AN ABSTRACT OF THE THESIS OF

David Lerman for the degree of Master of Science

in Electrical and Computer Engineering presented on December 5 1991

Title: A Model for the Generation and Study of Electromyographic Signals

Abstract approved: Redacted for Privacy

John L. Saugen

A computer model simulating the electrical activity of muscles of the upper arm during elbow motion is presented. The output of the model is an Electromyographic (EMG) signal. System identification is performed on the EMG signals using autoregressive moving average (ARMA) modelling. The calculated ARMA coefficients are then used as the feature set for pattern recognition. Pattern recognition is performed on the EMG signals to attempt to identify which of four possible motions is producing the signal. The results of pattern recognition are compared with results from pattern recognition of real EMG signals. The model is shown to be useful in predicting general trends found in the real data, but is not robust enough to predict accurate quantitative results. Simplifying assumptions about the filtering effects of body tissue, and about the size and position of muscles, are conjectured to be the most likely reasons the model is not quantitatively accurate.

A Model for the Generation and Study of Electromyographic Signals

by

David Lerman

A THESIS

submitted to

Oregon State University

in partial fulfillment of
the requirements for the
degree of

Master of Science

Completed December 5, 1991

Commencement June 1992

APPROVED:

Redacted for Privacy

Professor of Electrical and Computer Engineering in charge of major

Redacted for Privacy

Head of Department of Electrical and Computer Engineering

Redacted for Privacy

Dean of Graduate School

Date thesis is presented December 5, 1991

Typed by David Lerman

ACKNOWLEDGEMENTS

I would like to thank professor John Saugen for his long commitment to myself and my work on this thesis. This includes both supporting my work in the medical engineering field even though the ECE department does not formally support any biomedical curriculum, and also his enthusiasm to resume the work after my long hiatus from academic life.

I would also like to thank my wife, Jacquie, for all of the extra things she did so I could have enough time to do this work.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

A MODEL FOR THE GENERATION AND STUDY OF ELECTROMYOGRAPHIC SIGNALS

INTRODUCTION

Electromyography is the study of the signals generated by the electrical activity of muscle contraction. There are many reasons to study electromyographic (EMG) signals. These reasons range from the pure research concerns of determining when and how muscles act to perform motion, to clinical concerns of diagnosing muscle and nerve disease.

Since body movement via skeletal muscle action is fundamentally under the voluntary control of an individual, EMG signals are the end result of commands originating in the brain. If information about the original command can be determined from the EMG signals, these signals can be used as the inputs to controlling external devices such as prosthetic devices for amputees.

This thesis will introduce a computer model for the generation of EMG signals. System identification techniques are applied to the model as a test of how well it simulates actual physiological processes, and results from the model are compared to data acquired from a live subject.

Problem Statement:

EMG signal analysis is an important area of research which has received a great deal of attention from many researchers. Virtually all analysis is done on data acquired from live subjects.

There are several reasons why relying on real EMG data for analysis is not optimal. First of all, no two people are exactly the same. Even when the researcher takes great pains to gather data in exactly the same way from more than one person, differences in the individuals' anatomical and physiological makeup makes comparing results on the data inexact. If the researcher acquires data from a very large sample of people to statistically remove individual bias, the very real problems of data storage and management quickly arise.

Other drawbacks of acquiring real data is that it can be time consuming, expensive, and frequently uncomfortable or even painful for the subjects.

If a computer model existed that could simulate the generation of EMG signals, it would be helpful to researchers in the field. Different types of people or different states of disease or injury could be simulated with little effort by changing model parameters. The model would also prove useful in creating a testing database to evaluate proposed analysis algorithms a priori.

It is not proposed that a computer model can replace the need for testing real people. The final step to testing any methodology or product must still be done using real EMG data. A model can, however, greatly reduce the need for real EMG data in the early stages of research or product development, reducing negative aspects of data collection and maintenance mentioned earlier.

Review of Literature Search:

Over the past two decades, with the boom of computer technology and the continuing trend of more computing power at cheaper prices, mathematical processing of EMG signals has become an important field of research. The engineering science of signal analysis is often applied to EMG signals, and the engineering field of system identification is used toward the goals of prosthetic control and diagnosing nerve and muscle disease.

As early as 1974 Graupe [7] had applied autoregressive moving average (ARMA) modeling to EMG signals for the purpose of prosthetic control. ARMA modeling is a time domain based approach to time series analysis based primarily on the autocorrelations of the signal under test. In 1978 Graupe built a working system using his research and reported an 85% rate of success in determining motion information from EMG signals on one patient. Since this original work other researchers have repeatedly applied ARMA modeling to EMG signals [5,15,18]. Although widely used, ARMA modeling is sometimes found to give unacceptable results[18]. No concrete explanation can be found to explain the disparity of results using ARMA modeling, but the wide range of test subjects, operating conditions, desired outputs, and exact test methodologies are all probably complicit. A model such as the one presented in this thesis could be used to help understand this disparity.

Analysis of EMG signals has not been limited to ARMA modeling. Frequency domain analysis is often applied [4,10,11,18,19]. There have also been several ad hoc attempts using features such as zero crossings and various statistical moments of the signal as features for pattern discrimination [14].

Not all research in electromyography deals with the analysis of the EMG signal as measured by surface electrodes on the skin. Much of the research is done understanding the many subsystems that function during muscle contraction and motion. These subsystems are the nerve - muscle interface [18,19], the propagation of action potentials along muscle fibers and the mechanisms of ion flow [16], and the effects of measuring signals at the electrode - tissue interface for both surface and indwelling electrodes. Analysis of the filtering effects of the various tissues on the EMG signal as it passes through the body is an important link in the EMG signal generation process that is not investigated very thoroughly, probably because of the complicated nature of the process.

The common thread to all of the research is the use of real EMG signals for analysis. The only mention found of simulating the processes under test is by Zhang [19], who programmed a function generator to produce his mathematical model of the signal generated by a single motor neuron exciting a group of muscle fibers. This model did not expand to simulate the entire process of motion and the resultant EMG signals.

REVIEW OF PHYSIOLOGY

When a thought originates in the brain to produce a movement, it is processed by the central nervous system until the nerves that innervate the muscles responsible for the motion are excited. Motor neurons are the nerves that innervate skeletal muscle. Skeletal muscle is the type of muscle that is attached to the skeleton and is responsible for movement of the skeleton. The bodies of all of the motor neurons start in the spine and extend to the periphery by a part of the cell body called the axon. (It is a grouping of these individual nerve cell axons that is commonly referred to as a nerve). As a motor neuron's axon approaches the location of muscle it will innervate, it splits and forms junctions with several muscle fibers. While a motor neuron may innervate many muscle fibers, in the normal person any one muscle fiber is innervated by only one motor neuron. A single motor neuron and the group of muscle fibers it innervates is referred to as a motor unit (MU).

Each muscle fiber is a single cell. Muscles, as they are commonly regarded, are collections of between hundreds and many thousands of muscle fibers. The activity that occurs within any muscle fiber when it is excited by a motor neuron is almost identical to that of any other muscle fiber in the body. A motor neuron comes into very close proximity to a muscle fiber at a region of the fiber called the motor end plate. When a nerve excites the muscle fiber the membrane characteristics at the motor end plate change, producing a propagating wave of sodium ion flow, the muscle impulse, which travels in both directions away from the motor end plate. This electrical activity is transmitted to the interior of the muscle and causes the release of calcium from internal holding sacs. The released calcium triggers the process of muscle contraction. The calcium is eventually reabsorbed

into its holding sacs, at a much slower rate than it was released, enabling the muscle fiber to contract again. It is the ion flow that is associated with the muscle impulse that is the source of the EMG signals. Since a group of muscle fibers contract each time a motor neuron is excited, the fundamental unit of electrical activity is the sum of the outputs from all of the fibers of a motor unit; this is a motor unit action potential (MUAP).

The MUAP is a temporal spatial event that will "look" different depending on the sensing electrode and its relative position to the unit and, if small and close enough, to its relative position to the individual fibers within the unit. Each motor unit will fire at a frequency determined necessary to carry out the original motion command and controlled by sensory input from nerves internal to the body and frequently from other senses such as vision.

The MUAP is a high frequency event that is low pass filtered by body tissues as it propagates through the body. LeFEVER and DeLUCA [9] report in a study of MUAPs that they low pass filtered their data at 1000hz in order to remove the effects of nearby motor units whose action potentials have already been filtered to this lower frequency range by the surrounding tissues.

The EMG signal which is measured at the skin's surface by a surface electrode is a summation of hundreds to thousands of individual MUAPs, each one having been filtered by the tissues between the muscle fibers and the electrode. The resultant EMG signal is commonly low pass filtered at 1000hz to remove high frequency noise, since the tissue filtering does not allow frequencies higher than this.

Muscles are anatomically organized in the body to create motion. It is pretty obvious that a muscle, which is oriented to cause a certain motion, will be contracting while that motion occurs. A muscle with such an orientation can be

referred to as the prime mover, or the agonist, of the motion. What is not so obvious is that due to the complex mechanics of skeletal joints in motion, it is usually necessary for many other muscles to contract as well, at varying levels of force, in order to maintain smooth motion of the joint.

It is very common that the muscle which pulls directly opposite the prime mover will need to maintain a fair level of contraction for good joint mechanics. This muscle will be exerting a force, "contracting", even though it is lengthening at this time. The lengthening contraction of the muscle opposite the prime mover is often referred to as co-contraction. The muscle pulling in the opposite direction of the agonist is called the antagonist.

Muscles which contract during movement in order to stabilize a moving joint are called fixators.

THE COMPUTER MODEL

Introduction:

The computer model presented here encompasses the entire event of motion from the original thought to the EMG signal output that is one of its byproducts. The region specifically being modelled is the upper arm of a normal human being. The model region is shown in Figure 1.

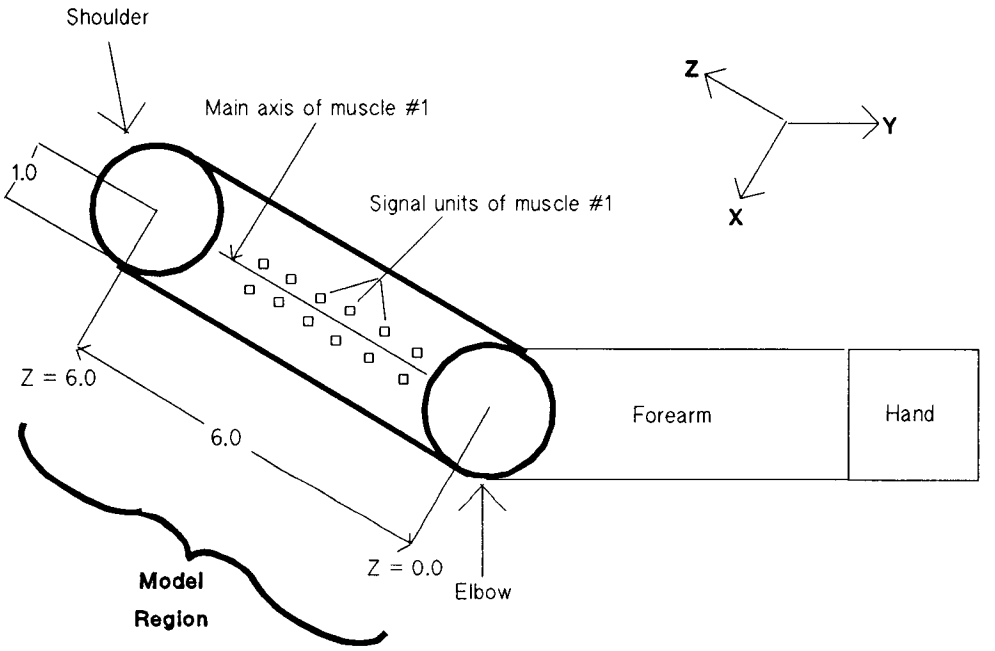


Figure 1: The Model Region

The inputs to a run of the model are: the motion to be performed (the thought), and the position of the sensing electrode. The parameters of the model

that are "fixed" are the output from motor units, the composition and location of muscles, the definitions of the type of contraction each muscle experiences during any specific motion, the filtering effects caused by tissue in different areas of the model region, and the parameters for determining what tissue filtering applies to the output of any motor unit based on electrode position. "Fixed" parameters can not be changed by the user when the model is run. All parameters can be changed and the system recompiled; this is currently the way to improve the model based on experimental results.

The model is written in C, using the Microsoft C6 compiler, and run on an IBM compatible PC.

The Model Structure:

The computer model is a discrete system simulating a sampled system with a sampling rate of 4000hz. 4000hz is chosen because it will accurately represent the highest frequencies in the simulation, and because it is the sampling rate of the instrument used to collect real data, making comparisons fairly easy.

The model area is a cylinder 6.0 units high with a radius of 1.0. The bottom of the area, z= 0.0, is at the level of the elbow joint and the top of the area, z=6.0, is at the level of the shoulder. The main axis of the model area is a line from point (x=0,y=0,z=0) to point (x=0,y=0,z=6.0).

The basic unit of signal generation within the model is referred to as a signal unit. All signal units act as point voltage sources and have the same characteristic output waveform. The signal unit's output wave and its frequency domain magnitude plot are shown in Figure 2. All signal units are contained within muscles and their position coordinates, which are initialized at the start of a run, remain constant during a run.

Muscles are cylindrical collections of signal units. Muscles are defined by their size, position in the model area and the number of signal units that they contain. Muscles can run either parallel or perpendicular to the main axis of the model area. A muscle's position and size are defined by the start and end points of the main axis of the muscle, and by the magnitude of the radius of the muscle. The signal units that belong to a muscle are equally distributed in groups of three along the main axis. Within a group of three the units are equally spaced around the center axis at a distance of one half the radius of the muscle from the center axis.

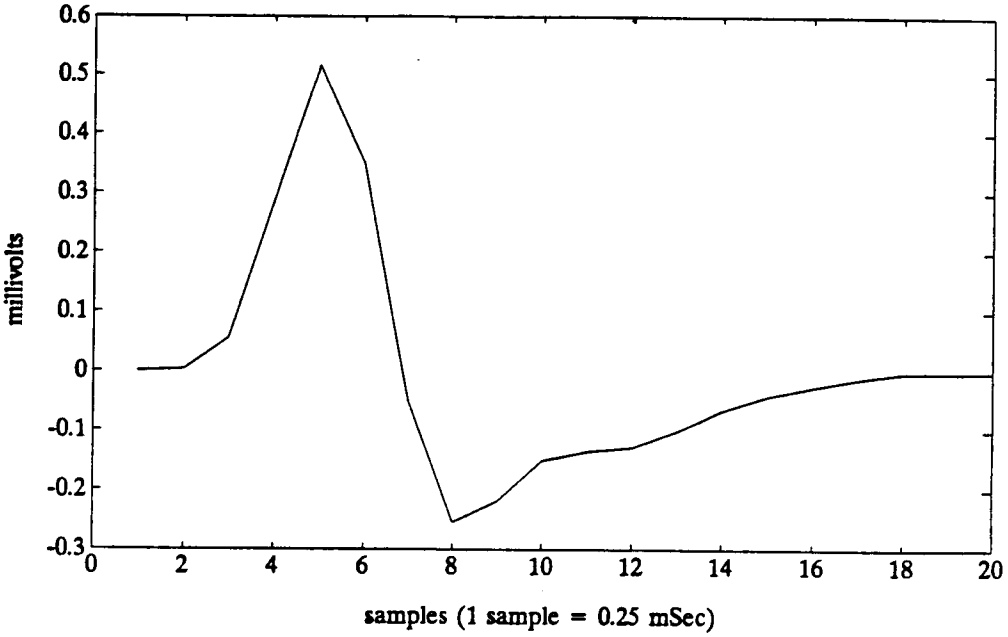


Figure 2a: Time domain plot of the Signal Unit's output waveform.

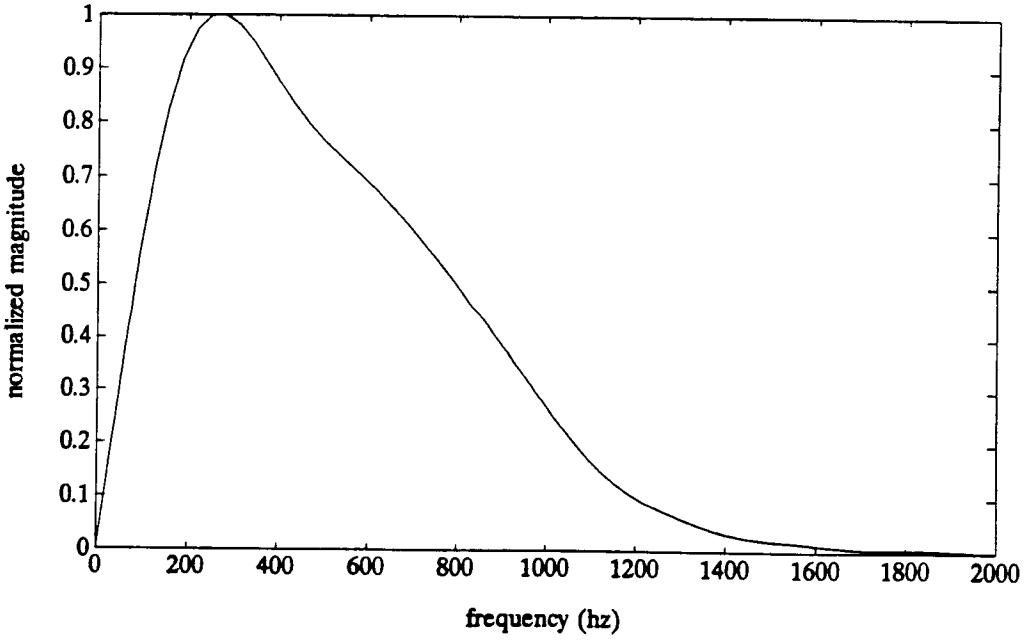


Figure 2b: Magnitude of the frequency spectrum of the Signal Unit's output waveform.

Three types of contraction are defined: 1) Regular Contraction, 2) Co-Contraction and 3) Weak Contraction. Each type of contraction is defined by a number which represents a force level of a muscle undergoing that type of contraction. For example, define Regular Contraction as 75. This would mean that a muscle undergoing Regular Contraction would deterministically fire 75% of its signal units during each time increment. (The actual number of units firing also has a random component, this is discussed later). The model has Regular Contraction = 75, Co-Contraction = 50, and Weak Contraction = 10. The relative values of these numbers are based on previous lab work done by this researcher.

Motions are defined as a list of what type of contraction each defined muscle is to perform. As an example, motion one might define muscle one in Regular Contraction, muscle two in Co-Contraction and muscles three and four in Weak Contraction. Four motions are defined in the model: 1) Flexion (bending the elbow), 2) Extension (straightening the elbow), 3) Supination (turning palm up) and 4)Pronation (turning palm down). These will be referred to as motions1-4 for simplicity.

Four muscles are defined in the model, one each as the prime mover of the four motions. They are referred to as muscles 1-4, where muscle 1 is the prime mover of motion 1 , muscle 2 is the prime mover of motion 2, etc. The defining parameters of the four muscles are listed in Table 1. A cross section of the model area is shown in Figure 3.

Table 1: Defining Parameters of Muscles

Muscle #	Center Axis Start(x,y,z)	Center Axis End(x,y,z)	Radius	# of Signal Units
1	(0.0, 0.5, 0.0)	(0.0 0.5, 6.0)	0.5	100
2	(0.0, -0.5, 1.0)	(0.0, -0.5, 6.0)	0.5	90
3	(-0.2, 0.4, 0.0)	(-0.2, 0.4, 6.0)	0.4	80
4	(-0.8, 0.0, 1.0)	(0.6, 0.0, 1.0)	0.5	70

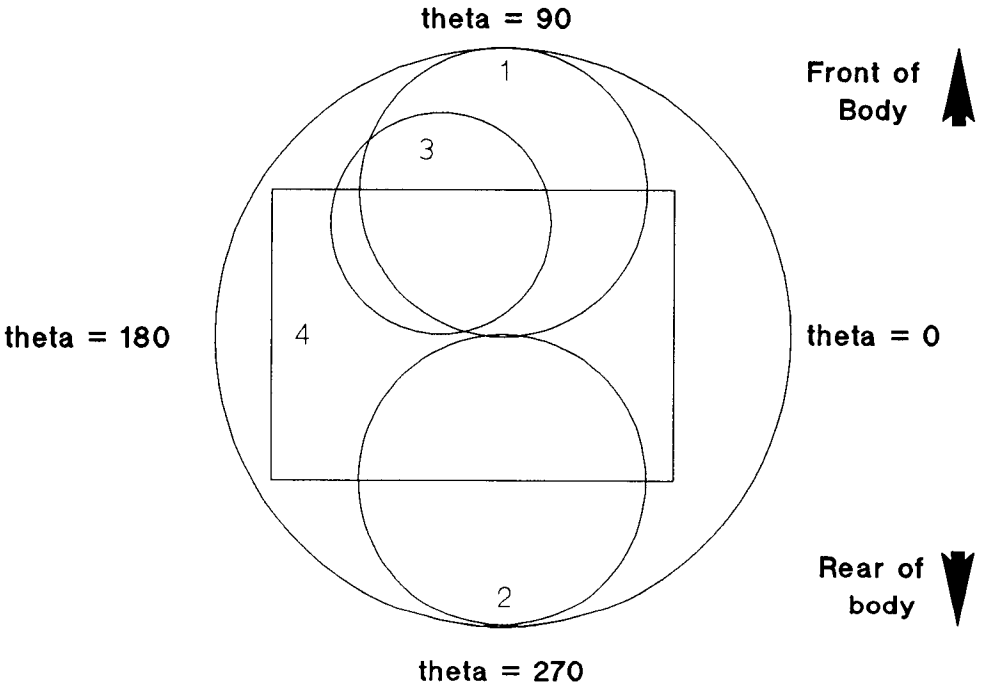


Figure 3: Top view of a cross section of the model region at z = 1.0. The numbers indicate the motion for which the muscle is the prime mover.

The effects of tissue filtering are modelled using four different, 2nd order, infinite impulse response (IIR) filters. All of the tissue filters have a 3db point at 300hz. The filters are purposely non-linear phase because the effects of tissue filtering are probably not linear phase. All of the filters have somewhat different

magnitude and phase responses since different tissue structures would have these characteristics. Which filter to use is dependent on the spatial relationship between the signal unit that is firing and the electrode position. Based on configurable thresholds in the x and y planes, the filter choice is made based on which thresholds are exceeded. For example, let the threshold in each plane be 0.4 units. If the absolute difference between the x coordinates of the signal unit and electrode is greater than 0.4 the x threshold is exceeded, with the same type of calculation for the y plane. Then, filter 1 is applied if neither threshold is exceeded, filter 2 if just x is exceeded, filter 3 if just y, and filter 4 if both thresholds are exceeded. The tissue structure is assumed to be homogeneous in the z direction with no extra filtering applied. The frequency domain characteristics of the four filters are shown in Figure 4. All filters were designed using the MATLAB signal processing software package for the PC.

The signal is attenuated by a linear function of the distance between the firing signal unit and the electrode.

Randomness is inserted into the model in two ways. First, a random component is added to the deterministic component of deciding how many units fire from each muscle at each time increment. The deterministic value is a function of what type of contraction is defined for the muscle during the motion under test. The random component is a uniform random variable that can be plus or minus 10% of the total signal units in the muscle. The second random input comes into play when deciding which units of a muscle fire at any given time increment. The actual units that fire are chosen randomly, again using a uniform random variable, from the units in the muscle.

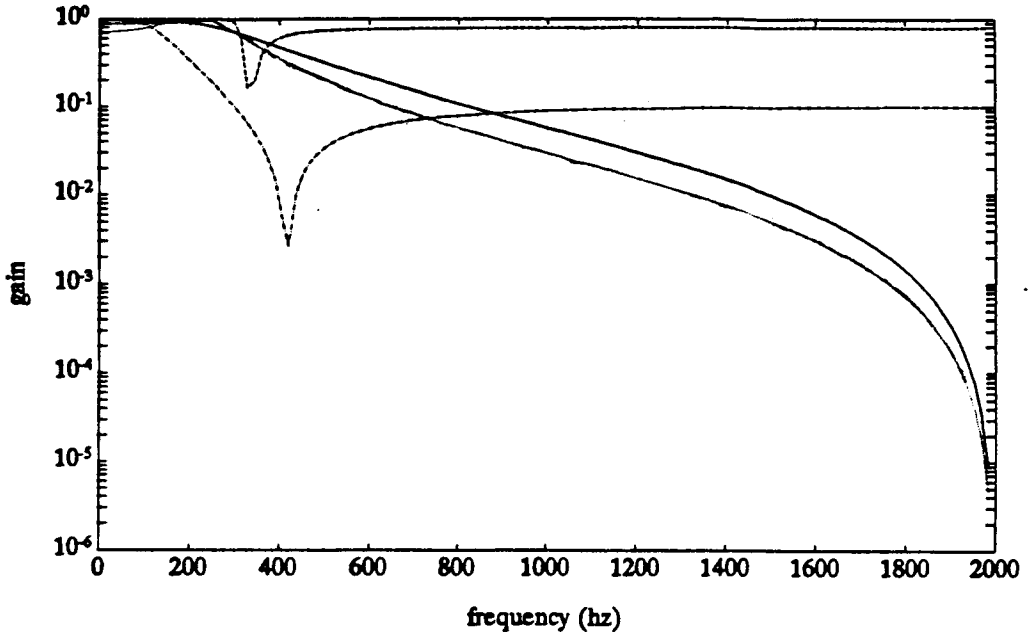


Figure 4a: Magnitude of the frequency spectrum of the four tissue filters.

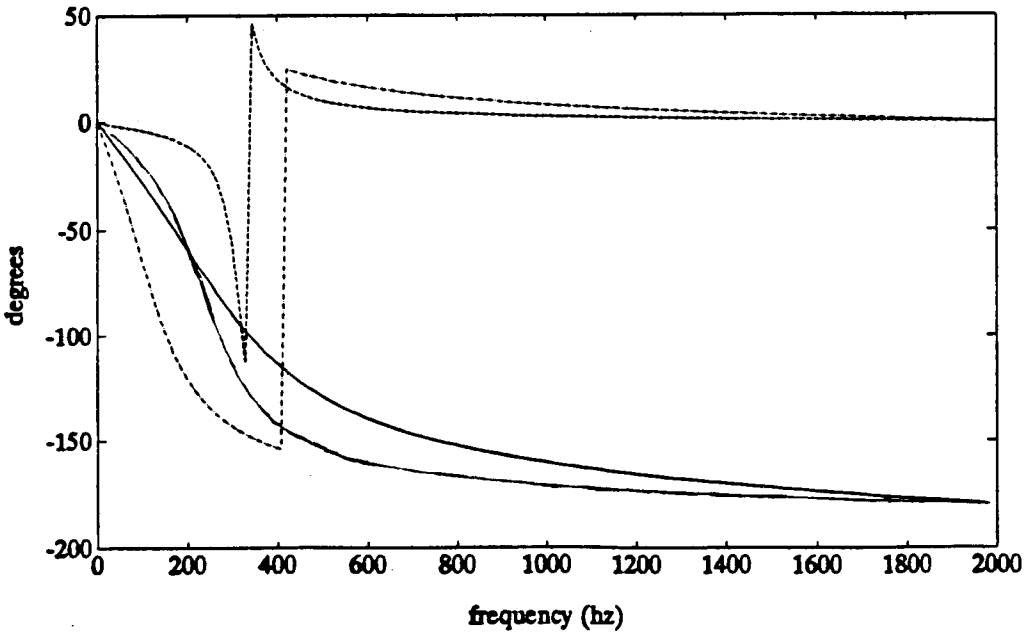


Figure 4b: Phase of the frequency spectrum for the four tissue filters.

The flow of a typical run of the model is outlined in the following psuedocode:

```
User Inputs: Electrode Position, Motion, Duration
Seed Random Number Generator, Calculate number of time increments.
FOR each time increment
BEGIN
        FOR each muscle
        BEGIN
                Calculate number of units that fire  // random component
                FOR each unit
                BEGIN
                        determine which unit //random decision
                        determine which filter to use
                        determine attenuation
                        add resulting waveform to output
                END
        END
END
```

The output from a run of the model is stored in a file as double precision floating points numbers in ASCII format. The output file is read by custom programs to calculate autocorrelation coefficients and autoregressive model parameters. The output file is also easily read by the MATLAB mathematics program to do model parameter estimation, frequency domain analysis and to produce hardcopy output.

An example of the output of a typical run of the model is shown in Figure 5.

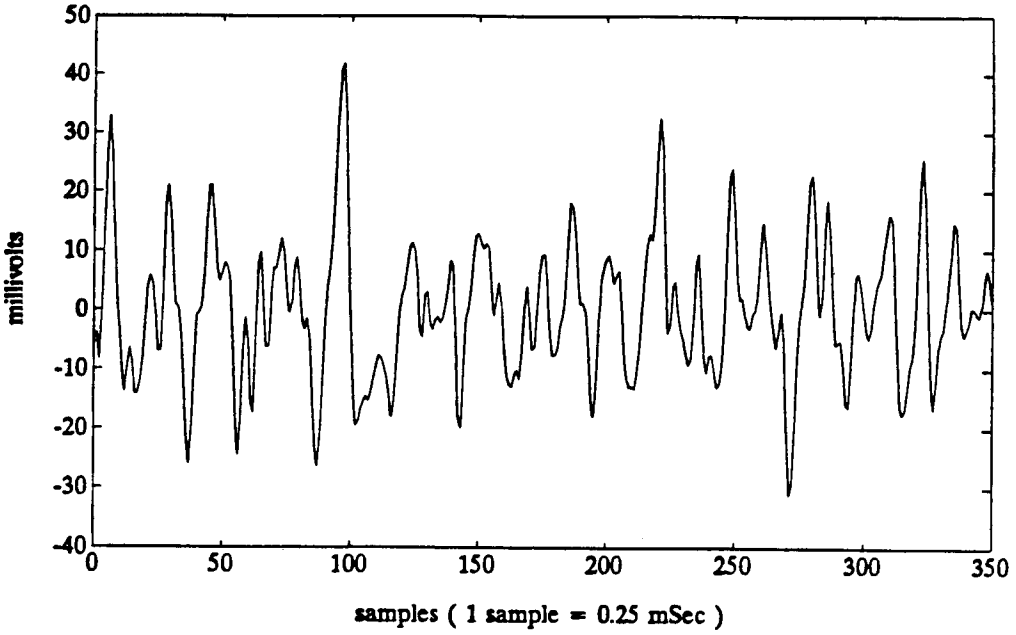


Figure 5a: Time domain plot of a sample of Model output.
Motion = Flexion(1).

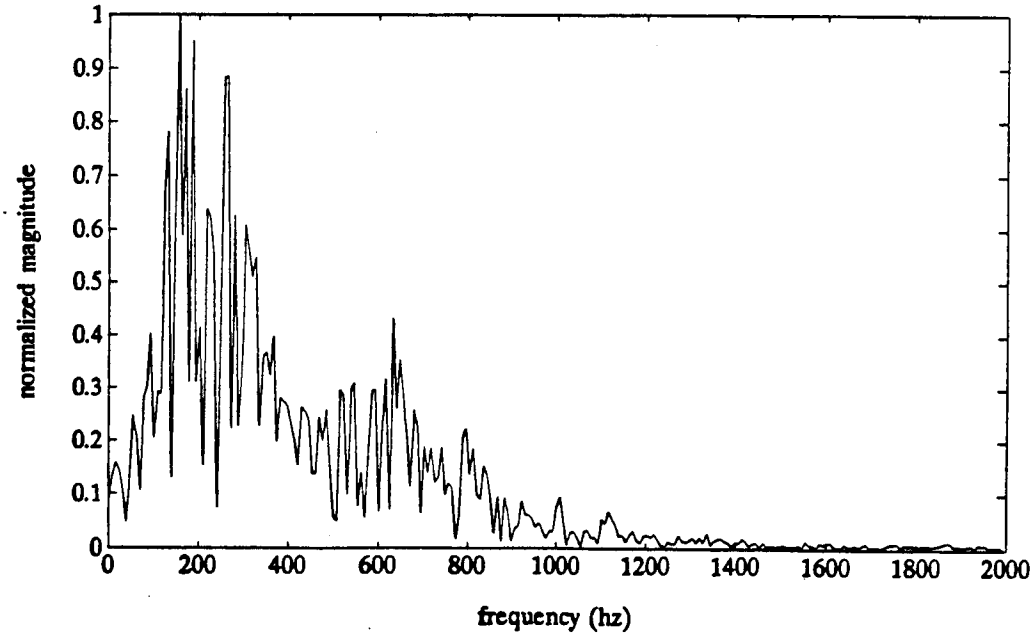


Figure 5b: Magnitude of the frequency spectrum for the sample of Model output plotted in figure 5a.

TESTING THE MODEL - PATTERN RECOGNITION

Now that a model has been proposed and run it must be tested somehow. The general goal of the model is to improve the quality of EMG analysis. A more specific first goal is to provide a model for developing and testing algorithms for pattern recognition of motions as the inputs to prosthetic control systems. The analysis and testing of the model will be done by comparing results of pattern analysis to published data, and to real data collected for testing purposes.

Pattern recognition is a multiple step process. The first step in a pattern recognition scheme is to identify features of the signal which can be used to determine what patterns are present in any given sample of the signal. Once that is done, the values of these features are determined for all of the patterns to be identified; this is the reference set of features. Then, the feature set of a specific instance of signal is determined and compared to the reference set to see what pattern is present.

In the example of the proposed model, the four motions, (1-4), are the patterns to be identified. The feature set representing each of the motions is determined during an offline learning session and stored as the reference set. Then, signals are generated from motions and the motion is determined from the EMG signal by fitting a feature set to the signal and comparing it to the stored reference set.

This thesis will use the system identification scheme of fitting autoregressive moving average (ARMA) models to the EMG signals as a means of determining a feature set for pattern recognition. Once an ARMA model is fit to the data, the

coefficients of the model themselves become the features upon which pattern identification is based.

ARMA modelling of EMG signals has been used as a standard method of determining a feature set for pattern recognition within the signals since Graupe first used the method successfully in the early 1970's. Although the method is sometimes reported to produce unacceptable results, Zhang [19], it is frequently tried as the first method of system identification.

Review of Box and Jenkins Time Series Analysis:

Nomenclature:

E()	Expectation
ρ	Autocorrelation coefficient
r	Estimated autocorrelation coefficient
γ	Autocovariance
c	Estimated autocovariance
θ	Autoregression coefficient
ψ	Moving average coefficient
z	Time series
a	White noise function
u	Mean of white noise function
B	Backward shift operator

The fitting of ARMA models to time series data is a subject rigorously formalized by Box and Jenkins in their book Time Series Analysis: Forecasting and Control [2]. This section will review the sections of that book pertinent to this thesis.

A time series is defined as a set of observations generated sequentially in time. The time series analysis done in the Box and Jenkins book is for discrete time series with a fixed time interval h between any two successive observations. Time in

the model is discrete with sample times, kT, being related to real time, t, as t=kT with k limited to integer values. For this thesis one increment of time is equal to 0.25msec. As a notational convention the discrete sample of the time series x at the current discrete time, t = kT with k=0, is represented as x_t. The value of x at the previous sample time, t = (k-1)T, is x_{t-1}, and the value of x at the next sample time is, t = (k+1)T, is x_{t+1}.

There are two useful models that can be employed when trying to characterize time series. The first is to model the time series as the output of a linear filter with a white noise process a_t, with mean u and variance σ_a^2, as the input. Using this model we have

$$z_t = u + a_t + \psi_1 a_{t-1} + \psi_2 a_{t-2} + \ldots\ldots \quad (1)$$

where z_t is the output at time t. The backward shift operator, B, is defined such that $Bz_t = z_{t-1}$ and $B^m z_t = z_{t-m}$. Using the backward shift operator (1) can be written as

$$z_t = u + \psi(B)a_t \quad (2)$$

where $\psi(B) = 1 + \psi_1 B + \psi_2 B^2 + \ldots$ is the transfer function of the linear filter. The sequence $\psi 1, \psi 2, \ldots$ may be finite or infinite. This model is referred to as a Moving Average model (MA). When the sequence is finite with q values the process is said to be a moving average process of order q (MA(q)).

The moving average process can be thought of as a finite impulse response (FIR) filter that transforms the white noise input to the time series output.

Another model that gets a lot of use is referred to as an autoregressive (AR) model. The AR model describes the output at any time t as a linear combination of past outputs and a white noise shock a_t. The equation for the model is

$$z_t = u + a_t + \theta_1 z_{t-1} + \theta_2 z_{t-2} + \ldots \quad (3)$$

or, with $\theta(B)$ defined as $\theta(B) = 1 - \theta_1 B - \theta_2 B^2 - \ldots$, eq. (3) can be written as

$$\theta(B) z_t = u + a_t. \quad (4)$$

When the set of autoregressive terms is finite with p members the process is said to be autoregressive of order p (AR(p)).

An autoregressive process can be thought of as an infinite impulse response (IIR) filter that transforms the white noise input into the time series output.

The partial autocorrelation coefficient (PAC) is a value used extensively by Box and Jenkins in their analysis of time series. The partial autocorrelation coefficient is defined as the pth coefficient of an AR(p) model.

The autocorrelation function (ACF) is the series of ACs for lags 0-n and the partial autocorrelation function (PACF) is the series of PACs for orders 0-n.

Although the series $\theta(B)$ and $\psi(B)$ can be infinite, the job of system identification tasks the researcher to search for a model with as few terms as possible in order to make analysis reasonable. This principle is referred to as parsimony. Attempting to be parsimonious, the series will be limited to order p for the AR process and order q for the MA process.

The mixed autoregressive moving average (ARMA) model, with AR order p and MA order q, is referred to as ARMA(p,q) and is written as

$$z_t = u + \theta_1 z_{t-1} + \theta_2 z_{t-2} + \ldots + \theta_p z_{t-p} + a_t + \psi_1 a_{t-1} + \ldots + \psi_q a_{t-q} \quad (5)$$

Assuming that the process to be modelled is stationary, with mean u, the autocovariance at lag k, γ_k, is defined as

$$\gamma_k = \mathrm{cov}[z_t, z_{t+k}] = E[(z_t - u)(z_{t+k} - u)] \tag{6}$$

and the autocovariance at lag zero, $\gamma 0$, is the variance of the process.

The autocorrelation coefficient (AC) at lag k, ρ_k, is

$$\rho_k = \frac{E[(z_t - u)(z_{t+k} - u)]}{\sqrt{E[(z_t - u)^2]E[(z_{t+k} - u)^2]}} \tag{7}$$

thus

$$\rho_k = \gamma_k / \gamma 0. \tag{8}$$

Equations 6 and 7 describe the calculation of the theoretical autocorrelation function. Since in practice we have only a finite time series of N observations, only estimates of the autocorrelations can actually be calculated. The estimated covariance at lag k, c_k, is calculated as:

$$c_k = \frac{1}{N} \sum_{t=1}^{N-k} (z_t - u)(z_{t+k} - u) \tag{9}$$

Thus the estimated autocorrelation at lag k , r_k, is:

$$r_k = \frac{c_0}{c_k} \tag{10}$$

Of practical importance to this thesis is the calculation of the autoregressive parameters of a model from the autocorrelation coefficients. To derive the necessary equations start from an AR(p) process, assume that the mean of the white noise process is zero for simplicity, and multiply by z_{t-k} to get:

$$z_t z_{t-k} = a_t z_{t-k} + \theta_1 z_{t-1} z_{t-k} + \theta_2 z_{t-2} z_{t-k} + \ldots + \theta_p z_{t-p} z_{t-k} \tag{11}$$

Taking the expectations of eq. 11 the result is

$$\gamma_k = \theta_1\gamma_{k-1} + \theta_2\gamma_{k-2} + \ldots + \theta_p\gamma_{k-p} \qquad k>0 \tag{12}$$

The first term, $E[a_t z_{t-k}]$, is zero for $k>0$, since z_{t-k} involves white noise inputs up to time t-k which are uncorrelated with a_t. Dividing eq. 12 by γ_0 yields the difference equation

$$\rho_k = \theta_1\rho_{k-1} + \theta_2\rho_{k-2} + \ldots + \theta_p\rho_{k-p} \qquad k>0 \tag{13}$$

By substituting k = 1,2,...,p into eq. 13 a set of linear equations, known as the Yule - Walker equations, are obtained.

$$\begin{aligned}
\rho_1 &= \theta_1 + \theta_2\rho_1 + \ldots + \theta_p\rho_{p-1} \\
\rho_2 &= \theta_1\rho_1 + \theta_2 + \ldots + \theta_p\rho_{p-2} \\
&\cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \\
&\cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \\
\rho_p &= \theta_1\rho_{p-1} + \theta_2\rho_{p-2} + \ldots + \theta_p
\end{aligned} \tag{14}$$

The Yule Walker equations are solved in a simple recursive program substituting the estimates for the autocorrelation coefficients for the theoretical values.

The Box and Jenkins method of fitting a model to data is based on an analysis of the autocorrelations and partial autocorrelations of the data. Their analysis has led to criteria upon which model fitting can be based. These criteria are: (In the following rules, a geometrically decaying series is referred to as a damped exponential).

1) For a pure AR(p) process the ACF is a mixture of damped exponentials and damped sine waves. The PACF will be nonzero up to order p and zero afterwards.

2) For a pure MA(q) process, the ACF function is zero for lags greater than q, while, in general, the PACF is dominated by a damped exponential term.

3) For a mixed ARMA(p,q) process with q-p <0, the entire ACF will consist of a mix of damped exponentials and sine waves. If q-p $>=0$, there will be q-p+1 initial values of the ACF that do not follow the general pattern of decaying exponentials and sine waves.

Fitting an ARMA Model to the Computer Model's Data:

The autocorrelations and partial autocorrelation coefficients for a sample of the model's output are plotted in Figure 6. The autocorrelation function decays rapidly, but does show some signs of damped sine wave or exponential behavior after lag 4. The partial autocorrelation function, on the other hand, is nonzero for orders of four and less but approaches zero very rapidly after order 4. Based on this data (considering many independent runs), an AR(4) model is chosen to represent the data.

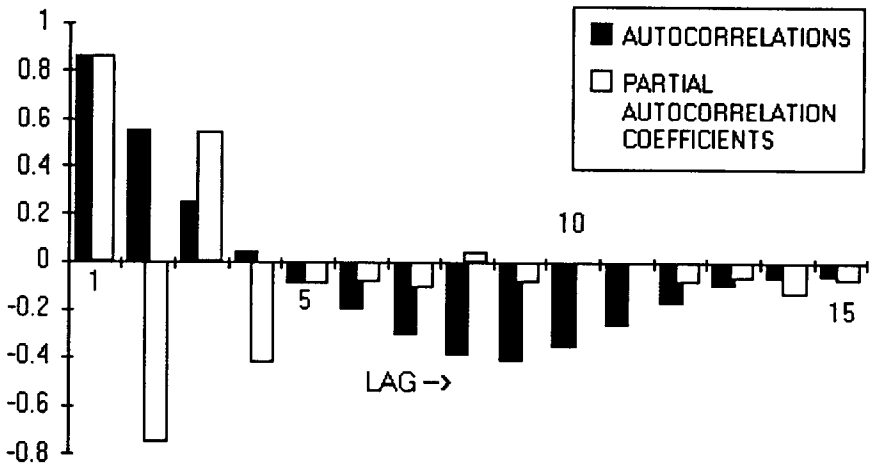


Figure 6: Autocorrelations and Partial Autocorrelation Coefficients for a sample of the Model's output data

Eq. (4) lends itself to determining a test of the choice of a pure AR model. If a pure AR model is correct, and if the order of the model is correct, the error between the actual data and calculated data based on the model should be white noise. White noise is characterized as being completely uncorrelated noise, therefore, the autocorrelation coefficients for any lag greater than zero of a white noise signal should be zero. The following procedure is performed:

1) Assume the process is best modelled as a pure AR process and fit pure AR models to the data of orders 1 - 10.

2) For each order of AR model, calculate the expected signal based on the autoregressive parameters.

3) For each order of AR model, calculate an error signal which is the difference between the actual and expected signals.

4) Calculate the autocorrelation coefficients of the error signals.

The results of this procedure are given in Table 2. The data shows that with a model of order 4 the first autocorrelation coefficient gets close to zero and that higher order models do not significantly improve the results. It can be concluded that the choice of an AR(4) model is correct and parsimonious.

Table 2: Autocorrelation Coefficients of Error Function - Model Data

AR Model Order	AC1	AC2	AC3
1	0.6616	0.0416	0.3067
2	0.4512	0.2693	0.3633
3	0.3218	0.19	0.07721
4	0.02727	0.1022	0.002037
5	0.02631	0.0942	0.002011
6	0.01272	0.016	0.0665
7	0.03239	0.0211	0.02181
8	0.03192	0.02635	0.02274
9	0.02679	0.0218	0.0199
10	0.02639	0.0239	0.01583

Pattern Recognition of Model Data:

When used as the input to controlling a prosthetic device, pattern recognition decisions must be made quickly for acceptable performance of the prosthesis. To allow time for on line processing and still achieve decent results, 0.1 seconds of data is used as the duration of the time signal to be analyzed.

The pattern recognition procedure works as follows:

1) Run the model 20 times for each motion at each electrode position. Calculate the AR4 parameters for each run and store the average of these parameters for the twenty runs as the reference set of features for each motion at each electrode position.

2) Run the model 10 times for each motion for each electrode position to generate the test set of data. For each of these test runs calculate the AR(4) parameters by doing a least squared error fit to the reference set for that electrode position. Determine if the recognized motion is the one actually performed.

Figure 7 shows the results on a cross section of the model region. The location of the results on the figure represents the electrode position. All results are at the level $z=2$ since it is found that changes in the z coordinate had little effect on the results. The figure gives a list of the percentage of correct identifications for each motion at each position.

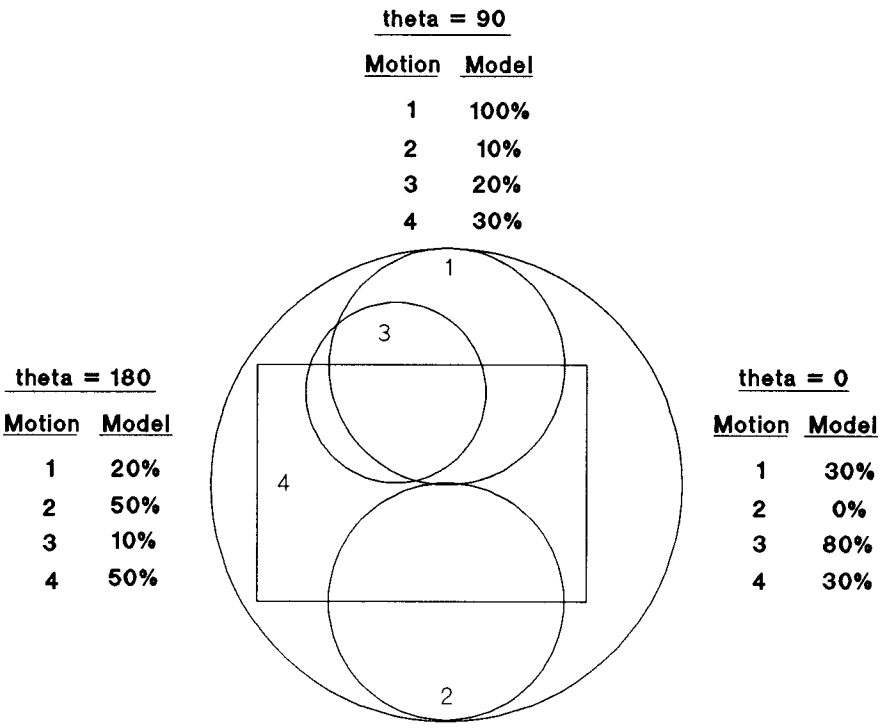


Figure 7: Results of pattern recognition on Model output superimposed on a cross section of the model region. The numbers indicate the percentage of correct pattern identifications for each pattern at each position.

Discussion of Results:

The results are not "good" in that they are certainly not of high enough quality to build a control system with this pattern identification as the input. There is a clear tendency to be able to predict a motion if the electrode is directly over the motion's prime mover, but if this "prime" motion is not the motion to be identified

the results of identification are poor. A comparison of these results to those found with real data are left until the next section.

One outcome that is not too surprising is that the results are effectively unchanged at electrode positions that differ only in the z direction. I attribute this to the fact that the effect of tissue filtering was not modelled in the z plane indicating that modelling the filtering effects of intervening tissues is a key component to the overall model.

ANALYSIS OF REAL DATA

Data is collected from a person for comparison to the model. The data is collected using a Hewlett-Packard XLi cardiograph running custom software which samples data at 4000hz. The data is high pass filtered at 40hz to remove unwanted body noise such as electrocardiographic signals. The high pass filter is a two pole, zero phase filter designed using the MATLAB software package.

The subject used to collect the data is normal and healthy. The electrodes used are Hewlett-Packard surface electrodes. The sites are the same as those used for model data, at four equally spaced locations around the upper arm approximately one third the distance from the elbow to the shoulder.

A sample of real data, and its frequency domain magnitude plot, is shown in Figure 8.

The processing of real data is the same as is done to the model's data. The first step is to determine the order of ARMA model to be used in the pattern recognition scheme. The autocorrelation coefficients and partial autocorrelations are plotted in Figure 9. The decaying nature of the autocorrelations and the sudden drop of the partial autocorrelations after lag 4 indicate an AR(4) model is appropriate. The autocorrelations of the error signals for several orders of AR models are shown in Table 3. This data indicates that an AR(3) or AR(4) model is satisfactory, an AR(4) model is used in the subsequent pattern identification.

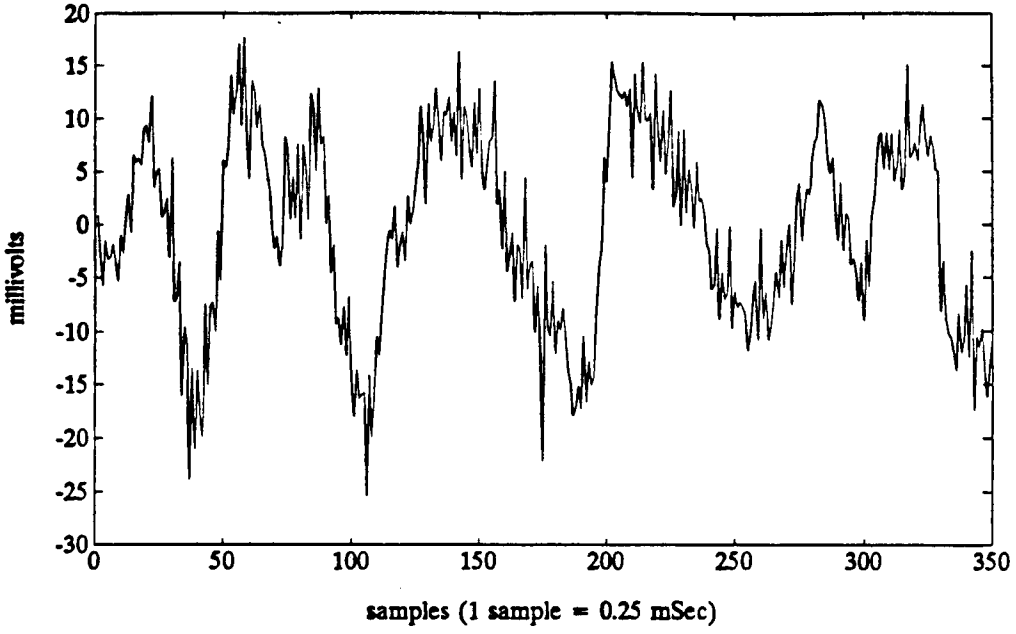


Figure 8a: Time domain plot of a sample of real EMG data.
Motion = Flexion(1).

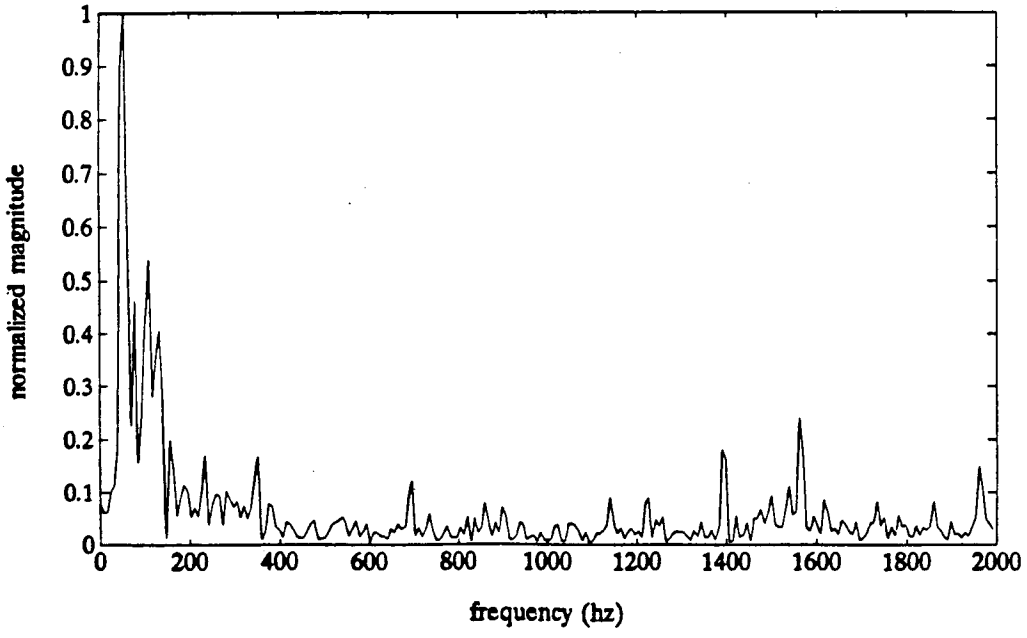


Figure 8b: Magnitude of the frequency spectrum for the sample of real EMG data plotted in figure 8a.

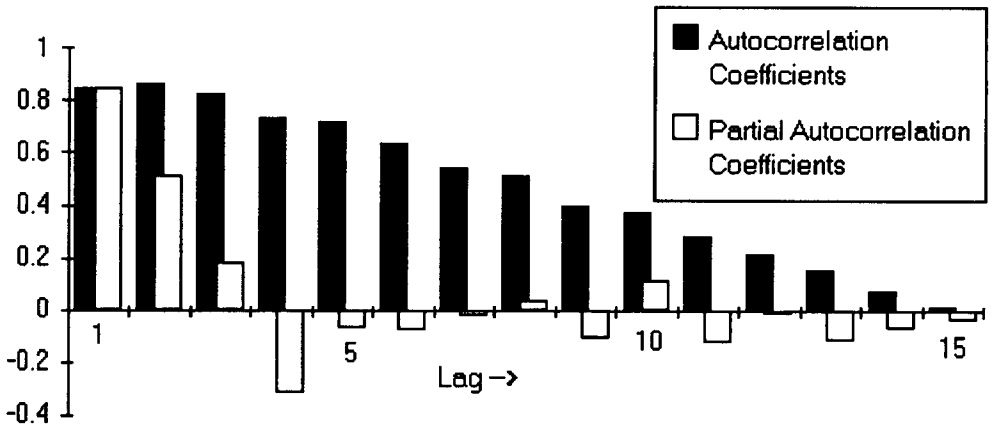


Figure 9: Autocorrelations and Partial Autocorrelation Coefficients for a sample of real EMG data

Table 3: Autocorrelation Coefficients of Error Function - Real Data

AR Model Order	AC1	AC2	AC3
1	0.436	0.224	0.26
2	0.103	0.0685	0.321
3	0.025	0.173	0.262
4	0.0428	0.00437	0.0181
5	0.0313	0.0213	0.0345
6	0.044	0.00081	0.0484
7	0.0243	0.0193	0.0539
8	0.0273	0.0241	0.057
9	0.0146	0.0149	0.0586
10	0.00887	0.00951	0.0442

The results of the pattern identification process are shown in Table 4. The values are the percentages of correct motion identifications for each motion at each position.

Table 4: Results of Pattern Identification for real data ($r = 1.0, \theta = ?, z = 2$)

Motion	theta = 0	theta = 90	theta = 180	theta = 270
1	0%	70%	30%	50%
2	30%	20%	70%	80%
3	50%	0%	40%	30%
4	30%	20%	90%	0%

COMPARISON OF RESULTS - MODEL VERSUS REAL DATA

Figure 10 shows the two sets of results, model and real data, on the cross sectional picture of the model area.

The first comparison that can be made is that the model is not a good predictor of the absolute results from the real data. The numbers from the model are both larger, the same as and smaller than from the real data leaving no good predictive conclusions as to absolute results.

A second comparison is more promising. For all but one case, motion 1 at $\theta = 90^{o}$, when the model predicts there is a 50% or more probability for success in identifying a specific motion at a specific location, the experimental results show a 50% or more success rate. The converse is also true. When the model predicts less than a 50% probability of success there is less than a 50% rate of success. This result encompasses the finding noted during model analysis; a motion is best detected by an electrode over the prime mover of the motion, but at this electrode position other motions are poorly identified.

A relevant question now is why are the results from the real data so different from those of Graupe, who reported an 85% success rate of motion identification. One plausible explanation is that Graupe did his work with actual amputees, and with one amputee in particular who had a lot of muscle and nerve damage in the location of the sensing electrode. (Graupe actually did a lot of his work of single electrode site pattern identification to deal with the real problem of having limited possible sites because of muscle and nerve damage). It is very possible that the anatomy and physiology of an amputee, especially one with muscle and or nerve damage, is different enough from those of a normal person to explain a large

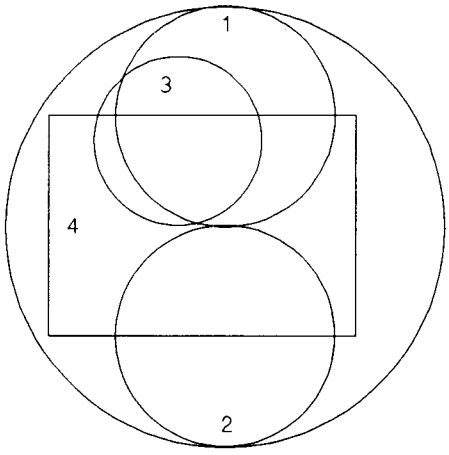
difference in results between the two populations. This explanation may be circumstantially corroborated by Zhang who reported unfavorable results using ARMA modelling. Zhang's research attempted to determine ligament damage by changes in EMG during ambulation, therefore, his test population would be considered minimally injured compared to a population of amputees.

theta = 90

Motion	Model	Real
1	100%	70%
2	10%	20%
3	20%	0%
4	30%	20%

theta = 180

Motion	Model	Real
1	20%	30%
2	50%	70%
3	10%	40%
4	50%	90%



theta = 0

Motion	Model	Real
1	30%	0%
2	0%	30%
3	80%	50%
4	30%	30%

theta = 270

Motion	Model	Real
1	0%	50%
2	100%	80%
3	30%	30%
4	10%	0%

Figure 10: Result of Pattern recognition for Model output and real EMG data.The numbers indicate the percentage of correct pattern identifications for each pattern at each position.

CONCLUSIONS AND RECOMMENDATIONS

It is concluded that the computer model for the generation of EMG signals presented in this thesis is adequate for predicting some general trends that will be found in the EMG of normal subjects. The model is not accurate enough to make precise quantitative predictions nor is it sophisticated enough to model any states of injury or disease at this time.

The results do show some promise and also point to areas where the model can be improved. The fact that the model does not change results with electrode position changes in the z plane, and that tissue filtering is assumed homogeneous in that plane, indicates that modelling the effects of the non homogeneous body tissues filtering EMG signals on the way to the body surface is vital to an accurate model. The simple attempt at modelling tissue filtering found in this model is not adequate to get accurate quantitative data from the model. The filters should have been designed such that the frequency content of model output more closely matched the frequency content of the real EMG data.

Another system in the model that is probably too simple to allow very accurate results is the ability to define muscles. Not all muscles run either parallel or perpendicular to any axis in the body and the ability to place muscles anywhere within the model area would be an improvement. Also, not all muscles are cylindrical. The ability to define more complex muscle geometries should also be added to the model.

The computer program that is the model should also have an improved user interface. Currently very few of the system parameters can be changed by the user when running the program. To be useful to researchers most if not all of the

parameters that define the system must be able to be set by the user when using the program.

The systems required to produce motion in a person are numerous and complex and have evolved over millions of years. The author has not completely modelled all of these systems with just several months of work, and in fact, it could take one or possibly many lifetimes to achieve an extremely accurate model. This thesis shows some promising results by building a first pass model that recognizes the existence of, and ties together into an integrated system, the most important physiological aspects of the generation of EMG signals. It is hoped that the structure of the model will lend itself to achieving better approximations of reality in the future.

BIBLIOGRAPHY

[1] Basmajian, J.V. Muscles Alive, Fourth Edition. Baltimore: Williams and Wilkins, 1979.

[2] G.E.M. Box and G.M. Jenkins, Time Series Analysis, Forecasting and Control. San Francisco, Calif.: Holden Day, 1976.

[3] Carlo J. DeLuca, "Physiology and Mathematics of Myoelectric Signals," IEEE Trans. Biomed. Engr., vol.BME-26, No.6, June 1979.

[4] G.C.DeAngeles, L.D.Gilmore, and C.J.DeLuca, "Standardized Evaluation of Techniques for Measuring the Spectral Compression of the Myoelectric Signal," IEEE Trans. Biomed. Engr., vol.BME-37, No.9, pp. 844-849., September 1990.

[5] P.Doerschuk, D.Gustafson, A.Willsky, "Upper Extremity Limb Function Dsicrimination Using EMG Signal Analysis," IEEE Trans. Biomed. Engr., vol.BME-30, pp.18-28, January 1983.

[6] P.Doerschuk, D.Gustafson, W.Jarisch, V.Jain, "Microprocessor Based Prosthetic Control," Final Report(April 1,1979 - July 31,1981), Prepared for the National Science Foundation, Report No.-NSF/ECS-82004, 1982.

[7] D.Graupe and W.K.Cline, "Functional Seperation of EMG Signals Via ARMA Identification Methods For Prosthesis Control Purposes," IEEE Trans. Syst., Man, Cybern., vol. SMC-5, pp.252-259, Mar. 1975.

[8] D.Graupe, J.Magnussen, and A.A.Beex, "A Microprocessor System For Multifunctional Control Of Upper-Limb Prostheses Via Myoelectric Signal Identification," IEEE Trans. Automat. Contr., vol. AC-23, pp. 538-544, Aug.1978.

[9] R.S.LeFever, and C.J.DeLuca, "A Procedure for Decomposing the Myoelectric Signal Into Its Constituent Action Potentials - Part 1: Technique, Theory, and Implementation," IEEE Trans. Biomed. Engr., vol. BME-29, pp.149-156, March 1982.

[10] N.Hogan, R.Mann, "Myoelectric Signal Processing: Optimal Estimation Applied To Electromyography-Part 1: Derivation of the Optimal Myoprocessor," IEEE Trans. Biomed. Engr., vol.BME-27, pp. 382-395, July 1980.

[11] Ibid- Part2:Demonstration Of Optimal Myoprocessor Performance," pp.396-410.

[12] S.Lee and G.N.Saridis, "The Control Of A Prosthetic Arm By EMG Pattern Recognition," IEEE Trans. on Automat.Contr., vol.AC-29, No.4, April 1984.

[13] Z.S.Pan, Y.Zhang, and P.A.Parker, "Motor Unit Power Spectrum and Firing Rate," Med. Biol. Eng. Comp., vol:27, pp. 14-18, Jan. 1989.

[14] G.N.Saridis and M.A.Gootee, "EMG Pattern Analysis And Classification For A Prosthetic Arm," IEEE Trans. Biomed. Engr., vol.BME-29, pp. 403-409, June 1982.

[15] R.J.Triolo, D.H.Nash, and G.D. Moskowitz, " The Identification of Time Series Models of Lower Extremity EMG for the Control of Prostheses Using Box-Jenkins Criteria," IEEE Trans. Biomed. Engr., vol.BME-35, pp. 584-594, August 1988.

[16] A.J.Vander, J.H.Sherman, D.S.Luciano, Human Physiology - The Mechanisms Of Body Function, Second Edition, New York: McGraw-Hill, 1970.

[17] M.Whitmarsh, D.Lerman, Development Of An EMG Signal Processor, An unpublished report prepared for Professors J.Saugen PhD. and L.Jensen, Department of Electrical and Computer Engineering, Oregon State University, 1985.

[18] Y.T.Zhang, P.A.Parker, and R.N.Scott, "Study of the effects of motor unit recruitment and firing statistics on the signal-to-noise ration of a myoelectric control channel," Med. Biol. Eng. Comp., vol:28, pp.225-231, May 1990.

[19] L.-Q.Zhang, R.Shiavi, M.A.Hunt, and J.-J. J.Chen, "Clustering Analysis and Pattern Discrimination of EMG Linear Envelopes," IEEE Trans. Biomed. Engr., vol. 38, pp.777-784, August 1991.