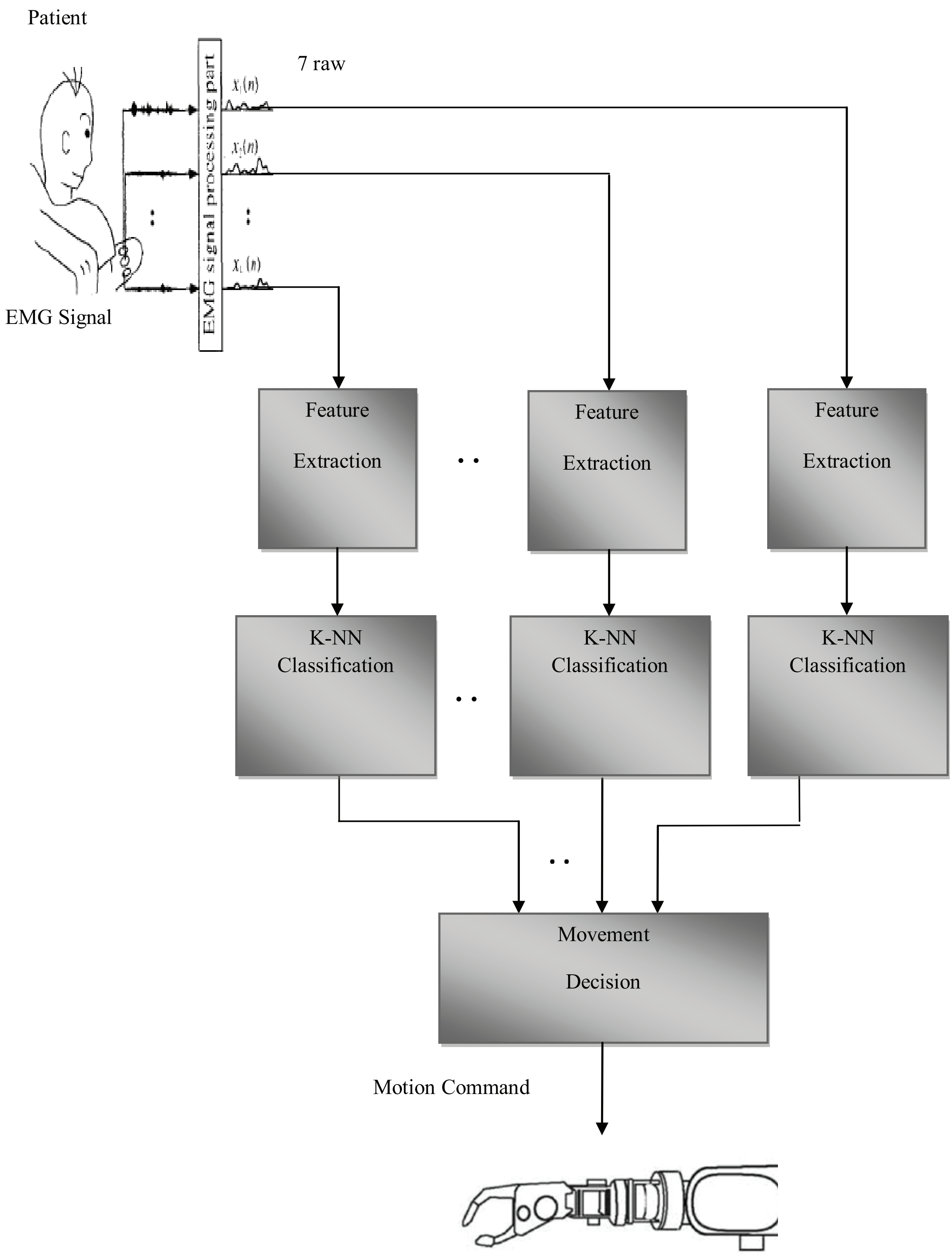

Fig. 5 Structure of the prototype system based on K-NN classifier

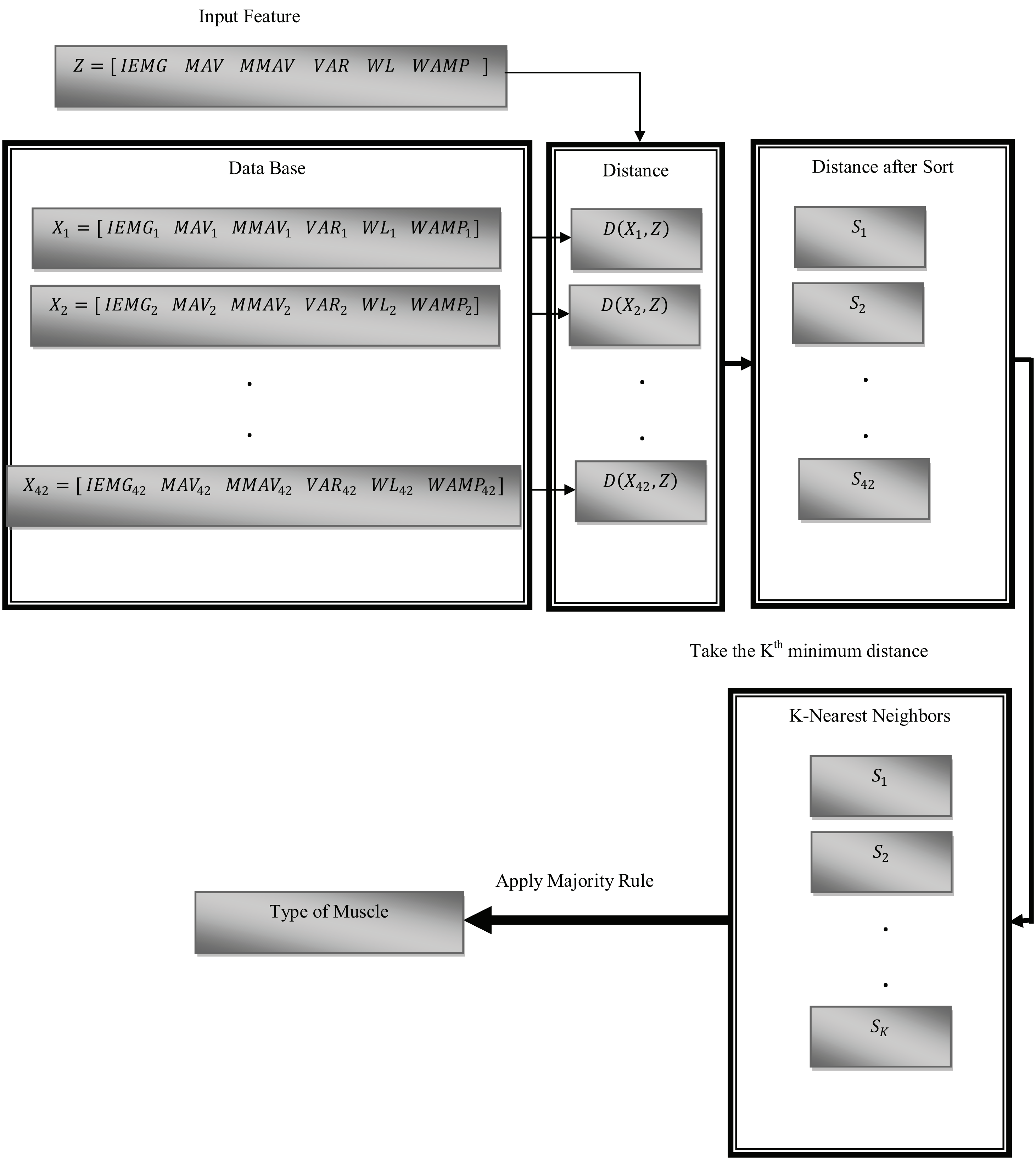

Fig. 6 . K-NN Classification Process