

Chapter 2

Methodology for Working with EMG Dataset

2.1 EMG Dataset

In our experiments, data acquisition system include PowerLab, 16sp, and Dual BioAmp manufactured by ADInstruments Ltd. and software Chart V5.0 with sampling rate adjusted at 2 kHz, recording signal amplitude 2 mV, primary low-pass filter 1 with cutoff frequency 500 Hz, and primary high-pass filter 2 with cutoff frequency 0.3 Hz. Data are outputted in txt or excel format which are readable in MATLAB for data processing. MATLAB 7.0 software installed on a Laptop with 2.2 GHz Core2Dual CPU is used for signal processing (Fig. 2.1).

Our samples are 20 normal and healthy subjects aged between 20 and 30 with almost similar physical power randomly selected from students of Biomedical Engineering Department, Amirkabir University of Technology, satisfying conditions of having enough sleep and appropriate nutrition, having no considerable physical activity before test, no sedative drug use for at least 24 h before test, with no bone fracture and musculoskeletal disorder close to test, and no pain sensed during tests by subjects. Each individual fills out a form requesting following items:

- Personal information as name, gender, age, height, and weight
- Types of recording signals
- Recording degrees of freedom
- Stimulation and motions
- Processing items requested
- Notes

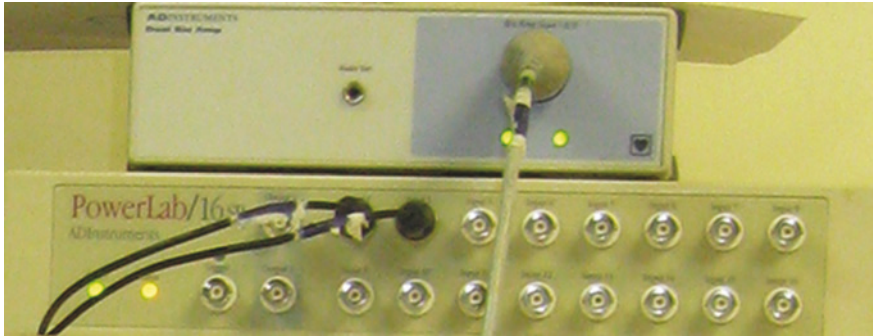


Fig. 2.1 Signal conversion devices PowerLab



Fig. 2.2 Biceps anatomy and electrode placement



Fig. 2.3 Deltoid anatomy and electrode placement

EMG signals of biceps, deltoid, triceps, tibialis anterior, and quadriceps muscles are recorded in three states of isometric contraction (ISO), maximum voluntary contraction (MVC), and dynamic contractions (Figs. 2.2, 2.3, 2.4, 2.5, and 2.6).

A preprocessing filtering process is then applied to recorded signals. A window consisting of 20,000 samples (10 s) is made cutoff for each signal to be processed and analyzed (Fig. 2.7).

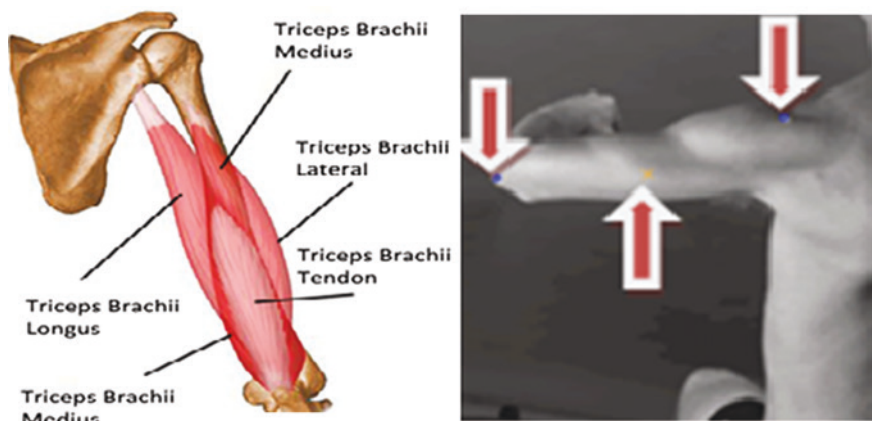


Fig. 2.4 Triceps anatomy and electrode placement

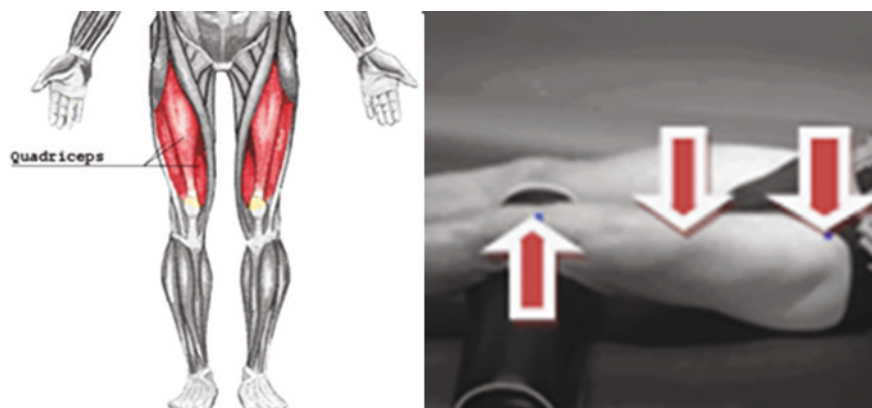
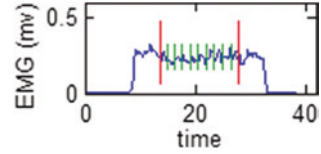


Fig. 2.5 Quadriceps anatomy and electrode placement



Fig. 2.6 Tibialis anterior anatomy and electrode placement

Fig. 2.7 Illustrated samples in 10 s



2.2 Feature Extraction

Feature extraction, which is step to measure features or properties from input data, is essential in pattern recognition system design. Goal of feature extraction is to characterize an object to be recognized by measurements whose values are very similar for objects in same category, and very different for objects in different categories. Computational complexity and class discrimination are two main factors for determining best feature set.

A set of features are listed in Table 2.1 along with their descriptions. The primary purpose of this work is to use these features to find an optimum set best describing and characterizing EMG signals. The nonlinear classifier used in this work is a five-layer neuro-fuzzy network; its inputs are selected features among feature list given in Table 2.2. In our experiments, it has been found that there

Table 2.1 Muscle test assessments by professional experts

Case number	Biceps	Triceps	Middle deltoid	Quadriceps	Tibialis anterior
1	5	+5	5	–	5
2	+5	+5	+5	–	+5
3	5	–5	+4	–	–5
4	–5	5	+4	–	–5
5	+4	4	4	–5	–5
6	4	4	+3	+3	–4
7	5	+4	+4	5	–5
8	5	5	5	–	5
9	+4	5	4	+4	5
10	+4	+4	+4	+4	5
11	5	5	+4	5	5
12	+5	+5	+5	–	–
13	+4	+4	+4	5	+4
14	4	+4	+4	5	+4
15	+4	4	4	5	+5
16	–	–	–	–	+4
17	+5	+5	+5	–	+5
18	+4	4	4	+4	+4
19	+5	+5	+5	–	+5
20	+4	+4	4	+4	5
21	+4	5	+4	–	+5

*Because of laboratory condition, quadriceps just had been evaluated in female

Table 2.2 Extracted features of recording EMG signals (k: sample number and N: total number of samples).

Item	Feature name	Feature definition	Description
1	Integration of absolute of EMG signal	$\text{IEMG} = \sum_{k=1}^N \text{emg}_k $	Related to muscle activity
2	Mean absolute value of signal	$\text{MAV} = \frac{1}{N} \sum_{k=1}^N \text{emg}_k $	Related to muscle contraction points
3	Root mean squared of signal	$\text{RMS} = \left(\frac{\sum_{k=1}^N \text{emg}_k^2}{N} \right)^{1/2}$	Related to muscle contraction indication with constant force before starting the muscle fatigue
4	Wave length of signal	$\text{WL} = \sum_{k=1}^{N-1} \text{emg}_{k+1} - \text{emg}_k $	–
5	Difference absolute mean value of signal	$\text{DAMV} = \frac{1}{N-1} \sum_{k=1}^{N-1} \text{emg}_{k+1} - \text{emg}_k $	Mean of WL
6	Variance of signal	$\text{VAR} = \frac{1}{N-1} \sum_{k=1}^N \text{emg}_k^2$	Related to signal power
7	Zero crossings of signal	$\text{ZC} = \sum_{k=1}^N \text{sgn}(-\text{emg}_k \text{emg}_{k+1})$ $\text{sgn}(x) = \begin{cases} 1 & \text{if } x > 0 \\ 0 & \text{otherwise} \end{cases}$	For measuring frequency shift and showing number of signal sign varying
8	Wilson amplitude of signal	$\text{WAMP} = \sum_{k=1}^N f(\text{emg}_k - \text{emg}_{k+1})$ $f(x) = \begin{cases} 1 & \text{if } x > \text{threshold} \\ 0 & \text{otherwise} \end{cases}$	An indication of action potential firing of motional unit and therefore muscle contraction level

(continued)

Table 2.2 (continued)

Item	Feature name	Feature definition	Description
9	Cepstrum coefficients of signal	$c_1 = a_1$ $c_n = -\sum_{k=1}^N \left(1 - \frac{k}{N}\right) a_k c_{N-k} - a_N$ $CC = [c_1 \ c_2 \ \dots \ c_N]$	Containing information about power spectrum of signal and belonging to static features
10	Simple square integral of signal	$SSI = \frac{1}{N} \sum_{k=1}^N emg_k ^2$	An indication of energy of signal
11	Conduction velocity of signal	$CV = \left(\frac{1}{N-1} \sum_{k=1}^N emg_k^2 \right)$	Similar to RMS feature
12	Mean absolute value slope of signal	$MAVS = MAV(i+1) - MAV(i)$	Related to muscle contraction variations

is a trade-off between classification accuracy and computational complexity. Therefore, for off-line signal processing high classification accuracy is followed through assigning more effective features and for online processing in which computation time is concerned minimum possible number of features should be chosen.

2.3 Neuro-Fuzzy Classifier

These days, neuro-fuzzy systems have been used in broad span of commercial and industrial applications that require analysis of indefinite and indecisive information [45, 47–49]. Hybrid integrated neuro-fuzzy is the major interest of research as it makes use of complementarities' strength of artificial neural network and fuzzy inference systems [45]. ANFIS, a neuro-fuzzy model, is used in this study, which is hybrid technology of integrated neuro-fuzzy model and a part of MATLAB's Fuzzy Logic Toolbox [46]. ANFIS is called a hybrid learning method since it combines gradient descent and least squares method. Gradient descent method is used for premising and tuning parameters that define membership functions. Least squares method is used for identifying parameters that define coefficients of each output equation [46]. The normal structure of ANFIS-like structure used in MATLAB's Fuzzy Logic Toolbox is employed in this research for its efficiency and applicability in clustering and classification problems. The ANFIS has a five-layer structure as described later in this section. Input layer accepts features so that the number of input nodes is same as the number of features. As it was mentioned earlier, ANFIS is composed of gradient descent and least squares methods for learning purpose (Fig. 2.8).

To represent fuzzy inference system, fixed number of layers is presented structurally. ANFIS in comparison with other neuro-fuzzy networks has high training speed, most effective learning algorithm, and simplicity of software [50]. ANFIS is the best function approximator and classifier among neuro-fuzzy models, and its fast convergence is comparable to other neuro-fuzzy models, although it was one of the first integrated hybrid neuro-fuzzy models [51]. Besides, ANFIS affords superior results when applied without any pre-training [52]. Most of the neuro-fuzzy inference systems are based on Takagi–Sugeno or Mamdani type. For model-based applications, Takagi–Sugeno fuzzy inference system is usually used [53, 54]. However, Mamdani fuzzy inference system is used for faster heuristics but with a low performance [55]. High accuracy and easy interpretation of Takagi–Sugeno system makes it a general tool for approximation. The generality of Takagi–Sugeno type is used for identification of complex systems [56]. These systems usually have expensive computations and require complicated learning approaches, but their performance is notable. Typical fuzzy rule for Takagi–Sugeno system is

If x_1 is MF_i^1 and x_2 is MF_i^2 , Then O is Γ_i ,

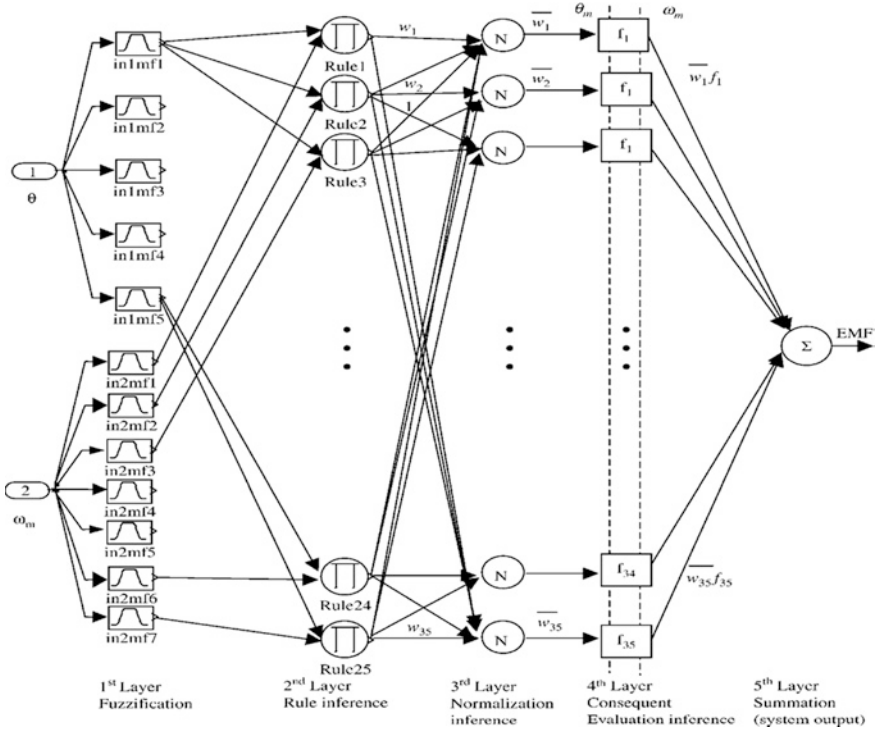


Fig. 2.8 The structure of neuro-fuzzy model

where MF_i^1 and MF_i^2 are fuzzy sets in antecedent and Γ_i is a crisp function in consequent. Usually, function O is a first order or a zero order for Takagi–Sugeno fuzzy inference [57]. In this study, first-order Takagi–Sugeno system is used for fuzzy inference part, and its structure is presented in Fig. 1.

Each one of the five layers of ANFIS performs a specific role for fuzzy inference system as follows:

Layer 1: First layer nodes are adaptive and generate membership grades for input set. Because of their smoothness and concise notation, Gaussian membership function, well known in fields of probability and statistics, is becoming increasingly popular function in fuzzy sets theory. In this study, Gaussian membership functions are used which can be automatically generated by ANFIS of MATLAB. The number of input nodes is the same as the number of features used for classification. Therefore, the number of features specifies the structure of neuro-fuzzy system and also complexity of learning procedure as it directly corresponds to the number of premise parameters as Gaussian membership functions.

Assume input is an N - D vector (same of extracted features) and we have 3 membership functions for each input array as

$$X = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_N \end{bmatrix}_{N \times 1}, \quad (2.1)$$

The Gaussian membership functions are given by

$$\text{GaussMF}_{D-\text{mf}}^{i,j} = \text{Exp} \left(-\frac{(x_i - C_j^i)^2}{2\delta_j^2} \right), \quad (2.2)$$

where x_i is an input array, and C_j^i and δ_j^i are antecedent parameters expressing centers and standard deviations of Gaussian membership functions of input vector, respectively. Output of this layer is a $M \times N$ matrix with M membership value for each of the N input variables.

With assumption mentioned above, for N input variables and 3 membership functions for each variable, output of first layer is

$$\text{MF}_{D=N, \text{mf}=3}^{\text{Gaussian}(X,C,\Sigma)} = \begin{bmatrix} \text{MF}_1^1 & \text{MF}_1^2 & \dots & \text{MF}_1^N \\ \text{MF}_2^1 & \text{MF}_2^2 & \dots & \text{MF}_2^N \\ \text{MF}_3^1 & \text{MF}_3^2 & \dots & \text{MF}_3^N \end{bmatrix}, \quad (2.3)$$

where MF_j^i is the membership value of the i th input to its j th membership function.

Layer 2: All nodes are fixed in this layer. All potential rules between the inputs are formulated applying fuzzy intersection (AND). The operation of product is used to estimate the firing strength of each rule. The output of this layer is a $M^N \times 1$ matrix (N input variables and M membership functions for each input variable), that is,

$$W_{3^N \times 1} = \left[w_{(i=1,2,\dots,3^N)}^i \right]_{3^N \times 1}, \quad (2.4)$$

Layer 3: In the third layer, the nodes are also fixed nodes. The nodes are symbolized by a notation of N , and the ratio of the i th rule's activation level to the total of all activation levels is computed. The output of this layer denominates as normalized firing strength

$$\bar{W}_{3^N \times 1} = \frac{W_{3^N \times 1}}{\sum_{i=1}^{3^N} w^i}, \quad (2.5)$$

Layer 4: The nodes are adaptive nodes in this layer. Each adaptive node i calculates the contribution of the i th rule toward the overall output as simply product of normalized firing strength and a first-order polynomial (for a first-order Sugeno model). Parameters in this layer are referred to as consequent parameters which shape output of this layer as

$$O^i = \Gamma^i \cdot \bar{w}^i \quad (2.6)$$

Layer 5: There is only one single fixed node in the fifth layer which calculates overall output as summation of contribution from each rule

$$O_{\text{mem.value}} = \Gamma_{1 \times 3^N} \bar{W}_{3^N \times 1}, \quad (2.7)$$

where

$$\Gamma_{1 \times 9} = [x_1 \ x_2 \ 1] \text{Coeff}_{3 \times 3^N}, \quad (2.8)$$

That $\text{Coeff}_{3 \times 3^N}$ is a matrix of consequent parameters of ANFIS used. It can be observed that there are two adaptive layers in this ANFIS structure, the first and fourth layers. There are two modifiable matrices of parameters $C_{3 \times N}$ and $\Sigma_{3 \times N}$ which shape input Gaussian membership functions. These parameters are so-called premise parameters. In the fourth layer, there is also a modifiable matrix of parameters $\text{Coeff}_{3 \times 3^N}$, pertaining to first-order polynomial. These parameters are so-called consequent parameters [58, 59].

Both premise and consequent matrices of parameters are adjusted during learning procedure aiming to make ANFIS output match training data. Least squares method can be used to identify optimal values of these parameters easily. When premise parameters are not fixed, search space becomes larger and convergence of training becomes slower. A hybrid algorithm combining least squares method and gradient descent method is adopted to solve this problem. Hybrid algorithm is composed of a forward pass and a backward pass. Least squares method (forward pass) is used to optimize consequent parameters with premise parameters fixed. Once optimal consequent parameters are found, backward pass starts immediately. Gradient descent method (backward pass) is used to adjust optimally premise parameters corresponding to fuzzy sets in input domain. Output of ANFIS is calculated by employing consequent parameters found in forward pass. Output error is used to adapt premise parameters by means of a standard back-propagation algorithm. It has been proven that this hybrid algorithm is highly efficient in training ANFIS [58, 59]. Once ANFIS is structured and learned, parameters are deterministic and classification can be executed.